Decarbonization, Irrigation, and Energy System Planning: Analyses in New York State and Ethiopia

Terence M. Conlon

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2023

Abstract

Decarbonization, Irrigation, and Energy System Planning:
Analyses in New York State and Ethiopia

Terence M. Conlon

This dissertation contains two collections of analyses, both broadly focused on energy system planning, but motivated by different research objectives in distinct geographic settings. Part I – Chapters I-III – evaluates decarbonization strategies in New York. These studies are characteristic of the primary energy-related challenge faced by the Global North: How can states cost-effectively meet time-bound emissions reduction targets? A series of linear programs are developed to answer this question, culminating in the System Electrification and Capacity TRansition (SECTR) model, a high-fidelity representation of the New York State energy system that characterizes statewide emissions and allows for comparative study of various decarbonization pathways. SECTR simulations indicate that prioritizing heating and vehicle electrification alongside an expansion of instate wind and solar generation capacity allows New York to meet recently legislated climate goals more affordably than through approaches that mandate substantial low-carbon electricity targets. Additional work also explores the optimal distribution of energy infrastructure within New York to meet specified decarbonization targets, along with the value of supply-side, demand-side, and bidirectional methods of system flexibility.

Part II of this dissertation – Chapters IV-VII – is concerned with the energy system challenges faced by the lowest income countries. Set in the Ethiopian Highlands, this work first aims to locate smallholder irrigated areas, as irrigation has attendant energy requirements that

are larger and more likely to generate supplementary sources of revenue compared to residential demands. Here, a novel classification methodology is developed to collect labeled data, train a machine learning-based irrigation detection model, and understand the spatial extent of model applicability. Across isolated plots of land as small as 30m by 30m, the resulting model achieves >95% prediction accuracy. Further studies then explore the system planning implications of simulated electricity demands associated with these irrigated areas.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I owe a great debt of gratitude to many people for supporting me throughout graduate school.

To my advisor, Professor Vijay Modi: None of this work would have been possible without your expertise and guidance. Over the last 6 years, your energy and enthusiasm bolstered my spirits whenever I felt discouraged. Thank you for allowing me to pursue many different problems throughout my time as a graduate student, and for sharing your knowledge of and passion for all types of energy engineering.

To my research colleagues: Thank you for the time you have spent with me discussing research ideas, formulating models, pouring over paper edits, and providing insights that invariably yielded better work outputs. Mike Waite, thank you for showing me how to conduct graduate-level research, for teaching me the value of clear, concise prose, and for your steady reassurance during the most frustrating stretches of the research process. Chris Small, thank you for introducing me to the world of remote sensing, and for setting an unimpeachable standard for scientific rigor that I strived to match. Yuezi Wu, thank you for your keen attention to detail and for making sure our model outputs were in fact what we claimed them to be. To all others in the Sustainable Engineering Lab community whom I have had the fortune to work alongside – thank you for your conversations and contributions to my understanding of energy systems, data science, and sustainable engineering. To everyone at Descartes Labs, especially Rose Rustowicz, thank you for your continuous mentorship, and for showing me that it is possible to make a career as a professional scientist.

To Simone Fobi: Thank you for being such a close friend and brilliant collaborator as we have navigated graduate school together. Aspiring to keep up with you across all our classes and research projects brought the best out of me, and your thoughtfulness and unflappability gave me strength when I needed it.

To my New York friends: Thank you all for the companionship and memories over the past six years. Our days and nights together sustained me throughout my studies, and I will always appreciate our shared meals, our travels, and the countless times you have made me laugh. To all my friends scattered across the world: Thank you for the regular words of encouragement that lifted me more than you know. You all contributed substantially to this effort, whether you realize it or not.

To Breanna: Thank you for all the love, support, and time we have spent together. You have been my closest friend and a constant source of joy in my life. Your energy, intelligence, and grace inspire me to be the best possible partner for you, and I cannot wait to share future adventures, unstructured downtime, and everything in between with you in this next chapter of our lives.

To my family: It is impossible to acknowledge and thank you for all you have given me. However, I hope you recognize how everything I have achieved has been a direct result of the love and stability you have provided me. Rae, thank you for your constant positivity and for reminding me that life needs to be lived vigorously. Mom and Dad, thank you for the example you continue to set. Your character, your generosity, and empathy for others have given me a blueprint for how to conduct myself, while your support has allowed me to chase every one of my goals – including this one.

For my parents.

Part I: Energy System Modeling in New York State

Introduction

The first half of this thesis assesses decarbonization strategies in New York State (NYS) via a series of energy system models. In 2019, NYS passed the Climate Leadership and Community Protection Act, a law mandating significant, quantifiable decarbonization targets in the years 2030, 2040, and 2050 [1]. To meet these goals, the state will need to install substantial capacities of renewable generation (primarily wind and solar) and convert existing heating and vehicle technologies to electric alternatives. As the cost-effectiveness of this transition will depend on NYS-specific nuances of legacy infrastructure, energy sources, and constraints, the following chapters introduce a set of frameworks for representing the NYS electricity grid and electrifiable heating and transport loads.

While the details of each analysis differ, there are a couple common findings across chapters regarding decarbonization in NYS. First, NYS can meet near-term grid decarbonization goals – i.e., achieving low-carbon electricity percents of 50-60%, compared to a current value of 40% – without substantial increases in per-unit electricity costs; these goals can be met by installing new solar and wind capacity without any supplementary integration technologies. However, when NYS reaches 70-80% low-carbon electricity, per-unit electricity costs increase significantly. Here, battery storage to shift renewable generation to hours when it can displace gas generation causes total system costs to rise. At percents low-carbon electricity of 70-80%, all types of flexibility measures – such as battery storage, pumped hydropower, increased transmission, or flexible electric vehicle (EV) demand – allow the state to integrate wind and solar

generation more cost-effectively by reducing curtailment. Lastly, flexibility measures are found to provide the most benefits to the grid when collocated with renewable generation, as collocation allows the system to avoid bottlenecks at major transmission interfaces during hours of plentiful low-carbon generation.

Chapter I represents the culmination of these energy system modeling efforts in NYS. Building upon previously developed models, this work introduces the System Electrification and Capacity Transition (SECTR) model, a cost-minimizing linear program with high fidelity to the existing electricity system that includes anticipated electrified heating and vehicle demands. SECTR is a spatially explicit model defined by individual nodes representing geographical sub-areas within the larger region of interest. To determine the least-cost infrastructure mix in future model scenarios, decision variables representing onshore wind, offshore wind, solar, battery storage, transmission, and new gas generation technologies are assigned node-specific costs; a full set of system costs is available in Appendix A of the chapter. Once parameterized, SECTR allows the investigation of trade-offs among increased low-carbon electricity generation and electrification of heating and vehicles for decarbonization in a regional system. It is applied to NYS (SECTR-NY) to deliver insights on pathways to meet state-specific climate goals.

Independent combinations of low-carbon electricity percents (LCPs) between 40-95% and heating and vehicle electrification (HVE) rates between 0-100% allow for scenario-based comparisons of different decarbonization strategies. In a system with 18.7 GW of average electricity demand, full heating and vehicle electrification corresponds to an additional average 6.7 GW of each type of load. The study's results are broadly consistent with previously published research that deep greenhouse gas (GHG) emissions reductions require both a significant LCP

and increases in HVE. However, results indicate that equivalent emissions reductions can be achieved at lower costs to the grid by prioritizing electrification with 40-70% low-carbon electricity supply instead of aiming for complete grid decarbonization. With 60% electrification and 50% low-carbon electricity, approximately 1/3 emissions reductions can be achieved at current supply costs; with only 20% electrification, 90% low-carbon electricity is required to achieve the same emissions reductions, resulting in 43% higher grid costs.

Chapter I also provides a set of insights for a system looking to electrify liquid fuel-based energy demands and increase the amount of low-carbon electricity on the grid. One finding is that first order GHG reductions from electrification occur due to efficiency improvements, regardless of whether additional low-carbon generation is brought online to support the new demands. At percents low-carbon electricity equivalent to current levels (approximately 40%), average heating emissions per unit heat delivered are 70% lower after electrification; average vehicle emissions are 56% lower per mile traveled. Moreover, three primary cost drivers are identified for a decarbonizing system: (1) decreasing per-unit costs of existing infrastructure with increasing electrified demand, (2) increasing in-state generation costs from low-carbon sources relative to gas-based and hydropower generation, and (3) increasing integration costs at high percentages of low-carbon electricity.

Additional SECTR-NY results reveal a shift in optimal capacity expansion strategies as decarbonization progresses. First, onshore wind generation is installed due to its lowest per-unit generation costs. Next, the system expands offshore wind capacity, a more expensive wind resource that feeds into high-density downstate nodes when cheaper upstate generation options have high utilization. Beyond low-carbon electricity percents of 80%, solar generation combined

with battery storage becomes the most cost-effective method of displacing remaining fossil fuel-based electricity generation. Further analysis demonstrates that solar is the most effective technology to pair with batteries, due to its regular cycling whereby generation during the day can meet demand and charge batteries, which in turn discharge overnight. In contrast, wind generation demonstrates less predictability, and can drop off considerably for multi-day periods, particularly in the summer.

Lastly, Chapter I compares model results from SECTR-NY simulations to initial analyses presented to the New York State Climate Action Council ("NYS study"). At a high level, the two sets of decarbonization strategies largely agree, with similar energy resource capacities recommended to meet NYS's 2030 climate goals. Deviations between the two can largely be attributed to differences in wind/solar potential timeseries and historical demand data, and to SECTR-NY's particular attention to low-temperature effects on heat pump and EV performance. However, SECTR-NY simulations and the NYS study differ significantly in the calculation of GHG emissions. SECTR-NY includes upstream natural gas leakage factors in line with recent research and quantifiable GHG effects, whereas the NYS study assumes lower leakage rates. As this distinction has significant impacts on the amount of electrification needed to meet NYS GHG reduction targets, firming up GHG emissions assumptions across stakeholders in the decarbonization planning process should be an important priority going forward.

Chapter II offers a more contained study compared to Chapter I, assessing only the impact of new transmission and battery storage in achieving low-carbon electricity percents between 50% and 80%. This work introduces the Renewable Target Model (RTM), a mixed-integer linear program parameterized with NYS-specific costs and existing capacities to determine the least

cost method of achieving a user-defined percent renewable electricity – i.e., electricity from wind, solar, and hydropower generation, but not nuclear. Similar to SECTR-NY, the RTM splits the New York Independent System Operator (NYISO) control area into four nodes based on major transmission interfaces within the state. The RTM includes onshore wind, offshore wind, battery storage, transmission, and new gas generation capacities as decision variables; solar capacity is set to a fixed quantity in all model scenarios. Simulations are conducted with and without 25% EV adoption, as opposed to an increasing rate between 0% and 100%; heating electrification is not considered in this chapter.

A primary finding of the study is that New York can achieve 50% renewable energy penetration at a $52/MWh levelized cost of electricity (in line with current generation costs) with only a buildout of new generation capacity: Onshore wind (13.7 GW), offshore wind (4.1 GW), and solar photovoltaics (3 GW). The presence of grid-scale battery storage, EVs, or additional behind-the-meter solar capacity does not markedly change the model-selected generation mix. Furthermore, the storage and generation requirements needed to achieve renewable generation targets above 70% lead to a substantial rise in total investment. Between 50% and 55% targets, the computed marginal levelized cost of electricity (LCOE) for new variable renewable energy is $94/MWh, compared to $592/MWh between 75% and 80%, suggesting alternative integration measures are likely necessary at such high penetration rates.

RTM results also indicate that expanded transmission plays a minimal role in meeting renewable generation targets between 50-60%. Here, realistic characterization of transmission costs, especially those for lines connecting the New York City Metro area to nodes upstate and on Long Island, correspond to lower amounts of installed transmission compared to other

decarbonization analyses. The difference between this finding and those of other studies is attributable to the tendency for other models to underestimate transmission costs or assume deployment that does not consider that new infrastructure will handle only the marginal increases in transmission. Across all simulations, new transmission never contributes more than 2% of total system costs.

Findings from Chapter II are likely generalizable to other regional electricity systems, particularly those along the U.S. Atlantic Coast that contain high population density areas with limited local or nearby renewable resources other than offshore wind power. Moreover, this type of system geography is likely replicated in inland states with access to expensive local solar generation and cheaper, more distant onshore wind generation. RTM simulations provide a useful reference point for similar settings: To achieve 50% renewable generation, 3-4 times more inexpensive generation distant from large load pockets (onshore wind) is required compared to more expensive generation collocated with high demand (offshore wind); however, the total investments in the different generation types are roughly equal.

Similar to results from Chapter I, Chapter II also finds that system dependence on battery storage increases at higher renewable generation targets. At 65% renewable electricity, the value of energy storage in NYS is 2.5-3 times greater than at 50%. As near-term targets transition to longer-term goals, the amount of computed battery capacity increases substantially: While no battery capacity is required at 50% renewable generation, storage equal to 2 hours of average demand (18.7 GW) is installed at 65%, a quantity that increases to 16 hours of storage at 80%. Such rapidly increasing requirements indicate that some portion of storage capacity may be met by alternatives to batteries, including hydrogen storage.

Chapter III contains a study regarding integration of up to 30 GW of wind capacity in NYS. First, NYS hydropower resources are characterized using publicly available generation quantities, rated power capacities and reservoir sizes, and available streamflow data. Through optimization models that minimize gas generation in NYS, Chapter III next assesses the integration benefits of three types of system flexibility: flexible hydropower supply, flexible electricity supply and demand provided by pumped hydropower storage, and flexible demand provided by EVs. Flexibility is measured in terms of how much additional wind generation can be integrated into the NYS system with either 10 GW or 30 GW of wind capacity installed; solar capacity is set to 600 MW in all scenarios considered. To compare across the flexibility types, a novel quantity is introduced, "Potential Flexible Energy" (PFE), a metric that measures average energy throughput (in MWh/h, which reduces to MW) for each type of system flexibility.

For flexible supply, PFE equals the average amount of energy that flows into the hydropower reservoir over the analyzed time period. For flexible supply and demand, PFE equals the product of the facility's power generation capacity and a maximum capacity factor, equal to 0.45 after accounting for charging and discharging efficiencies. For flexible demand, PFE is set equal to the average hourly EV load. Flexibility types are compared for PFE = 1440 MW, a quantity equal to the average hydropower generation of all NYS hydropower resources besides the four largest which are individually characterized in the first portion of the chapter. Flexible hydropower and pumped storage power capacity is set to 3.2 GW, as this quantity multiplied by a capacity factor of 0.45 equals 1440 MW; reservoir capacity for these two types of storage is set equal to 24 times maximum generation (76.8 GWh). Here, 76.8 GWh of storage represents approximately 4 hours of average statewide load (18.7 GW). For demand flexibility, power

capacity is determined to be 6 times the average hourly EV load (8.64 GW), with the system able to meet daily load equal to 34.56 GWh (24 times the average hourly EV load) anytime between 7pm and 7am.

With 30 GW of wind capacity installed in NYS – corresponding to average uncurtailed generation of 9.4 GW in a system with average electricity demand of 18.7 GW – flexible hydropower supply increases average wind utilization by 470 MW, compared to a no-flexibility baseline (5.3 GW). In contrast, flexible supply and demand in the form of pumped hydropower storage integrates an additional 660 MW of wind generation on average, an approximate 10% increase. Accordingly, bidirectional pumped hydropower flexibility allows for increased utilization compared to unidirectional supply side flexibility. Here, the costs of 76.8 GWh of bidirectional storage are estimated to be approximately $10B, and 10% of the 30 GW of wind capacity to be approximately $4.5B. Together, these numbers indicate that pumped hydropower storage will not be a cost-effective method of integrating wind generation in NYS. In contrast, flexible hydropower supply may prove a more affordable option, for while this method of flexibility only integrates an additional 470 MW of wind power with 30 GW of installed capacity, the costs of unidirectional flexibility are much lower than bidirectional flexibility.

Flexible EV demand (roughly equivalent to the daily use of 3.4 million passenger EVs or 25% of statewide light-duty vehicles) increases wind utilization by 840 MW, both compared to the no-flexibility baselines. Compared to flexible hydropower and pumped storage, an equivalent amount of flexibility in the form of additional EV demand yields the largest increase in wind utilization, as this demand provides new sinks for wind generation during hours when it would

otherwise be curtailed. As such, we conclude that additional EV demand creates the largest amount of financial value for wind integration.

Nomenclature

Fixed variables and parameters

$A_{P_x,j}$	capital annualization rate for annualization period *P*, technology *x*, and interest rate *j* [years^{-1}]
$C_{existing-cap}$	total cost of existing transmission and generation capacity over entire analysis period [$]
C_{fuel}	total fuel cost for fossil fuel-based generation over entire analysis period [$]
$C_{generation}$	total generation cost over entire analysis period [$]
$C_{new-cap}$	total new capacity cost over entire analysis period [$]
$c_{bio,i}$	biofuel generated electricity price at node *i* [$/MWh]
$c_{existing-ramp}$	existing fossil fuel-based generation ramping cost [$/MW-h]
$c_{ff,i}$	fossil fuel price at node *i* [$/MMBTU]
$c_{hydro,i}$	hydropower generated electricity price at node *i* [$/MWh]
$c_{imp,i}$	imported electricity price at node *i* [$/MWh]
$c_{new-ramp}$	new fossil fuel-based generation ramping cost [$/MW-h]
$c_{nuc,i}$	nuclear generated electricity price at node *i* [$/MWh]
$CAP_{batt-e,i}$	battery storage energy capital cost at node *i* [$/MWh]
$CAP_{batt-p,i}$	battery storage power capital cost at node *i* [$/MW]
$CAP_{ff,i}$	new fossil fuel-based generation capital cost at node *i* [$/MW]
$CAP_{h2-e,i}$	hydrogen storage energy capital cost at node *i* [$/MWh]
$CAP_{h2-p,i}$	hydrogen storage power capital cost at node *i* [$/MW]
$CAP_{on,i}$	onshore wind power capital cost at node *i* [$/MW]
$CAP_{off,i}$	offshore wind power capital cost at node *i* [$/MW]
$CAP_{us-solar,i}$	utility-scale solar generation capital cost at node *i* [$/MW]
$CAP_{tx,ii'}$	capital cost of upgraded transmission from node *i* to adjacent node *i'* [$/MW-mi]
$D^t_{elec,i}$	existing electricity demand at node *i* and timestep *t* [MWh]
$D^t_{heat,i}$	electrified heating demand at node *i* and timestep *t* [MWh]
$D^{t,full}_{heat,i}$	full electrified heating demand at node *i* and timestep *t* [MWh]
$D^t_{veh,i}$	vehicle charging demand at node *i* and timestep *t* [MWh]
$D^{t,full}_{veh,i}$	full electric vehicle charging demand at node *i* and timestep *t* [MWh]
$D^t_{veh-fix,i}$	vehicle fixed charging demand at node *i* [MWh]
$D^{daily}_{veh,i}$	daily vehicle charging demand at node *i* [MWh]
$D^{daily}_{veh-fix,i}$	daily vehicle fixed charging demand at node *i* [MWh]
$D^{daily}_{veh-flex,i}$	daily vehicle flexible charging demand at node *i* [MWh]
$D^{daily}_{veh-tot,i}$	total daily vehicle charging demand at node *i* (fixed plus flexible) [MWh]
$d_{ii'}$	distance between node *i* and adjacent node *i'* [mi]
$EX_{cap,i}$	annual cost of maintaining existing generation capacity at node *i* [$/MW-yr]
$EX_{tx,i}$	annual cost of existing transmission at node *i* [$/MWh-yr]
F	quantity of fuel consumed [MJ]

f_{c-s}	cubic spline function
$H_{fix,i}^{t}$	fixed hydropower electricity generation at node *i* and timestep *t* [MWh]
$H_{flex,i}^{max}$	flexible hydropower maximum electricity generation at node *i* [MWh]
$H_{flex,i}^{daily}$	daily flexible hydropower generation at node *i* [MWh]
$H_{fix,i}^{monthly}$	monthly fixed hydropower generation at node *i* [MWh]
$H_{flex,i}^{monthly}$	monthly flexible hydropower generation at node *i* [MWh]
$H_{tot,i}^{monthly}$	monthly total hydropower electricity generation at node *i* (fixed plus flexible) [MWh]
$h_{veh-start}$	electric vehicle charging start time
$h_{veh-end}$	electric vehicle charging end time
$h_{veh-min}$	minimum number of hours required for full daily electric vehicle charging [hours]
I	set of all nodes in study region
i	single node in the study region
i'	node adjacent to *i*
j	interest rate
L_i^{daily}	daily biofuel generation at node *i* [MWh]
L_i^{max}	biofuel maximum generation at node *i* [MWh]
l	transmission loss rate
m	day index
N_i^t	nuclear-generated electricity at node *i* [MWh]
n_{years}	number of years in the analysis [years]
omf_{ff}	new fossil fuel-based generation fixed operations and management cost [$/MW-yr]
omf_{h2}	hydrogen storage fixed operations and management cost [$/MW-yr]
omf_{on}	onshore wind power fixed operations and management cost [$/MW-yr]
omf_{off}	offshore wind power fixed operations and management cost [$/MW-yr]
$omf_{us-solar}$	utility-scale solar power fixed operations and management cost [$/MW-yr]
$omf_{tx,ii'}$	fixed operations and management cost of upgraded transmission from node *i* to adjacent node *i'* [$/MW-yr]
omv_{ff}	new fossil fuel-based generation variable operations and management cost [$/MWh]
P	annualization period [years]
S	full set of onshore wind capacity sites
s	single onshore wind capacity site
T	total number of hourly time steps in analysis
t	hourly time step
$U_{tx-flow,i}^{existing}$	annual existing intranodal transmission flow at node *i* [MWh]
V_i^{max}	maximum hourly electricity import limit at node *i* [MWh]
$W_{off,i}^{t}$	potential offshore wind-generated electricity at node *i* and timestep *t* [$MWh_{generation}/MW_{installed}$]
$W_{on,i}^{t}$	potential onshore wind-generated electricity at node *i* and timestep *t*

	[$MWh_{generation}/MW_{installed}$]
$W^t_{on,s,i}$	potential onshore wind-generated electricity at site *s* in node *i* and timestep *t* [$MWh_{generation}/MW_{installed}$]
$W^t_{btm-solar,i}$	potential behind-the-meter solar-generated electricity at node *i* and timestep *t* [$MWh_{generation}/MW_{installed}$]
$W^t_{us-solar,i}$	potential utility-scale solar-generated electricity at node *i* and timestep *t* [$MWh_{generation}/MW_{installed}$]
$X_{btm-solar,i}$	capacity of behind-the-meter solar generation (existing and newly simulated) at node *i* [MW]
$X^{existing}_{batt-e,i}$	capacity of existing battery energy at node *i* [MWh]
$X^{existing}_{batt-p,i}$	capacity of existing battery power at node *i* [MW]
$X^{existing}_{bio,i}$	capacity of existing biofuel generation at node *i* [MW]
$X^{existing}_{cap,i}$	capacity of existing generation with associated maintenance costs at node *i* [MW]
$X^{existing}_{ff,i}$	capacity of existing fossil fuel-based generation at node *i* [MW]
$X^{existing}_{hydro,i}$	capacity of existing hydropower generation at node *i* [MW]
$X^{existing}_{off,i}$	capacity of existing offshore wind generation at node *i* [MW]
$X^{existing}_{on,i}$	capacity of existing onshore wind generation at node *i* [MW]
$X^{existing}_{nuc,i}$	capacity of existing nuclear generation at node i [MW]
X_{ps}	hydropower pumped storage power capacity [MW]
$X^{existing}_{tx,ii'}$	capacity of existing transmission between node *i* and adjacent node *i'* [MW]
$X^{existing}_{us-solar,i}$	capacity of existing utility-scale solar generation at node *i* [MW]
y	fraction
η_{batt}	one-way battery storage efficiency
η_{hydro}	one-way hydropower generation efficiency
η_{h2}	one-way hydrogen storage efficiency
$\eta_{ff-existing}$	fossil fuel-based generation efficiency of existing capacity
η_{ff-new}	fossil fuel-based generation efficiency of new capacity
η_{veh}	electric vehicle charging efficiency
ε	emissions [CO_2e]
θ	emissions rate [CO_2e/unit energy]
κ	storage self-discharge
σ	fossil fuel-based generation reserve requirement
$\varphi_{p2e-batt-min}$	minimum possible battery storage power-to-energy ratio
$\varphi_{p2e-batt-max}$	maximum possible battery storage power-to-energy ratio
$\varphi_{p2e-h2-min}$	minimum possible hydrogen storage power-to-energy ratio
$\varphi_{p2e-h2-max}$	maximum possible hydrogen storage power-to-energy ratio

Decision variables
All variables are constrained to be greater than or equal to 0.

$D^t_{veh-flex,i}$	vehicle flexible charging demand at node *i* and timestep *t* [MWh]
$H^t_{flex,i}$	flexible hydropower electricity generation at node *i* and timestep *t* [MWh]
$E^t_{batt,i}$	aggregate battery storage state of charge at node *i* and timestep *t* [MWh]
$E^t_{h2,i}$	aggregate hydrogen storage state of charge at node *i* [MWh]
$G^t_{existing,i}$	fossil fuel-based generation from existing capacity at node *i* and timestep *t* [MWh]
$G^t_{existing-diff,i}$	absolute value of the difference in fossil fuel-based generation from existing capacity at node *i* between time steps *t* and *t-1* [MWh]
$G^t_{new,i}$	fossil fuel-based generation from new capacity at node *i* and timestep *t* [MWh]
$G^t_{new-diff,i}$	absolute value of the difference in fossil fuel-based generation from new capacity at node *i* between time steps *t* and *t-1* [MWh]
L^t_i	biofuel generation at node *i* and timestep *t* [MWh]
NL^t_i	netload at node *i* and timestep *t* [MWh]
V^t_i	imported electricity at node *i* and timestep *t* [MWh]
$X_{batt-e,i}$	battery storage energy capacity installed at node *i* [MWh]
$X_{batt-p,i}$	battery storage power capacity installed at node *i* [MW]
$X_{ff,i}$	capacity of fossil fuel-based generation installed at node *i* [MW]
$X_{h2-e,i}$	hydrogen storage energy capacity installed at node *i* [MWh]
$X_{h2-p,i}$	hydrogen storage power capacity installed at node *i* [MW]
$X_{off,i}$	capacity of offshore wind generation installed at node *i* [MW]
$X_{on,i}$	capacity of onshore wind generation installed at node *i* [MW]
$x_{on,s,i}$	capacity of onshore wind generation installed at site *s* in node *i* [MW]
$X_{us-solar,i}$	capacity of utility-scale solar generation installed at node *i* [MW]
$X_{tx,ii'}$	capacity of new transmission from node *i* to adjacent node *i'* [MW]
$Z^t_{ii'}$	electricity transmitted from node *i* to adjacent node *i'* at timestep *t* [MWh]
$\gamma^t_{batt,i}$	increase in battery storage state of charge at node *i* and timestep *t* [MWh]
$\gamma^t_{h2,i}$	increase in hydrogen storage state of charge at node *i* and timestep *t* [MWh]
$\delta^t_{batt,i}$	decrease in battery storage state of charge at node *i* and timestep *t* [MWh]
$\delta^t_{h2,i}$	decrease in hydrogen storage state of charge at node *i* and timestep *t* [MWh]
ζ^t_i	renewable generation curtailment at node *i* and timestep *t* [MWh]

Scenario configuration parameters

RGT	renewable electricity generation target: Fraction of total demand that must be met by renewable energy (combined wind, water, and solar power)
LCP	low-carbon electricity generation percent: Fraction of total demand that must be met by low-carbon energy (combined nuclear, wind, water, and solar power)
PFE	potential flexible electricity

$p_{heat,i}$	fraction of full heating electrification demand simulated at node *i*
$p_{veh,i}$	fraction of full vehicle electrification demand simulated at node *i*
ω	percent reduction in total greenhouse gas emissions

Subscripts and superscripts

(Note: Some fixed variables and parameters defined above are used in subscripts and superscripts. These terms are not redefined here.)

batt	battery storage
bg	Blenheim-Gilboa
bio	biofuel
btm	behind-the-meter
diff	difference
elec	electricity
fix	fixed
ff	fossil-fuel
fh	flexible hydro
flex	flexible
gas	motor gasoline
heat	heating
h2	hydrogen storage
imp	imports
ind	industrial sector
lew	Lewiston
max	maximum
nc	new capacity
nia	Niagara
ng	natural gas
off	offshore wind
on	onshore wind
other	out-of-scope
ps	pumped storage
p2e	power-to-energy
sh	small hydro
stl	St. Lawrence
tot	total
transp	transportation sector
tx	transmission
us	utility scale
veh	vehicle

Acronyms and abbreviations

CEM	capacity expansion model
CF	capacity factor
CFM	comparative flexibility model
FHM	flexible hydropower model
GHG	greenhouse gas
HVE	heating and vehicle electrification
ISO	independent system operator
LCOE	levelized cost of electricity
NYS	New York State
NREL	National Renewable Energy Laboratory
PFE	potential flexible electricity
PV	photovoltaic
RGT	renewable generation target
RTO	regional transmission organization
SECTR	System Electrification and Capacity TRansition
SECTR-NY	System Electrification and Capacity TRansition – applied to New York State
VRE	variable renewable energy

Chapter I: Assessing trade-offs among electrification and grid decarbonization in a clean energy transition: Application to New York State

Abstract

A modeling framework is presented to investigate trade-offs among decarbonization from increased low-carbon electricity generation and electrification of heating and vehicles. The model is broadly applicable but relies on high-fidelity parameterization of existing infrastructure and anticipated electrified loads; this study applies it to New York State where detailed data is available. Trade-offs are investigated between end-use electrification and renewable energy deployment in terms of supply costs, generation and storage capacities, renewable resource mix, and system operation. Results indicate that equivalent emissions reductions can be achieved at lower costs to the grid by prioritizing electrification with 40-70% low-carbon electricity supply instead of aiming for complete grid decarbonization. With 60% electrification and 50% low-carbon electricity, approximately 1/3 emissions reductions can be achieved at current supply costs; with only 20% electrification, 90% low-carbon electricity is required to achieve the same emissions reductions, resulting in 43% higher grid costs. In addition, three primary cost drivers are identified for a system undergoing decarbonization: (1) decreasing per-unit costs of existing infrastructure with increasing electrified demand, (2) higher in-state generation costs from low-carbon sources relative to gas-based and hydropower generation, and (3) increasing integration costs at high percentages of low-carbon electricity.

1. Introduction

The United States is at a clean energy crossroads. Economically, per-unit costs of new solar and wind generation have become lower than coal and gas generation in parts of the country [2]. Policy-wise, several states have recently passed major climate legislation [3]. Public opinion mirrors these changes: A growing consensus acknowledges that a clean energy transition would have numerous social [4] and economic benefits [5]. As a result, support for sweeping federal action has reached new heights [6]. Even so, the cost-effectiveness of this transition will be influenced by region-specific nuances of legacy infrastructure, energy sources, and constraints [7]. This chapter proposes an open-source framework that offers a means to evaluate decarbonizing the electricity grid while considering electrification of heating and vehicles. The framework is then to New York State (NYS) to highlight trade-offs among dominant decarbonization options emblematic of a region with a well-defined electricity system and a variety of climates, renewable energy resources, and existing fossil fuel end-use needs.

There is widespread consensus that coupling electrification of heating and vehicles with renewable energy expansion is the best approach to reducing energy-related greenhouse gas (GHG) emissions [8]. In fact, it is infeasible to meet deep decarbonization targets without both cleaning the grid and replacing current fossil fuel transportation and heating technologies with low-carbon alternatives [9]. However, less well understood are how prioritizing fossil fuel end-use electrification or the percentage of electricity from low-carbon sources influences the cost-effectiveness of emissions reductions, electrification's potential benefits to the electricity system, and how transitioning existing heating and transportation infrastructure impacts hourly energy system operation.

Many energy system models seek to determine economically optimal technology mixes for future electricity scenarios, including those set in NYS [10]. Modeling unit commitment and dispatch [11] at the scale of individual generators [12] under varying degrees of foresight [13] can provide detailed operational understanding for a fully defined system. Capacity expansion models (CEMs) generally aggregate generators with similar characteristics in order to avoid the significant computational requirements of high spatial and temporal resolution models with capacities as decision variables [14]. The improved tractability of CEMs (often called "macro-energy system models"[15] when applied to regional systems) allows them to incorporate a larger number of system characteristics [16]. CEMs have expanded to include additional technological options, demonstrating that higher fidelity to existing systems results in more accurate capacity expansion scenarios [17]. By modeling resource stochasticity, other CEMs find that optimal system design changes under uncertainty [18]. Moreover, the inclusion of environmental considerations shifts the deployment of renewable generation capacity compared to CEMs that do not account for land use limitations [19]. CEMs that simulate interconnected energy systems such as transportation [20] and heating [21] have modeled sector-wide clean energy transitions, showing that the interplay of different energy demands is critical in understanding decarbonization pathways. Nevertheless, because characterizing actual systems can be time-consuming (if sufficient information and data is even available), CEMs often do not contain high-fidelity parameterizations of all existing system conditions [21]. These shortcomings are particularly problematic for regional energy systems (e.g. at the Regional Transmission Organization (RTO) or Independent System Operator (ISO) scale) with unique existing

infrastructure and resource mixes that are likely to affect deep decarbonization efforts, as well as intra-regional heterogeneity that may not be captured in larger-scale models [22].

While CEMs have previously been used to investigate the impact of electrified loads on least-cost model decisions, there remain opportunities for improvements in methods and applications. A group of CEM-based studies by the U.S. National Renewable Energy Laboratory (NREL) explores the effects of electrification and decarbonization on model-selected energy infrastructure capacities [23], electricity cost [24], emissions [25], variable renewable electricity (VRE) integration [26], and electricity demand curves [27] in the continental US. These NREL studies use representative time slices in place of continuous time series to solve models with high spatial resolution, but this approach precludes thorough investigation of system operation. Similarly, a recent study on achieving net-zero emissions in the continental U.S. through expanded low-carbon electricity and end-use electrification simulates power sector operations at an hourly resolution for 41 representative days [22]; as with the NREL studies, representative time slices prevent a full accounting for system operation over a continuous time period. Other studies include continuous supply and demand time series to evaluate power flow for discrete scenarios (i.e. with fixed infrastructure capacities rather than optimal capacity expansion decision-making) to evaluate the effects of electrification on VRE integration [28]. Another study of this type applies a grid model introduced in [29] to evaluate the effects of electrified heating demand in California on both GHG emissions and grid resource capacity needs. Here, resource mixes are exogenously defined, and electricity costs in future electrification scenarios are not presented [30].

Recent studies of NYS have found that deep decarbonization is feasible using existing technologies, and that different pathways exist to a carbon neutral future [31]. One such report issued by New York's Climate Action Council concludes that substantial progress on heating and vehicle electrification is required by 2030, and that nearly 100 GW of renewable generation capacity is required for full energy sector decarbonization by 2050 [32]. Related work uses a capacity expansion model and representative timeseries to show that battery storage will be required to ensure electricity reliability during a low-carbon transition [33]. However, these studies also list areas for future research, including incorporation of an updated GHG emissions assumptions accounting [31].

A gap in the literature thus remains: An evaluation of both cost-optimal capacity expansion and system operation for a well-characterized existing regional energy system, under various combinations of electrification and low-carbon electricity adoption rates, using multiple years of real data, with improved emissions assumptions. Chapter I addresses this gap by introducing an open-source System Electrification and Capacity TRansition (SECTR) modeling framework. To determine optimal system characteristics, SECTR computes the lowest total cost of electricity generation, transmission, and storage resource mix for specified combinations of: (a) low-carbon electricity supply percentage, (b) building end-use and vehicle electrification, and (c) percent GHG emissions reduction. SECTR is designed to replicate existing system characteristics: spatially heterogeneous hourly electricity demands, generation technologies, and capital and operating costs; inter-nodal transmission limits; energy storage; temperature-dependent electric vehicle charging demands; and electrified heating demand time series [34]. Agriculture and industrial emissions are included in GHG computations, but SECTR does not

endogenously model changes in those sectors. In this chapter, the SECTR framework is applied to New York State's energy system (SECTR-NY). Lastly, for the SECTR-NY application, this chapter includes an emissions accounting that improves upon the accounting contained in current NYS reports, as it incorporates methane leakage and adopts the longer duration GHG warming potentials specified by a recent state climate law.

2. Methodology

Section 2 contains a description of the SECTR model general formulation, and the motivation for its application to New York State.

2.1 System Electrification and Capacity Transition model general formulation

A SECTR model study region is defined by individual nodes, *i,* representing geographical sub-areas within the larger region of interest. Along with existing electricity demand, each node contains electrified heating[i] and vehicle charging loads at each timestep, *t*, within the overall time period simulated, *T.* To determine the least-cost infrastructure mix in future model scenarios, decision variables are assigned node-specific costs. SECTR uses a characterization of the region's energy related GHG emissions as both a reference quantity for GHG emissions reduction computations and to compute the emissions impact of reduced fossil fuel usage associated with heating and vehicle electrification; the model does not consider improved efficiency or growth of fossil fuel end-uses.

[i] Note that SECTR incorporates the ability to model shifts of any fossil fuel-based building end-use, which generally depend on heat in some form: In US residences, 93% of natural gas, 86% of propane, and 98% of fuel oil consumption is used for either space or water heating [332]; in commercial buildings, 78% of natural gas and 70% of fuel oil consumption is used for space or water heating [333]. As such, "heating" is used for short.

SECTR evaluates different low-carbon electricity supply and end-use electrification scenarios by computing the total cost of new and existing infrastructure capacity and maintenance, fuels, and resource operation to estimate the total annual cost of electricity generation and transmission; these returned costs do not include delivery expenses (primarily distribution system costs). The modeling framework does not include the cost of replacing current fossil fuel-based building systems and vehicles or electricity distribution system costs; as such, SECTR cost computations can be considered those that typically constitute the "supply" portion of a utility customer's bill.

The remainder of Section I-2.1 contains a subset of the SECTR governing equations that establish the model configuration, along with additional equations that define how costs and emissions are calculated. Section I-2.2 contains SECTR equations not used in any of the presented analysis; these constraints allow for additional model functionality in future applications.

Objective function

SECTR's objective function minimizes the total annual electricity system supply cost based on specification of two of the following three configuration parameters: (1) minimum percent of in-state electricity generated from low-carbon resources, LCP; 2) minimum percent electrification of current fossil fuel-based heating, p_{heat}, and vehicle electrification, p_{veh}; and (3) minimum GHG emissions reduction requirement, ω. Eqs. (I-1) – (I-4) describe the objective function, where $C_{new-cap}$ is the total cost of new capacity, $C_{generation}$is the total cost of generation, and $C_{existing-cap}$ is the total cost of maintaining existing capacity:

$$obj = minimize\left(C_{new-cap} + C_{generation} + C_{existing-cap}\right)$$

(I-1)

$$C_{new-cap} = n_{years} * \sum_{i \in I} \Big[(A_{P_{on},j} * CAP_{on,i} + omf_{on}) * X_{on,i} + (A_{P_{off},j} * CAP_{off,i} + omf_{off}) * X_{off,i} + (A_{P_{us-solar},j} * CAP_{us-solar,i} + omf_{us-solar}) * X_{us-solar,i} + (A_{P_{batt},j} * CAP_{batt-e,i}) * X_{batt-e,i} + (A_{P_{batt},j} * CAP_{batt-p,i}) * X_{batt-p,i} + (A_{P_{ff},j} * CAP_{ff,i} + omf_{ff}) * X_{ff,i} + \sum_{i'} (A_{P_{tx},j} * CAP_{tx,ii'} * d_{ii'} + omf_{tx,ii'}) * X_{tx,ii'} \Big]$$

(I-2)

$$C_{generation} = \sum_{i \in I} \sum_{t \in T} \Big[c_{hydro,i} * \left(H^t_{fixed,i} + H^t_{flex,i} \right) + c_{nuc,i} * N^t_i + c_{bio,i} * L^t_i + c_{imp,i} * V^t_i + 3.412 * c_{ff,i} * \left(\frac{G^t_{existing,i}}{\eta_{ff-existing}} + \frac{G^t_{new,i}}{\eta_{ff-new}} \right) + omv_{ff} * G^t_{new,i} + c_{existing-ramp} * G^t_{existing-diff,i} + c_{new-ramp} * G^t_{new-diff,i} \Big]$$

(I-3)

$$C_{existing-cap} = n_{years} * \sum_{i \in I} \left[EX_{cap,i} * X^{existing}_{cap,i} + EX_{tx,i} * U^{existing}_{tx-flow,i} \right]$$

(I-4)

Levelized cost of electricity calculations

The levelized cost of electricity (LCOE) is calculated per Eq. (I-5):

$$LCOE = \frac{C_{new-cap} + C_{generation} + C_{existing-cap}}{\sum_{t \in T} \sum_{i \in I} \left[\mathrm{D}^t_{elec,i} + \mathrm{D}^t_{heat,i} + \mathrm{D}^t_{veh,i} - X_{btm-solar,i} * W^t_{btm-solar,i} \right]}$$

(I-5)

Note that the LCOE is simply the total electricity supply cost divided by the total electricity demand, after subtracting contributions from behind-the-meter (BTM) solar generation. LCOE is used as a general comparative metric between scenarios.

Capital cost annualization

For a given technology, *x*, the annualization rate ($A_{P_x,j}$) associated with the capacity cost, CAP_x, is computed from a technology-specific annualization period (P_x) and a 5% interest rate (j), per Eq. (I-6).

$$A_{P_x,j} = \frac{j * (1+j)^{P_x}}{((1+j)^{P_x} - 1)} \tag{I-6}$$

Heating and vehicle electrification

Hourly demands for electrified heating, $D^t_{heat,i}$, are based on the nodal percentage of heating electrification, $p_{heat,i}$, and user-provided nodal electricity demands for full heating electrification, $D^{t,full}_{heat,i}$, per Eq. (I-7).

$$D^t_{heat,i} = p_{heat,i} * D^{t,full}_{heat,i} \tag{I-7}$$

Electric vehicle demand at each time step, $D^t_{veh,i}$, is based on the nodal percentage of vehicle electrification, $p_{veh,i}$, and user-provided nodal electricity demands for full vehicle electrification, $D^{t,full}_{veh,i}$, per Eq. (I-8).

$$D^t_{veh,i} = p_{veh,i} * D^{t,full}_{veh,i} \tag{I-8}$$

Energy balance constraint

The nodal energy balance is constrained by the following inequality, with all variables defined in the Nomenclature:

$$(X_{on,i} + X_{on,i}^{existing}) * W_{on,i}^{t} + (X_{off,i} + X_{off,i}^{existing}) * W_{off,i}^{t} + (X_{us-solar,i} + X_{us-solar,i}^{existing}) * W_{us-solar,i}^{t} + X_{btm-solar,i} * W_{btm-solar,i}^{t} + H_{flex,i}^{t} + H_{fixed,i}^{t} + N_{i}^{t} + G_{existing,i}^{t} + G_{new,i}^{t} + L_{i}^{t} + V_{i}^{t} - \gamma_{batt,i}^{t} + \delta_{batt,i}^{t} + \sum_{i'} [(1-l) * Z_{i'i}^{t} - Z_{ii'}^{t}] \geq \mathrm{D}_{elec,i}^{t} + \mathrm{D}_{heat,i}^{t} + D_{veh,i}^{t} \quad \text{(I-9)}$$

The low-carbon electricity generation curtailment is computed from the slack in this constraint at each node.

Low-carbon electricity generation targets

For certain SECTR configurations, the user selects a low-carbon percent (LCP) – a minimum percentage of in-state electricity supply from onshore and offshore wind, hydropower, solar, and nuclear power after subtracting out contributions from BTM generation; the electricity generated from fossil fuels and biofuels over the full simulation period is thus constrained per Eq. (I-10).

$$\sum_{t\in T}\sum_{i\in I}\left(G_{existing,i}^{t} + G_{new,i}^{t} + L_{i}^{t}\right) \leq (1-LCP) * \sum_{t\in T}\sum_{i\in I}\left[\mathrm{D}_{elec,i}^{t} + \mathrm{D}_{heat,i}^{t} + \mathrm{D}_{veh,i}^{t} - V_{i}^{t} - X_{btm-solar,i} * W_{btm-solar,i}^{t}\right] \quad \text{(I-10)}$$

Emission reduction calculations and assumptions

In-region electricity generation emissions are calculated with emissions rate of fossil fuel-based generation, θ_{ff}, and generation from existing, $G_{existing,i}^{t}$, and new, $G_{new,i}^{t}$, fossil fuel plants, after accounting for their respective efficiencies, $\eta_{ff-existing}$ and η_{ff-new}. Emissions from imported electricity are determined by the product of the emissions rate of imports, $\theta_{imp,i}$,

and the quantity of imports, V_i^t. Together, emissions from in-region generated electricity and imports are summed over all nodes *i* and timesteps *t* to compute total electricity related GHG emissions, ε_{elec}, for each scenario, per Eq. (I-11).

$$\varepsilon_{elec} = \sum_{t\in I}\sum_{i\in I}\left[\theta_{ff} * \left(\frac{G^t_{existing,i}}{\eta_{ff-existing}} + \frac{G^t_{new,i}}{\eta_{ff-new}}\right) + \theta_{imp,i} * V_i^t\right] \quad \text{(I-11)}$$

GHG emissions of remaining fossil fuel heating, ε_{heat}, are equal to product of the complement of the heating electrification fraction simulated, $p_{heat,i}$; the blended emissions rate for heating, θ_{heat}; and the total quantity of heating fuel consumed $F_{heat,tot,i}$. This quantity is summed over all nodes *i* and is computed per Eq. (I-12):

$$\varepsilon_{heat} = \sum_{i\in I}\left(1 - p_{heat,i}\right) * \theta_{heat} * F_{heat,tot,i} \quad \text{(I-12)}$$

GHG emissions of non-electrified vehicles, ε_{veh}, are calculated per Eq. (I-13). This accounting is analogous to that for heating emissions, using the fraction of vehicle electrification simulated, $p_{veh,i}$; the blended emissions rate for vehicles, θ_{veh}; and the total quantity of vehicle fuel consumed, $F_{veh,tot,i}$. Total transportation sector emissions also include existing transportation emissions outside the scope of the current analysis, $\varepsilon_{transp,other}$, per Eq. (I-14):

$$\varepsilon_{veh} = \sum_{i\in I}\left(1 - p_{veh,i}\right) * \theta_{veh} * F_{veh,tot,i} \quad \text{(I-13)}$$

$$\varepsilon_{transp} = \varepsilon_{veh} + \varepsilon_{transp,other} \quad \text{(I-14)}$$

Industrial sector emissions from energy consumption, ε_{ind}, are added to compute total GHG emissions. Emissions from the incineration of waste are excluded from the specific formulations of future energy scenarios. To compute the overall percent reduction in GHG emissions, ω, SECTR compares total computed emissions to the user-provided reference quantity, $\varepsilon_{reference}$ per Eq. (I-15).

$$\omega = \frac{\varepsilon_{reference} - (\varepsilon_{elec} + \varepsilon_{heat} + \varepsilon_{transp} + \varepsilon_{ind})}{\varepsilon_{reference}} \tag{I-15}$$

Characterization of fossil fuel generation

Fossil fuel-based electricity generation from existing, $X_{ff,i}^{existing}$, and new, $X_{ff,i}$, capacity is modeled. In scenarios where $X_{ff,i}$ is selected, all new generation is provided by simple cycle gas turbines, because of the very low load factors of peak load increases with heating and vehicle electrification [35]. Existing fossil fuel-based generation efficiency, $\eta_{ff-existing}$, is determined from historical data; new gas turbine efficiency, η_{ff-new}, is based on advanced combustion turbines [36]. Fossil fuel generation costs are computed per Eq. (I-16).

$$C_{fuel} = \sum_{t \in T} \sum_{i \in I} 3.412 * c_{ff,i} * \left(\frac{G_{existing,i}^{t}}{\eta_{ff-existing}} + \frac{G_{new,i}^{t}}{\eta_{ff-new}} \right) \tag{I-16}$$

A capacity reserve margin on $X_{ff,i}^{existing}$ and $X_{ff,i}$ is also imposed:

$$X_{ff,i}^{existing} \geq (1 + \sigma) * G_{existing,i}^{t} \tag{I-17}$$

$$X_{ff,i} \geq (1 + \sigma) * G_{new,i}^{t} \tag{I-18}$$

To avoid significant increases in computation time, fossil fuel-based generation start-up costs are linearized as ramping costs, $c_{ff-ramp}$, on a per-MW per-hour basis ($/MW-h); this quantity is applied to $G^t_{new-diff,i}$ and $G^t_{existing-diff,i}$, variables which represent the absolute value of the hourly change in gas generation (Eqs. (I-19 and I-20)). Ramping limitations are not imposed on the gas generators [37].

$$G^t_{existing-diff,i} = | G^t_{existing,i} - G^{t-1}_{existing,i} | \tag{I-19}$$

$$G^t_{new-diff,i} = | G^t_{new,i} - G^{t-1}_{new,i} | \tag{I-20}$$

Wind capacity

Both new onshore, $X_{on,i}$, and offshore, $X_{off,i}$, wind capacities are simulated, and are limited by resource availability and maximum capacity available at each node (onshore, Eq. (I-21)) or within the study region (offshore, Eq. (I-22)):

$$X^{existing}_{on,i} + X_{on,i} \leq X^{max}_{on,i} \tag{I-21}$$

$$\sum_{i \in I} \left(X^{existing}_{off,i} + X_{off,i} \right) \leq X^{max}_{off} \tag{I-22}$$

Solar capacity

Node-specific BTM solar capacity, $X_{btm-solar,i}$, produces fixed generation at each node equal to the product of user-imposed capacity and the supplied generation potential time series, $W^t_{btm-solar,i}$. BTM solar is treated as must-run.

Utility-scale solar capacity is constrained per Eq. (I-23):

$$X_{us-solar,i}^{existing} + X_{us-solar,i} \leq X_{us-solar,i}^{max} \tag{I-23}$$

Internodal transmission

The cost of maintaining existing transmission capacity is based on user inputs for historical transmission costs and flows. Costs of new transmission capacity are defined for each internodal interface. Transmission losses of 3% between adjacent nodes are assumed, and a nominal cost of transmission ($0.01/MWh) is applied. Eq. (I-24) limits internodal transmission flow, $Z_{ii'}^{t}$, to the combined capacity of existing, $X_{tx,ii'}^{existing}$,and new, $X_{tx,ii'}$, transmission:

$$Z_{ii'}^{t} \leq X_{tx,ii'}^{existing} + X_{tx,ii'} \tag{I-24}$$

Battery storage

Energy storage is based on lithium-ion batteries and is modeled as bulk storage at each node. Modeled batteries are constrained to a power-to-energy ratio, $\varphi_{p2e-batt}$, and a single efficiency, η_{batt}, applied on both charge and discharge. A nominal $0.01/MWh cost is attached to battery charge, $\gamma_{batt,i}^{t}$, and discharge, $\delta_{batt,i}^{t}$; storage self-discharge, κ, is also included. Battery storage constraints are presented in Eqs. (I-25) – (I-29).

$$\frac{\delta_{batt,i}^{t}}{\eta_{batt}} - \eta_{batt} * \gamma_{batt,i}^{t} = (1-\kappa) * E_{batt,i}^{T} - E_{batt,i}^{t} , \qquad \forall t = 0 \tag{I-25a}$$

$$\frac{\delta_{batt,i}^{t}}{\eta_{batt}} - \eta_{batt} * \gamma_{batt,i}^{t} = (1-\kappa) * E_{batt,i}^{t-1} - E_{batt,i}^{t} \ , \qquad \forall t > 0$$

(I-25b)

$$E_{batt,i}^{t} \leq X_{batt-e,i} + X_{batt-e,i}^{existing}$$

(I-26)

$$\gamma_{batt,i}^{t} \leq X_{batt-p,i} + X_{batt-p,i}^{existing}$$

(I-27)

$$\delta_{batt,i}^{t} \leq X_{batt-p,i} + X_{batt-p,i}^{existing}$$

(I-28)

$$\varphi_{p2e-batt-min} * (X_{batt-p,i} + X_{batt-p,i}^{existing}) \leq X_{batt-e,i} + X_{batt-e,i}^{existing} \leq \varphi_{p2e-batt-max} * (X_{batt-p,i} + X_{batt-p,i}^{existing})$$

(I-29)

In the SECTR formulation, storage self-discharge and nominal storage charge and discharge costs are included to limit the number of unique model solutions, thereby allowing the model to find an optimal solution more quickly. In the case where excess low-carbon generation is available over a period of hours, storage self-discharge reduces the number of ways to fully charge the storage to a single, unique schedule. As storage technologies undergo self-discharge in reality, the self-discharge parameter better allows SECTR to simulate likely battery operation. Moreover, when excess low-carbon generation is available and battery storage is fully charged, without nominal storage charge and discharge costs, nothing prevents the model from discharging the batteries, curtailing that energy, and then using the excess generation to recharge the batteries. Nominal charge and discharge costs prevent this type of unnecessary operation.

Nuclear generation

Nodal nuclear generation, N_i^t is modeled as constant based on a user input value and is treated as must-run.

Hydropower generation

SECTR includes modules for both fixed and flexible hydropower operation per [10]. Monthly hydropower generation is split into fixed, $H_{fix,i}^{monthly}$, and flexible, $H_{flex,i}^{monthly}$ quantities based on the nodal fraction of hydropower to be considered fixed, $y_{fix,i}$, as shown in Eqs. (I-30) – (I-31); both monthly generation quantities are fit with cubic splines, f_{c-s}, per Eqs. (I-32) – (I-33):

$$H_{fix,i}^{monthly} = y_{fix,i} * H_{tot,i}^{monthly} \tag{I-30}$$

$$H_{flex,i}^{monthly} = (1 - y_{fix,i}) * H_{tot,i}^{monthly} \tag{I-31}$$

$$H_{fix,i}^{t} = f_{c-s}(H_{fix,i}^{monthly}) \tag{I-32}$$

$$H_{flex,i}^{daily} = f_{c-s}(H_{flex,i}^{monthly}) \tag{I-33}$$

While fixed hydropower generation time series, $H_{fix,i}^t$, are treated as must-run, flexible hydropower generation, $H_{flex,i}^t$, can vary throughout the day to meet a daily nodal total, $H_{flex,i}^{daily}$, per Eqs. (I-34) – (I-35).

$$\sum_{t=1+24m}^{24*(m+1)} H_{flex,i}^{t} = H_{flex,i}^{daily}, \qquad m = 0..\frac{T}{24} - 1 \tag{I-34}$$

$$H_{flex,i}^{t} \leq H_{flex,i}^{max} \tag{I-35}$$

Biofuel generation

Biofuel generation, L_i^t, is assumed to have flexible operation, and can meet up to a set amount of daily generation, L_i^{daily}, without exceeding a nodal limit, L_i^{max}, at any time step per Eqs. (I-36) – (I-37):

$$\sum_{t=1+24m}^{24*(m+1)} L_i^t \leq L_i^{daily}, \qquad m = 0..\frac{T}{24} - 1 \tag{I-36}$$

$$L_i^t \leq L_i^{max} \tag{I-37}$$

Interregional imports

Electricity imports into the study region, V_i^t, are allowed at each node. All interregional imports are subject to a maximum limit, V_i^{max}, per Eq. (I-38):

$$V_i^t \leq V_i^{max} \tag{I-38}$$

Existing generation capacity costs

A fixed cost, $EX_{cap,i}$, is applied to eligible existing generation capacity, $X_{cap,i}^{existing}$, per Eq. (I-4). All existing hydropower, nuclear, fossil-fuel, and biofuel capacity is included in this approach, per Eq. (I-39).

$$X_{cap,i}^{existing} = X_{hydro,i}^{existing} + X_{nuc,i}^{existing} + X_{gt,i}^{existing} + X_{bio,i}^{existing} \quad \text{(I-39)}$$

2.2 Model framework additional modeling capabilities

The SECTR framework has additional modeling capabilities not used in any of the SECTR-NY results. These capabilities are detailed in the following paragraphs.

Objective function

With the inclusion of hydrogen storage energy and power capacity as SECTR decision variables, the total cost of new capacity is presented in Eq. (I-40):

$$\begin{aligned} C_{new-cap} = n_{years} \\ * \sum_{i \in I} \Big[& (A_{P_{on},j} * CAP_{on,i} + omf_{on}) * X_{on,i} + (A_{P_{off},j} * CAP_{off,i} + omf_{off}) \\ & * X_{off,i} + (A_{P_{us-solar},j} * CAP_{us-solar,i} + omf_{us-solar}) * X_{us-solar,i} \\ & + (A_{P_{batt},j} * CAP_{batt-e,i}) * X_{batt-e,i} + (A_{P_{batt},j} * CAP_{batt-p,i}) * X_{batt-p,i} \\ & + (A_{P_{h2},j} * CAP_{h2-e,i} + omf_{h2}) * X_{h2-e,i} + (A_{P_{h2},j} * CAP_{h2-p,i}) * X_{h2-p,i} \\ & + (A_{P_{ff},j} * CAP_{ff,i} + omf_{ff}) * X_{ff,i} \\ & + \sum_{i'} (A_{P_{tx},j} * CAP_{tx,ii'} * d_{ii'} + omf_{tx,ii'}) * X_{tx,ii'} \Big] \end{aligned} \quad \text{(I-40)}$$

The second SECTR objective function minimizes the levelized cost of electricity (LCOE) according to Eq. (I-41), where LCOE is defined in Eq. (I-5). When this second objective function is applied, the user specifies a greenhouse gas (GHG) emission reduction, and SECTR determines the combination of low-carbon electricity percent (LCP) and heating and vehicle electrification rate (HVE) that allows for the lowest LCOE:

$$obj_2 = \ minimize(LCOE)$$

(I-41)

Energy balance constraint

With the inclusion of hydrogen storage charge and discharge capabilities, nodal energy balance is constrained per Eq. (I-42):

$$\begin{aligned}(X_{on,i} + X_{on,i}^{existing}) * W_{on,i}^t + (X_{off,i} + X_{off,i}^{existing}) * W_{off,i}^t + (X_{us-solar,i} + X_{us-solar,i}^{existing}) \\ * W_{us-solar,i}^t + X_{btm-solar,i} * W_{btm-solar,i}^t + H_{flex,i}^t + H_{fixed,i}^t + N_i^t \\ + G_{existing,i}^t + G_{new,i}^t + L_i^t + V_i^t - \gamma_{batt,i}^t + \delta_{batt,i}^t - \gamma_{h2,i}^t + \delta_{h2,i}^t \\ + \sum_{i'} [(1-l) * Z_{i'i}^t - Z_{ii'}^t] \geq \mathrm{D}_{elec,i}^t + \mathrm{D}_{heat,i}^t + D_{veh,i}^t\end{aligned}$$

(I-42)

Renewable electricity generation targets

In SECTR simulations, users can also select a renewable generation target (RGT) – a minimum percentage of electricity from onshore and offshore wind, hydropower, and solar. Accordingly, the maximum allowable electricity generated from fossil fuels, biofuels, and nuclear power over the full simulation period is constrained per Eq. (I-43).

$$\begin{aligned}\sum_{t\in T}\sum_{i\in I}\left(G_{existing,i}^t + G_{new,i}^t + L_i^t + N_i^t\right) \leq (1-RGT) * \\ \sum_{t\in T}\sum_{i\in I}\left[\mathrm{D}_{elec,i}^t + \mathrm{D}_{heat,i}^t + D_{ev,i} - V_i^t - X_{btm-solar,i} * W_{btm-solar,i}^t\right]\end{aligned}$$

(I-43)

Flexible charging of electrified vehicle demand

SECTR includes another formulation for electric vehicle charging in which $\mathrm{D}_{veh,i}^{t}$ can be computed as the sum of a fixed electric vehicle demand, $D_{veh-fix,i}^{t}$, and a flexible electric vehicle demand, $D_{veh-flex,i}^{t}$, per Eq. (I-44):

$$\mathrm{D}_{veh,i}^{t} = D_{veh-fix,i}^{t} + D_{veh-flex,i}^{t} \tag{I-44}$$

This formulation uses a daily nodal vehicle electricity requirement, $E_{veh,i}^{daily}$, calculated as the product of the nodal percentage of vehicle electrification (user-defined or computed, depending on model configuration), $p_{veh,i}$, and user-provided daily nodal electricity requirement for full vehicle electrification, $E_{veh,i}^{daily,full}$, per Eq. (I-45).

$$E_{veh,i}^{daily} = p_{veh,i} * E_{veh,i}^{daily,full} \tag{I-45}$$

Here, SECTR allows flexibility in meeting daily vehicle electrification energy requirements. Users can split daily vehicle electricity energy demand, $E_{veh,i}^{daily}$, into flexible, $E_{veh-flex,i}^{daily}$, and fixed, $E_{veh-fix,i}^{daily}$, portions based on a provided fraction of daily vehicle electricity requirement allowed to be flexible, $y_{veh-flex}$, as shown in Eqs. (I-46).

$$E_{veh-flex,i}^{daily} = y_{veh-flex} * E_{veh,i}^{daily} \tag{I-46}$$

$$E_{veh-fix,i}^{daily} = (1 - y_{veh-flex}) * E_{veh,i}^{daily} \tag{I-47}$$

In determining hourly flexible vehicle charging demand, $D_{veh-flex,i}^{t}$, SECTR requires that the user provide a timestep for the hour at which daily charging can start, $h_{veh-start}$ and a

timestep indicating the last hour at which charging is allowed, $h_{veh-end}$. The standard SECTR formulation establishes a lower limit of 4 hours, $h_{veh-min}$, for full daily flexible EV charging. The flexible vehicle charging and power constraints are shown below:

$$\sum_{t=\ h_{veh-start}}^{h_{veh-end}} D^t_{veh-flex,i} = \frac{E^{daily}_{veh-flex,i}}{\eta_{veh}}\ , \qquad for\ m = 0..\frac{T}{24} - 1 \tag{I-48}$$

$$D^t_{veh-flex,i}\ \leq \frac{E^{daily}_{veh-flex,i}}{h_{veh-min}} \tag{I-49}$$

To determine the hourly fixed vehicle charging demand, $D^t_{veh-fix,i}$, the daily fixed vehicle charging load is split equally across the same charging period. The fixed charging constraint is shown in Eq. (I-50):

$$D^t_{veh-fix,i} = \frac{E^{daily}_{veh-fix,i}}{\eta_{veh}{}^{*}\,(h_{veh-end} - h_{veh-start}+1)}\ , for\ t = (h_{veh-start} + 24m)\,..\,(h_{veh-end} + 24m),\ for\ m = \ 0..\frac{T}{24} - 1 \tag{I-50}$$

Hydrogen storage

Long-term energy storage capabilities are modeled based on potential future system costs of grid-scale power-to-gas (P2G) with hydrogen (H2) gas: H2 produced by electrolysis, $\gamma^t_{h2,i}$; availability of a low-cost gas storage reservoir, $E^t_{h2,i}$; and electricity generated by H2 combustion in a gas turbine, $\delta^t_{h2,i}$. Nodal per-unit power capacity, $CAP_{h2-p,i}$, and energy capacity, $CAP_{h2-e,i}$, cost components are assigned. Hydrogen storage efficiency, η_{h2}, is applied on both charge and discharge. A self-discharge rate, κ, is also included.

SECTR places no constraints on the hydrogen storage power-to-energy ratio. Hydrogen storage energy balance, $E_{h2,i}^{t}$; power capacity, $X_{h2-p,i}$; energy capacity, $X_{h2-e,i}$; charging, $\gamma_{h2,i}^{t}$; and discharging, $\delta_{h2,i}^{t}$, constraints are shown in Eqs. (I-51) – (I-54).

$$\frac{\delta_{h2,i}^{t}}{\eta_{h2}} - \eta_{h2} * \gamma_{h2,i}^{t} = (1-\kappa) * E_{h2,i}^{T} - E_{h2,i}^{t} \ , \qquad \forall t = 0$$

(I-51a)

$$\frac{\delta_{h2,i}^{t}}{\eta_{h2}} - \eta_{h2} * \gamma_{h2,i}^{t} = (1-\kappa) * E_{h2,i}^{t-1} - E_{h2,i}^{t} \ , \qquad \forall t > 0$$

(I-51b)

$$E_{h2,i}^{t} \leq X_{h2-e,i}$$

(I-52)

$$\gamma_{h2,i}^{t} \leq X_{h2-p,i}$$

(I-53)

$$\delta_{h2,i}^{t} \leq X_{h2-p,i}$$

(I-54)

Identical to the treatment of battery storage, hydrogen storage self-discharge and nominal charging and discharging costs are included to limit the number of unique model solutions for a given SECTR configuration.

Flowchart summary for running a SECTR model scenarios

Figure I-1 presents a flowchart that summarizes the main steps for a user – broadly defined as anyone defining or executing a SECTR configuration – to instantiate and solve SECTR model scenarios. In short, after defining the fixed variables and parameters (see *Nomenclature*), specifying two of the three scenario configuration parameters – low-carbon electricity percent,

LCP; heating and vehicle electrification (HVE) rates p_{heat} and p_{veh}; and GHG reduction, ω – allows SECTR to determine the cost-optimal energy system design for a future decarbonization scenario.

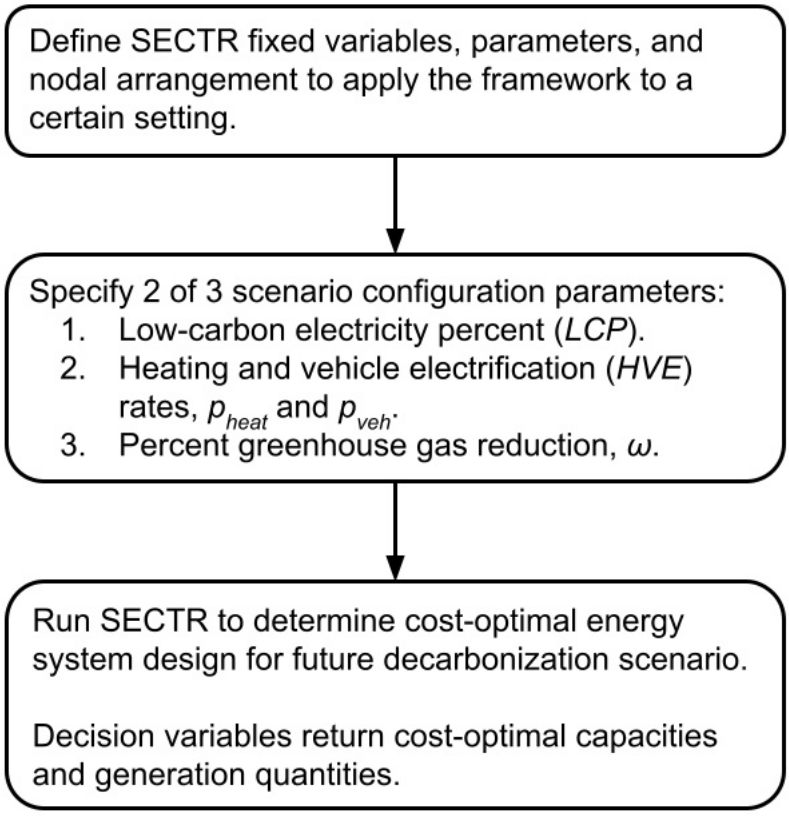

Figure I-1: Flowchart for instantiating and solving SECTR general formulation model scenarios.

2.3 Application to New York State

The SECTR framework is applied to New York State (SECTR-NY), which provides a useful study area for several reasons, including:

- A 2019 law [1] mandating significant, quantifiable decarbonization targets in the years 2030, 2040, and 2050.

- A single electricity supply system operator and market administrator – the New York Independent System Operator (NYISO) – covering the extent of New York State.
- Well-defined transmission interfaces, both internal (between NYISO zones) and external (imports/exports between NYISO and other load areas).
- Diverse and geographically heterogeneous loads and potential renewable resources.
- Definable effects of population and built environment density on current system costs and documented costs of new infrastructure capacity.
- Extensive data availability for the current electricity system and statewide GHG emissions.

Four nodes are defined for NYS by grouping NYISO zones based on the state's major transmission interfaces (see Figure I-1). The existing system is generally defined by the most recent reference data available; however, load and weather time series data for 2007-2012 are used in the model formulation because the reference model data for hourly wind and solar resource potential are available for only those years. Monthly characteristics of electricity supply and demand time series over the six modeled years are shown in Figure I-2.

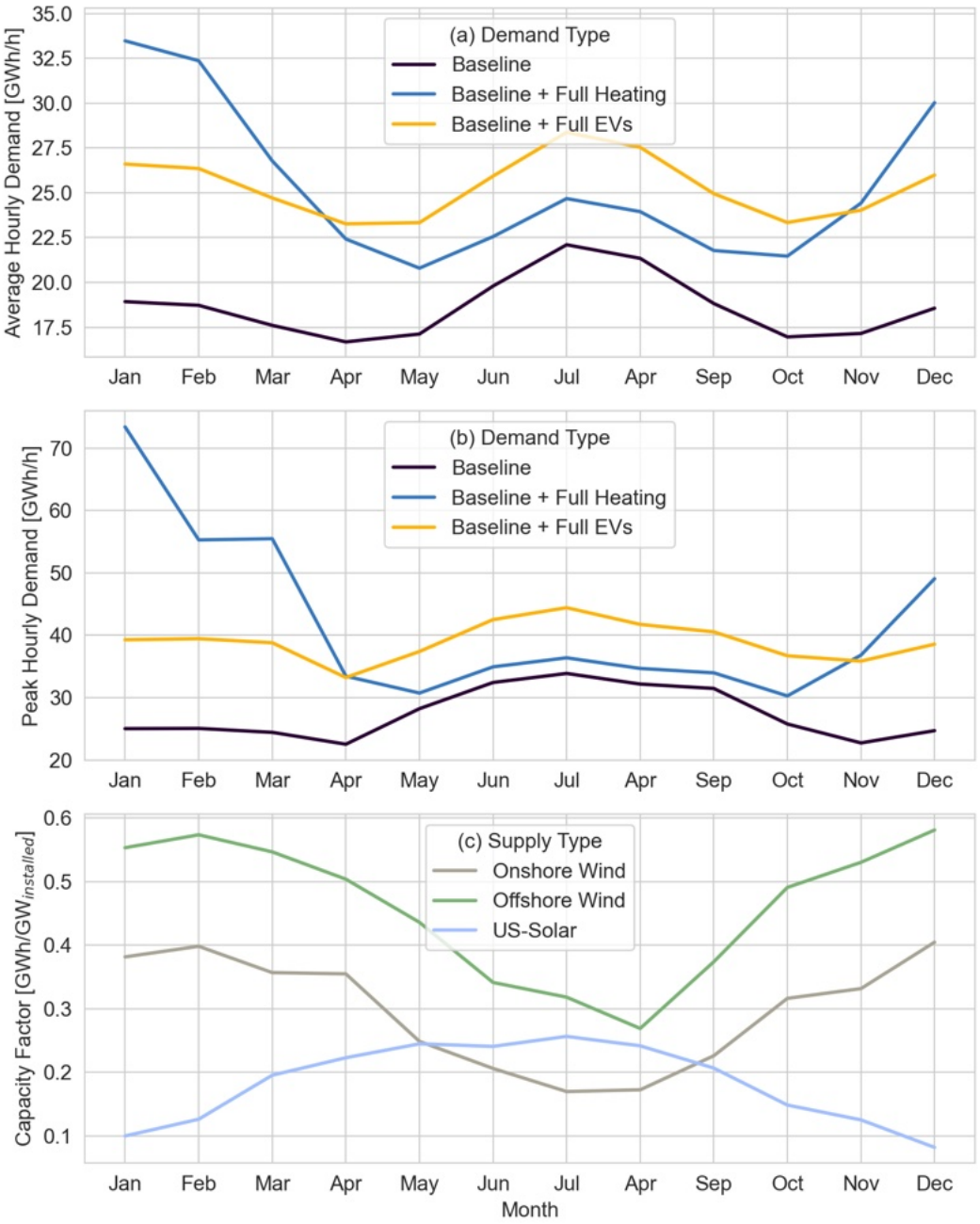

Figure I-2: (a) monthly averages of hourly electricity demand, (b) monthly peak of hourly electricity demand, and (c) monthly capacity factors for wind and solar resources in NYS.

The subsections below detail the SECTR-NY parameterization, including descriptions of all data sources used and model data development. In SECTR-NY, New York State (NYS) is split into

four nodes based on the major transmission interfaces of the New York Independent System Operator (NYISO) control area; these nodes are shown in Figure I-1.

In all simulations, low cost estimates are adopted for the technologies with multiple estimates available. All new generation technologies are annualized with a 20-year annualization period; all storage technologies are annualized with a 10-year annualization period. All model constraints presented in Sections I-2.1 and I-2.2 that contain variables with nodal indexing are applied over all nodes in the study region; constraints which contain variables with temporal indexing are applied over all timesteps in the study period.

Nodal electricity demands

The existing electricity demand used is the 2007-2012 demand in each NYISO load zone [38], aggregated at each node per Figure I-1; the average existing statewide demand is 18,655 MWh/h. Table I-1 shows average and peak electricity demands at each node. Current electricity demands include some amount of electricity usage for heating and very limited use for passenger vehicles. Here, new electricity demands from converting current fossil fuel end-uses in buildings and on-road vehicles to electric technologies are also considered. (As discussed in the Section I-2.1, fossil fuel end-uses in buildings are thermal and dominated by space heating, “heating” is used for short.)

Nodal electricity demands for heating fossil fuel conversion to electric heat pumps (EHPs) are based on a nationwide building heating model described in detail in a recently published paper [34] and applied to 2007-2012 temperature data [39]. To convert fossil fuel demands to thermal loads, current average fossil fuel equipment efficiencies of 82% for space heating and 58% for DHW are assumed based on average values for “Installed Base” equipment from the US

Energy Information Administration (EIA) [40]. The temperature-dependent coefficient of performance (COP) of new EHPs is based on the 90th percentile performance of EHPs in a regularly updated database of "cold climate" EHPs [41] and modeled per [39]. The COP of domestic hot water (DHW) EHPs was assumed to be a constant 2.32 based on the highest field-validated product performance from an National Renewable Energy Laboratory (NREL) study [42]. Full heating electrification results in a computed statewide average additional electricity load of 7573 MWh/h; however, the conversion of existing electric resistance heating to EHPs is also considered, which reduces statewide average heating electricity demand to 6716 MWh/h. Regional and statewide computed average and peak electrified heating values are shown in Table I-1.

To parameterize potential electric vehicle charging demand, the total 2018 volumetric sales of gasoline and diesel to New York transportation customers [43] are converted to miles driven using an assumed 21.0 miles per gallon (mpg). The latter assumption is based on an average vehicle age of 11.8 years in 2019 per the Bureau of Transportation Statistics [44] and the corresponding average "Real World" fuel economy of 2008 model year vehicles per the US Environmental Protection Agency (EPA) [45]. The nodal distribution of the fuel sales is assumed to be equal to the distribution of 2016 county level gasoline sales aggregated to the nodal level [46]. This mileage is then converted into daily temperature-dependent EV charging profiles using NREL's EVI-Pro model API [47] assuming 1/3 100-mile range EVs and 2/3 250-mile range EVs (based on a fixed ratio of the NREL model); weekends and weekdays are treated identically, using

a 5:2 weighted average of weekday and weekend profiles for each day[ii]. This approach results in a computed average statewide EV demand of 6769 MWh/h. Because of the many assumptions involved and the closeness of this value to the net additional potential demand from heating electrification, the EV demand series is scaled to an equivalent 6716 MWh/h average demand to facilitate more direct comparison between the two. Regional and statewide computed average and peak electrified vehicle values are also shown in Table I-1.

Table I-1: Existing and potential new nodal electricity demands.

Node	Existing Electricity Demand [MWh/h][a]		Computed Potential Net New Heating Electricity Demand [MWh/h][b]		Computed Potential New Electric Vehicle Demand [MWh/h][b]	
	Average	Peak	Average	Peak	Average	Peak
1	6383	10,467	2178	20,982	2458	5471
2	2495	4795	1059	11,347	1182	2641
3	7211	13,623	2376	13,303	1667	3642
4	2567	5933	1103	6601	1409	3081
Statewide	18,655	33,876	6716	51,088	6716	14,836

[a] NYISO [38].
[b] See the text of this section.

Internodal transmission

In SECTR-NY, both existing internodal transmission limits and costs are characterized. Existing transmission limits assumptions shown in Table I-2 are those assumed by NYISO for the year 2021 in recent system reliability simulations [48].

[ii] The NREL tool requires selections among fixed options for various inputs, the following of which were selected: 80% sedans, 20% SUVs; middle option of 80% for home charging preference; middle option of 75% for home charging access; equal usage of Level 1 and Level 2 home charging; 80% of work charging using Level 2 chargers; and minimum delay in charging at both home and work locations.

Table I-2: SECTR internodal existing transmission limits and costs of existing and new transmission.

Interface	Miles[a]	Existing Limits [MW][b]		New Transmission		
		West to East	East to West	Cost of New Transmission Capacity [$/MW-mi][c]		New Transmission O&M costs [$/MW-yr][d]
				$/MW-mi	$/GW	
1: Node 1 to 2	300	5000	3400	2400	720	2806
2: Node 2 to 3	150	7000	7000	4800	720	2357
3: Node 3 to 4	60	1613	220	12,000	720	277

[a] Distance between nodes taken as the distance between the representative cities of Buffalo, Albany, New York City, and Brentwood, per Google Maps.
[b] NYISO [48].
[c] See the text of this section.
[d] NREL [49].

Projecting costs of specific large-scale transmission upgrades is difficult. To evaluate the effect of transmission prices on future energy scenarios, public information on the costs of recent and proposed transmission projects in NYS was reviewed, as well as cost assumptions used in other studies of the region. References used in this assessment include: For Interface 1 (Node 1 to 2), Table I-2 shows the approximate average of $1400/MW-mi for simulated aboveground High Voltage Direct Current (HVDC) [50]; and $3614/MW-mi for underground HVDC in a NYISO study of the region [51]. For Interface 2 (Node 2 to 3), the Table I-2 value is approximately ¾ of the cost of $6567/MW-mi for a recent NYS underground HVDC transmission installation [52] (adjusted downward due unique challenges surrounding this project). For Interface 3 (Node 3 to 4), a transmission upgrade cost of $12,000/MW-mi is assumed based on a previous underground HVDC transmission project between New Jersey and Long Island [53]. With the above per-(MW-mi) costs of upgraded transmission and the assumed distances between each node's representative city, per-GW costs of new transmission are equal at every interface.

The NREL Jobs and Economic Development Impact (JEDI) Transmission Line Model [49] is used to compute internodal O&M costs for new transmission; new transmission capacity between Nodes 1 and 2 assumes O&M costs for 500 kVAC lines, capacity between Nodes 2 and 3 assumes O&M costs for 345 kVAC lines, and capacity between Nodes 3 and 4 assumes O&M costs for HVDC reinforcements.

The annual cost of maintaining existing transmission capacity is assumed to be the total costs recovered through electricity sales based on EIA data [54]: Based on the 2019 transmission contribution to electricity unit costs ($16.9/MWh at Nodes 1 and 2; $27.3/MWh at Nodes 3 and 4) and 2019 total electricity sales (69.683 TWh at Nodes 1 and 2; 75.52 TWh at Nodes 3 and 4), total annual cost for existing transmission was computed to be approximately $3.239B.

Characterization of fossil fuel-based electricity generation

SECTR uses a simplified characterization of the existing NYS fossil fuel electricity generation fleet and new generation capacity at each node without modeling individual generators; relevant assumed values described in this section are summarized in Table I-3. As natural gas provides 96% of fossil fuel-based electricity generation in NYS [55] and generators that burn natural gas (alone or as part of dual fuel capabilities) produce 99% of NYS fossil fuel-based electricity generation [56], only existing gas-fueled electricity generation capacity (including dual fuel generators) are considered, equal to the nameplate capacity operational at the end of 2019 per NYISO [57]. The assumed cost of existing electricity generation capacity – all existing generation modeled, including natural gas, hydropower, biofuel and nuclear – at each node is derived from capacity market costs used in a recent New York State Energy Research and

Development Authority (NYSERDA) study[iii]. Generator start-up costs are assumed to be $79/MW-h, the value for combined cycle gas turbines (CCGT) in a recent NREL study [37]. An electricity generation efficiency of 42.8% is assumed for existing natural gas generation based on NYS electric power sector total natural gas consumption [58] and natural gas-based electricity generation [59] for 2019. Modeled natural gas prices for electricity generation at each node are derived from regional natural gas avoided costs in a recent NYSERDA study[iv].

New gas-fueled generation costs are adopted based on industrial frame gas turbines (GTs) per EIA's 2020 Annual Energy Outlook [36]. These GTs have node-specific capital costs, statewide fixed and variable operations and maintenance costs, and constant 34.4% efficiency. New generator start-up costs are assumed to be $69/MW-h, the value for GTs in a recent NREL study [37]. Natural gas prices for new generators are assumed to be the same as those for existing generators at each node. Existing and new natural gas-based electricity generation capacity are constrained to be a minimum 1.189 times larger than peak generation, based on NYISO's 18.9% statewide capacity reserve margin for the 2020-2021 capability year [60].

[iii] The reference study [61] contains capacity market costs for New York City (NYC), Long Island (LI), Lower Hudson Valley (LHV) and Rest of State (ROS). Here, Node 1 is assumed to be 100% ROS; Node 2 to be 50% LHV and 50% ROS per the approximate actual capacity distribution [57]; Node 3 to be 87% NYC and 13% LHV per the reference study; and Node 4 to be 100% LI.

[iv] The reference study [61] contains natural gas avoided costs for Upstate/Western NY (UWNY), Hudson Valley (HV), and New York City and Long Island (NYC-LI). Node 1 is 100% UWNY, Node 2 is 100% HV and Node 4 is 100% NYC-LI. Node 3 is assumed to be 87% NYC-LI and 13% HV per the reference study.

Table I-3: Nodal gas-fueled electricity generation assumptions.

Node	Wholesale Nat. Gas Prices [$/MMBTU][a]	Existing Gas-Fueled Generation			New Gas-Fueled Generation			
		Capacity [MW][b]	Capital Cost [$/kW-yr][a]	Start-up Costs [$/MW-h][c]	Capital Cost [$/kW][d]	Fixed O&M Cost [$/kW-yr][d]	Variable O&M Cost [$/MWh][d]	Start-up Costs [$/MW-h][c]
1	2.89	3934.2	27.640	79	772	6.97	4.48	69
2	4.04	8622.5	53.440	79	772	6.97	4.48	69
3	3.67	10,249.9	101.303	79	1034	6.97	4.48	69
4	3.62	4192.7	104.600	79	1034	6.97	4.48	69

[a] NYSERDA [61].
[b] NYISO [57].
[c] Bloom et al. [37].
[d] EIA [36].

Wind power capacity and generation

Existing onshore wind capacities at each node are those active by the end of 2019 [57] as shown in Table I-4.

Wind power potential capacity and power output are based on model data developed by NREL for 126,000 potential wind sites [62,63]. First, onshore wind power potential time series data were adjusted to account for consistent over-predictions based on historical output of existing sites in NYS [64]. After this adjustment, a single wind potential timeseries was produced for each of the two upstate nodes[v] by computing the capacity-weighted potential timeseries of all NREL-modeled sites in each node.

To determine the offshore wind potential timeseries, potential timeseries for all NREL modeled wind sites within NYS maritime boundaries are collected; these timeseries are then weighted by modeled site capacity to return a single potential timeseries. This single timeseries

[v] Onshore wind capacity is ignored for downstate nodes 3 and 4 due to space constraints and the likelihood of a large buildout of offshore wind capacity connected to these nodes.

is adjusted based on a previously published logit transform method [64] so that the new capacity factor equals the estimate from a more recent NREL wind energy resource assessment [65], after subtracting electrical and wake losses[vi]. This adjusted timeseries is applied to both downstate nodes.

High and low costs are computed for onshore (only available in upstate Nodes 1 and 2) and offshore (only available in downstate Nodes 3 and 4) wind capacity. Based on the average of costs from three recent NREL wind technology reports [66–68] and predicted cost reductions [69], a high cost of $1992/kW and a low cost of $1698/kW are assumed for onshore wind capacity. For onshore wind, fixed O&M costs of $18.10/kW-yr are applied per the 2018 Bloomberg New Energy Finance Wind Operations and Maintenance Pricing Index [70]; installations are limited to the maximum capacities given in the NREL data set [62]. Based on a review of the costs of wind energy [71], along with cost reduction estimates [69], the high cost of offshore wind capacity is set to $3583/kW; a cost curve fit to a NREL estimates of offshore wind LCOE in 2030 [72] (5% interest, 20 year lifetime) yields a low cost estimate of $2256/kW. A fixed operations and management cost of $38/kW-yr is applied for offshore wind [73], and total offshore wind installations are capped to 57.9 GW based on potential capacity in water depths less than 60m as identified by NREL [65] (See Table I-4).

[vi] From the offshore wind resource assessment [65], the potential capacity (Appendix B) and resource energy with losses (Appendix D) in water depth less than 60m areas are collected, keeping electrical losses and wake losses but removing 6% fixed losses (Appendix J). This results in a NYS offshore wind average capacity factor of 45.9%.

Utility-scale solar capacity and generation

Existing utility-scale solar capacities at each node are those active by the end of 2019 [57] as shown in Table I-4.

The utility-scale solar potential generation time series for each node is determined by (1) identifying the capacity and location of all NYS potential grid-scale solar PV sites in a NREL model solar data set [74]; (2) computing hourly solar PV potential output using NREL's System Advisory Model [75], assuming single-axis tracking, tilted at latitude; (3) adjusting the system efficiency according to protocols specified by the California Energy Commission [76]; and (4) aggregating the individual site time series at each node, weighted by each site's capacity per the NREL data set.

High costs of new utility-scale solar PV capacity of $1341/kW at Nodes 1 and 2, and $1593/MW at Nodes 3 and 4 are adopted based on location-specific capital cost inputs to EIA's Annual Energy Outlook [36]. Low cost estimates are computed by applying a 25% cost reduction to high cost estimates, which is approximately the average of the cost reductions seen for onshore (15%) and offshore (37%) wind capacity, described above: $1006/kW in Nodes 1 and 2 and $1195/kW in Nodes 3 and 4. A statewide $10.4/kW-yr fixed O&M cost is set for new solar capacity based on a recent NREL benchmark for utility-scale tracking PV [77]. To account for space limitations, the maximum potential utility-scale solar PV capacity is determined by county and then aggregated to the nodal level, per Table I-4. For each county, the maximum capacity is based on 1) the smaller quantity of (a) existing cropland, per the 2017 USDA Census of Agriculture [78], or (b) 10% of the county's total land area; and 2) an assumed 8.5 MW/acre [79].

Table I-4: Nodal existing and maximum wind power and utility-scale solar capacities.

	Existing Capacity [MW]		Maximum Potential Capacity [MW]		
Node	Onshore Wind[a]	Utility-scale Solar[a]	Onshore Wind[b]	Offshore Wind[c]	Utility-scale Solar[d]
1	1985.25	0	32,402	0	212,710
2	0	0	4376	0	44,899
3	0	0	0	57,938	481
4	0	56.5	0		2743

[a] NYISO [57].
[b] Draxl et al. [62].
[c] Musial et al. [65].
[d] See the text of this section.

Behind-the-meter solar capacity and generation

Nodal BTM solar capacity is imposed exogenously on the optimization based on a user-provided year and a nodal capacity distribution, itself determined by a NYISO-projected 9 GW solar capacity scenario [80]. Statewide BTM solar capacity is based on a logistic growth function of the general form shown in Eq. (I-55) fit to historical capacity data for the years 2000-2019 [81]:

$$\sum_{i \in I} X_{btm-solar,i} = \frac{K}{1 + Qe^{-B(year-M)*1/v}} \tag{I-55}$$

where K = 10,982.023; Q = 1.680925e-4; B = 0.1202713, M = 1995.067; v = 4.955324e-6.

Existing nodal capacity as of the end of 2019 [81] and projected distribution computed per Eq. (I-55) for example years are shown in Table I-5.

Table I-5: Nodal behind-the-meter (BTM) solar capacity

	BTM Solar capacity (MW) for given year			
Node	Current[a]	2030	2040	2050
1	562	2109	3009	3348
2	523	2364	3372	3752
3	293	1096	1564	1740
4	259	1039	1482	1649

[a] At the end of 2019 per NYSERDA [81].

The BTM PV generation time series for each node is determined by (1) selecting a representative city for each NYISO zone from those in the NREL National Solar Radiation Database [82]; (2) computing hourly solar PV potential output using NREL's System Advisory Model [75], assuming a fixed axis, tilted at latitude; (3) adjusting the system efficiency according to protocols specified by the California Energy Commission [76]; and (4) aggregating zonal time series at each node weighted by zonal capacities in the NYISO-projected 9 GW solar capacity scenario [80].

Energy storage

Existing battery storage power capacity was extracted from the EIA energy mapping system [83], and existing battery storage energy capacity was determined from news reports and websites corresponding to recently installed projects[vii]; these quantities are presented in Table I-6. Although the SECTR General Formulation allows per-unit power capacity and per-unit energy capacity cost components, for the present analyses only energy capacity costs are included. High and low costs are set based on the "Mid" and "Low" cost projections for 2030 from NREL [84]: $208/kWh and $144/kWh, respectively. A power-to-energy ratio of 0.25 kW/kWh is assumed

[vii] Node 1: East Pulaski BESS [87] and Lockheed Martin RMS [88]. Node 2: KCE NY 1 assumed to be 4 hour battery system [89]. Node 4: East Hampton Energy Storage Center [90] and Montauk Energy Storage Center [91].

based on common 4-hour battery systems, with 94.6% charge and discharge efficiencies based on the 89.5% roundtrip efficiency of a commercially available battery storage system [85]. A 10-year lifetime [86] is adopted for modeled batteries. Batteries are also assigned a self-discharge rate of 0.1%/hr.

Table I-6: Existing nodal battery energy and power capacity.

	Existing Battery Capacity	
Node	Battery Energy [MWh]	Battery Power [MW]
1	5.2[a]	3[a]
2	80	20[b]
3	0	0
4	65[c]	10[c]

[a] Battery capacities taken from [87,88].
[b] Key Capture Energy [89]; the facility is assumed to be a 4 hour battery system.
[c] Battery capacities taken from [90,91].

For long-term storage, the use of hydrogen electrolysis and combustion in a gas turbine is assumed, with model-selected deployment analogous to battery storage based on cost components for both power capacity and energy capacity. A power capacity cost of $3013/kW is adopted based on a recent study [92] for Nodes 1 and 2; the same capital cost adjustment for GTs is then applied for Nodes 3 and 4, resulting in $4036/kW. For hydrogen storage capital costs, a per-unit energy cost of $0.35/kWh is set for geologic storage in Node 1 based on an NREL study (and adjusting from 2008 dollars to 2020 dollars) [93]. For other nodes, hydrogen storage is assumed to occur in carbon fiber storage tanks given the lack of geologic formations for storage and higher population density; a storage cost of $8.29/kWh is applied based on annually updated Department of Energy hydrogen storage cost analysis [94]. A fixed operations cost of $48.87/kW-yr is assumed based on an earlier study [95]. Charge and discharge efficiencies of 59.2% are

adopted based on 35% roundtrip efficiency in a recent NREL analysis [96] referencing an earlier study [97]. A self-discharge rate of 0.1%/hr is set.

Nuclear power

The nuclear power landscape in NYS is evolving, as the last operational generator of the Indian Point two-generator plant in Node 3 shuttered on April 30, 2021 [98], and nuclear generators in Node 1 have been subsidized in recent years. To investigate the impact of capacity retirements, the SECTR-NY formulation can either include or ignore these nuclear generators. Nuclear capacity is distributed across all four model nodes per NYISO [57] as shown in Table I-7 (which for clarity shows no nuclear at Nodes 2 and 4). Electricity generation is assumed to be constant throughout the simulation period and equal to the average electricity production of those generators in 2019 according to NYISO [57]. The price of nuclear electricity at each node is computed from the average 2019 day-ahead locational based marginal pricing (LBMP) [38] of each nuclear generator at each node, weighted by the 2019 total net electricity generation [57] of each of those generators. The price at Node 1 is increased to account for subsidies of the nuclear generators at that node, funded by Zero Emission Credits (ZECs). Per Eq. (I-56), the per energy unit subsidy is computed from the 2020 compliance year ZEC rate [99], NYISO's 2020 baseline demand forecast [57], and the constant output of nuclear electricity at Node 1 from Table I-7.

$$\left\{\begin{matrix}Nuclear\ Price\ Subsidy \\ at\ Node\ 1\end{matrix}\right\} = \frac{\left\{\begin{matrix}2020\ Compliance \\ Year\ ZEC\ Rate\end{matrix}\right\} \times \left\{\begin{matrix}NYISO\ 2020\ Baseline \\ Demand\ Forecast\end{matrix}\right\}}{\{Constant\ Electricity\ Generation\ at\ Node\ 1\}} \quad \text{(I-56)}$$

The assumed cost of existing nuclear electricity generation capacity at each node is the same as described above under "Characterization of fossil fuel-based electricity generation."

Table I-7: Nodal existing nuclear power characteristics

Node	Generation Capacity [MW][a]	Constant Electricity Generation [MWh/h][a]	Capacity Cost [$/kW-yr][b]	Electricity Price [$/MWh][c]
1	3536.8	3207	27.640	37.94
2	0	0	N/A	N/A
3	2311	1906	101.303	26.82
4	0	0	N/A	N/A

[a] NYISO [57].
[b] NYSERDA [61].
[c] See the text of this section.

Hydropower

The methodology for creating hydropower fixed and flexible generation time series is described in a recent paper [10]. Actual monthly hydropower output by facility is collected for 2007-2012[viii] from EIA Form 923 [100], and then is aggregated at each node. The two largest NYS hydropower facilities (both located at Node 1) operate near their maximum capacity given available stream flows; accordingly, fixed hourly time series are provided for these generators. The remaining hydropower generation and capacity in Nodes 1 and 2 are considered to be flexible with provided daily total electric energy generation requirements. Total fixed and flexible hydropower capacities are computed from the nameplate capacities operational at the end of 2019 [57]. Hourly generation is determined endogenously by the hydropower methodology detailed in the General Formulation. Hydropower-generated electricity prices are based on

[viii] Monthly generation quantities for 2007-2012 are used to align with the wind, solar, and demand time series.

recent prices for such electricity in NYISO's day-ahead market[ix]. The assumed cost of existing hydropower electricity generation capacity at each node is the same as described above under "Characterization of fossil fuel-based electricity generation." The values described here are summarized in Table I-8.

Table I-8: Existing hydropower characteristics

	Average Generation [MWh/h][a]		Capacity [MW][b]		Capacity Cost [$/kW-yr][c]	Electricity Price [$/MWh][d]
Node	Fixed	Flexible	Fixed	Flexible		
1	2395	328	3948	769.4	27.640	18.47
2	0	270	0	608.7	53.440	28.02
3	0	0	0	0	N/A	N/A
4	0	0	0	0	N/A	N/A

[a] EIA [100].
[b] NYISO [57].
[c] NYSERDA [61].
[d] See footnote *vii*.

Biofuel-based electricity generation

SECTR-NY classifies various electricity generation feedstocks as "biofuels": wood and wood waste, biogas, and solid waste. In NYS, biofuel capacity is distributed across all four model nodes per NYISO [57] as shown in Table I-9. Intraday biofuel electricity generation is flexible as described in Section I-2.1; maximum daily electricity generation is assumed to be constant throughout the simulation period and equal to the average daily electricity production of those generators in 2019 according to NYISO [57]. Biofuel-generated electricity prices are based on

[ix] All based on 2019 hourly day-ahead LBMP [38] and weightings by total 2019 electricity production [57]: Node 1 cost is the weighted average LBMP for Moses Niagara and St. Lawrence hydropower facilities; Node 2 is the weighted average LBMP of the four highest producing hydropower facilities at that node (62% of total hydroelectricity production at that node).

recent prices for such electricity in NYISO's day-ahead market[x]. The assumed cost of existing biofuel-based electricity generation capacity at each node is the same as described above under "Characterization of fossil fuel-based electricity generation."

Table I-9: Nodal existing biofuel characteristics

Node	Generation Capacity [MW][a]	Daily Electricity Generation [MWh][a]	Capacity Cost [$/kW-yr][b]	Electricity Price [$/MWh][c]
1	258.0	3289.041	27.640	20.66
2	45.0	473.425	53.440	27.41
3	59.7	1046.575	101.303	27.05
4	142.2	2445.479	104.600	32.29

[a] NYISO [57].
[b] NYSERDA [61].
[c] See footnote *ix*.

External imports

NYISO currently imports significant quantities of low-carbon electricity from Hydro-Quebec (HQ), a net average of 1247 MWh/h in 2019 [57]; as such, electricity imported at this interface with Node 1 is included as a decision variable constrained to the maximum interface limit specified by NYISO (1.5 GW) [38]. A cost of $22.13/MWh is attributed to this imported electricity based on average 2019 day-ahead LBMP [101] and including capacity market payments for 1114 MW capacity per NYISO [102].

NYS regulators are nearing approval for plans for the Champlain Hudson Power Express, a 1250 MW HVDC transmission line that would bring hydropower-produced electricity from

[x] All based on 2019 hourly day-ahead LBMP [38] and weightings by total 2019 electricity production [57]: Node 1 cost is the average LBMP for the four highest producing biofuel facilities at that node (58% of total biofuel electricity production at that node); Node 2 is the average of Zone F and G LBMP; Node 3 is the average LBMP for the 1 biofuel facility at that node; Node 4 is the weighted average LBMP of all four biofuel facilities at that node.

Quebec to New York City [52], which is also included in recent NYC local legislation [103]. As such, additional electricity import into Node 3 is included in future energy system scenarios. The precise cost of this electricity supply is unknown; however, a price of $70/MWh is adopted based on publicly available information, personal and public conversations about the project, and various possible financing parameters[xi]. The line is assumed to provide 1125 MWh/h continuous based on the approximate 90% capacity factor of existing upstate Hydro-Quebec import lines [38] and an understanding of the project from public and personal conversations.

Imports from other external control areas are ignored to avoid characterizing or modeling future developments in regions that currently rely largely on fossil fuel-based electricity generation.

Greenhouse gas emissions

As accounted by NYSERDA, NYS energy sector emissions constitute 84% of total statewide GHG emissions (measured in equivalent global warming potential of carbon dioxide, CO_2e) as of 2016 [104]. The remaining 16% of GHG emissions comes from industrial processes, agriculture, and waste.

In New York's Climate Leadership and Community Protection Act (CLCPA), statewide GHG emissions accounting includes GHGs produced in NYS and GHGs produced outside NYS that are associated with imported electricity and fossil fuels [1]. Table I-10 shows emissions factors for carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O) compiled from a variety of sources;

[xi] Our calculations are generally in the $65-70/MWh range based on the project website's lower bound capital cost [52], higher potential upfront costs that have been discussed publicly, various annualization periods, average HQ export revenues ($1441M on 33.7 TWh in 2019 [334]), and the approximate 90% capacity factor of existing upstate Hydro-Quebec import lines [38].

the table also includes values for carbon dioxide equivalent (CO_2e). CO_2e is a single metric that combines the effect of multiple GHGs based on their global warming potential (GWP). CLCPA requires GWP values based on the amount of warming impact relative to CO_2 when integrated over a 20-year time frame. Here, respective GWPs of 86 for CH_4 and 264 for N_2O are used, in accordance with the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) [105]. CH_4 emissions, particularly for natural gas, are largely dependent on venting at wellheads and leakage in transmission and distribution infrastructure; understanding these effects is the subject of ongoing research, but recent efforts focused on New York State provide a reliable reference point [106].

Table I-10: Emissions factors [g/MJ] for GHG contributors

Energy source	CO_2	CH_4	N_2O[f]	CO_2e
Coal	92[a]	0.185[c]	$1.52 \cdot 10^{-3}$	108.31
Petroleum	73[b]	0.093[d]	$5.69 \cdot 10^{-4}$	81.15
Natural Gas	55[a]	0.641[e]	$9.48 \cdot 10^{-5}$	110.18

[a] Based on high-heating values per Hayhoe et al. [107] as documented by Howarth et al. [106].
[b] CO_2 emission factor for petroleum is the high-heating value from Howarth et al. [108] as reported by Howarth et al. [106].
[c] As computed by Howarth et al. [106] based on the ratio of total methane emissions during coal mining and total coal production in the U.S. in 1990 from IPCC reporting [109], with a coal heating value of 27 MJ/kg [108].
[d] Based on CH_4 emissions from petroleum production from the National Energy Technology Laboratory (NETL) [110] as documented by Howarth et al. [106].
[e] Computed from assumptions of: CH_4 emission rate of 3.6% as used for NYS in Howarth et al. [106] based on a range computed by Alvarez et al. [111] and Howarth et al. [108]; natural gas to be 93% CH_4 [112]; and a high-heating value of 52.2 MJ/kg for natural gas in the U.S. market [113].
[f] EPA [114].

Per the targets set in the CLCPA [1], emissions reductions relative to a 1990 reference value are computed. Reference CO_2 and CH_4 emissions for electricity, buildings, industrial, transportation are calculated by using the 1990 EIA fuel consumption estimates [43] and emission factors in Table I-10; CO_2 and CH_4 emissions for electricity imports in 1990 are taken

directly from Howarth et al. [106]; CO_2 and CH_4 emissions for waste incineration and all N_2O emissions in 1990 are from the NYSERDA inventory of GHG in NYS [104]. Thus an $\varepsilon_{baseline}$ of 302.770 $MMtCO_2e$/year is computed, per Table I-11. Table I-11 further delineates emissions that are fixed in the model and those that are variable: variable emissions can change as computed by the model for a given user-defined scenario and as described by Eqs. (I-11) – (I-15) in Section I-2.1.

Current NYS electricity emissions are calculated by using SECTR to model a "current scenario". The current scenario includes all existing NYS energy infrastructure parameterized and discussed above, and assumes current capacities of wind and solar power, no additional electrification of vehicle or heating demand, and no generation from the Indian Point nuclear facility. Using the natural gas emissions factors in Table I-10 and the model-returned amount of natural gas generation needed to meet the existing electricity demand, current electricity emissions of 84.889 $MMtCO_2e$/year are computed, per Table I-11. Since SECTR-NY assumes the modeled electricity sector can be fully decarbonized, these emissions are considered variable.

Total fossil fuel usage for heating, $F_{heat,tot,i}$, is computed from the heating model [34] described above; portions of this fossil fuel usage are attributed to natural gas, fuel oil, and propane based on the 2018 residential and commercial usage of these fuels [43]. Annual heating GHG emissions, ε_{heat}, are calculated as 110.853 $MMtCO_2e$/year averaged over the six-year computation period based on the emissions factors for natural gas and petroleum (for fuel oil and propane) in Table I-10. As SECTR-NY assumes that NYS heating demand can be fully electrified, these emissions are considered variable in Table I-11.

Transportation sector emissions are determined from the 2018 EIA fuel consumption estimates [43] that are used to calculate statewide vehicle energy demand as described above. Gasoline and diesel consumption for transportation that can be electrified, $F_{veh,tot}$, and the petroleum emissions factor in Table I-10 are used to compute 73.703 $MMtCO_2e$/year variable emissions for vehicles included in the model electrification scope, ε_{veh}. Aviation fuel, hydrocarbon gas liquids, jet fuel, lubricants, residual oil, and natural gas consumption for transportation listed in the same EIA dataset [43] are considered fixed and constitute transportation emissions outside the scope of the model, $\varepsilon_{transp,other}$. From this usage data and the appropriate emissions factors from Table I-10, $\varepsilon_{transp,other}$ is computed to be 21.956 $MMtCO_2e$/year, a quantity fixed in every model run.

Industrial emissions, ε_{ind}, are calculated from the 2018 EIA fuel consumption estimates [43] and the appropriate emissions factors from Table I-10. Here, coal, natural gas, and petroleum products result in computed total emissions of 19.365 $MMtCO_2e$/year; this quantity is fixed in every model run.

Table I-11 also displays emissions from waste incineration in New York for the current system. In SECTR-NY model scenarios, waste incineration is excluded per the CLCPA [1]; as this emissions quantity is set to 0 in all model runs, it is presented as variable.

Table I-11: Relevant aggregate greenhouse gas emissions ($MMtCO_2e$/year)

Emissions Source	1990 Reference Emissions	Current Emissions as Modeled	
		Variable	Fixed
Electricity	86.772	84.889[a]	0
Electricity Imports	1.909	0	0[b]
Heating (Buildings)	100.468	110.853	0
Industrial	32.824	0	19.365
Transportation	79.532	73.703	21.956
Waste Incineration	1.265	2.784[c]	0
Total	302.770	272.229	41.321
		313.550	

[a] Based on SECTR-NY model of current system as described in this section.
[b] Electricity imports are only considered from hydropower generation.
[c] This quantity represents the 2016 value from NYSERDA inventory for waste incineration [104]. In SECTR model runs, it is set to zero.

Seasonal distribution of demand and renewable generation potentials

Existing electricity demand, electrified demand, and the renewable generation potentials of wind and solar resources used in SECTR-NY simulations all demonstrate substantial seasonal variability. Figure I-2 contains monthly values for the mean and maximum existing electricity demand; the means and maximums of existing electricity demand combined with either electrified heating or transport; and the means of onshore wind, offshore wind, and utility-scale solar generation potentials. From the top and middle panels in Figure I-2, one observes that electrified heating increases average and peak electricity demands in the winter months: Full electrification of heating increases average load by up to 15 GWh/h and peak load by up to 52 GWh/h. In contrast, electrification of transport has smoother effect. With 100% transport electrification, average load rises by 6 to 8 GWh/h in all months of the year, with the larger increases coming during the winter due to the inverse relationship between temperature and EV

charging demand. The effect on peak load is similarly consistent: Peak electricity demand increases by 12 to 15 GWh/h in all months.

Wind and solar generation potentials in NYS also display a strong seasonal dependence. Offshore and onshore wind potentials both peak in the winter months, reaching an average of 0.57 $MWh_{generation}/MW_{installed}$ and 0.40 $MWh_{generation}/MW_{installed}$, respectively; in the summer months, average generation for each decreases by approximately 50%. In contrast, utility-scale solar capacity offers peak generation potentials during the summer months, up to an average 0.26 $MWh_{generation}/MW_{installed}$, while winter months see this quantity drop to 0.10 $MWh_{generation}/MW_{installed}$.

Taking all three panels of Figure I-2 together, a clear seasonal alignment is identified between electrified heating demand and wind generation potential, indicating that electrified heating may prove effective in integrating large amounts of installed wind capacity. Moreover, summer-peaking solar generation is well-suited to meet summer loads in NYS, both those that currently exist and those that are increased by transport electrification.

3. Results

Section 3.1 establishes and distinguishes between a “Current” model configuration that mirrors existing NYS system characteristics, and a “Baseline” configuration for decarbonization scenario comparison. Section 3.2 presents the results of SECTR-NY Baseline configuration simulations for different combinations of in-state low-carbon electricity generation percentages (LCP) and heating and vehicle electrification rates (HVE). Section 3.3 compares SECTR-NY results to those published in recent NYS studies on decarbonization pathways. Section 3.4 presents additional results for the SECTR Baseline configuration. Section 3.5 contains results that

investigate the impact of SECTR system parameterization assumptions. Section 3.6 presents results contained in Section 3.3 for different heating and vehicle electrification rates (HVEs) and low carbon electricity percents (LCPs).

All results are presented for SECTR-NY simulations solved over the entire 6-year time period modeled; all specified generation and demand quantities are presented as hourly averages in Wh/h over the full 6-year simulation period.

3.1 Current system validation and Baseline configuration

The Baseline configuration deviates from the Current system configuration in three ways summarized in Table I-12: The Baseline configuration excludes existing nuclear power at Node 1, includes an additional 5 GW of solar BTM capacity corresponding to a simulation year of 2030, and simulates an additional planned 1.25 GW of hydropower import capacity into New York City (NYC). For comparative purposes, Table I-12 also includes a “Baseline with Nuclear” scenario. Table I-12 scenarios exclude any additional HVE beyond current electric heating and vehicles.

Table I-12: 'Current', 'Baseline with Nuclear', and 'Baseline' system configuration comparisons.

	Configuration Parameters			Specified System Characteristics[a,b]				Model-returned System Characteristics[b]				
Configuration	% GHG[c]	% HVE[d]	% LCP[e]	Instate Hydro [GW]	Nuclear [GW]	BTM Solar [GW]	Hydro Imports [GW]	Onshore Wind [GW]	Utility-Scale Solar [GW]	Battery [GWh]	Wind and Solar LCOE [$/MWh]	Total LCOE [$/MWh]
Current	3.6	0	38.2	5.3	3.5	1.6	1.5	2.0[f]	0.1[f]	0.2[f]	69.7	65.3
Baseline w. Nuclear	-2.0	0	42.4	5.3	3.5	6.6	2.8	2.0	0.1	1.1	69.3	68.6
Baseline	-1.6	0	40	5.3	0	6.6	2.8	9.1	2.6	2.0	67.8	72.1

[a] See Section I-2.1 for existing system characteristics and the text of this section for any modifications for the specific configuration.
[b] Besides LCOE values, all system characteristics presented indicate capacities.
[c] '% GHG' refers to the percent change in greenhouse gas emissions compared to the 1990 reference quantity. A positive value indicates a computed increase in emissions, a negative value indicates a reduction.
[d] '% HVE' refers to the percent of additional heating and vehicle electrification simulated; some heating electrification (and a very small amount of vehicle electrification) currently exists in NYS.
[e] '% LCP' refers to the percent of in-state electricity supply from low-carbon sources.
[f] Indicates model capacities that are constrained to existing capacity in the 'current' configuration.

The model-computed Current configuration LCOE of $65.3/MWh compares favorably to the actual system. An actual NYS electricity supply cost of $69.1/MWh is estimated, based on 2019 NYS generation and transmission costs [115], electricity sales [115], and total zonal electricity demands; this actual cost would include ancillary service and NYISO operation costs of approximately $2/MWh [116] that are not included in SECTR-NY. Despite the difference between these two values, the close alignment in computed costs supports SECTR-NY's applicability to the NYS system and its suitability for further analyses.

The Current configuration computes an LCP of 38.2% and a 3.6% increase in GHG emissions compared to the 1990 reference quantity. Total emissions increase because CO_2

reductions from natural gas displacing coal and fuel oil combustion are offset by GHG increases from larger transportation energy demands, methane leakage associated with natural gas production and transmission, and the retirement of a large nuclear power plant; these effects are more pronounced due to the use of the 20-year GWP value for methane in place of 100-year GWP value. Moreover, the calculated LCP of 38.2% is lower than the 2019 fraction of NYS electricity demand met by low-carbon sources (62.3%) for two reasons: 1) per the language of the CLCPA, LCP only considers in-state generation, and does not account for substantial hydropower imports from Canada; and 2) SECTR-NY does not include nuclear generation from Indian Point, as this facility was fully closed on April 30, 2021[xii].

The Baseline with Nuclear configuration – adding BTM PV and NYC hydropower imports to the Current configuration – computes a 2% GHG reduction and $68.6/MWh LCOE; the $3.3/MWh higher LCOE is due to the higher cost of hydropower imported into NYC and the reduction of regional demands due to solar BTM (i.e., existing system capacity costs are distributed over less load). Removing all nuclear capacity establishes the Baseline configuration; a 40% LCP is set for round number comparison in subsequent sections that is close to the current 38.2%. Approximately 10 GW of solar and wind capacity are installed to replace the nuclear generation, resulting in a slightly lower reduction in GHG (-1.6%) and a slightly higher LCOE ($72.1/MWh). Given the reasonable deviations from the current system model, the Baseline configuration is adopted for future scenario evaluations.

[xii] The "Current" and "Baseline with Nuclear" configurations do include generation from NYS nuclear plants besides Indian Point, as these plants remain operational as of this dissertation's publication.

3.2 Analysis of low-carbon electricity and end-use electrification scenarios

For a series of SECTR-NY simulations with different combinations of LCPs and HVEs[xiii], relationships among LCOE, GHG emissions, HVE, LCP, and renewable energy capacity are shown in Figure I-3. Here, computed LCOEs represent the total costs for supply (primarily generation, storage, and transmission), excluding delivery costs (primarily distribution system costs). HVE rates refer to new heating and vehicle electrification, as some heating (and a small share of vehicles) currently uses electricity. Note that the 40% LCP and 0% HVE scenario presented in Table I-12 is located in the bottom-left of the figure; for comparison beyond NYS, 39.7% of US electricity generation was from low-carbon sources in 2020 [117].

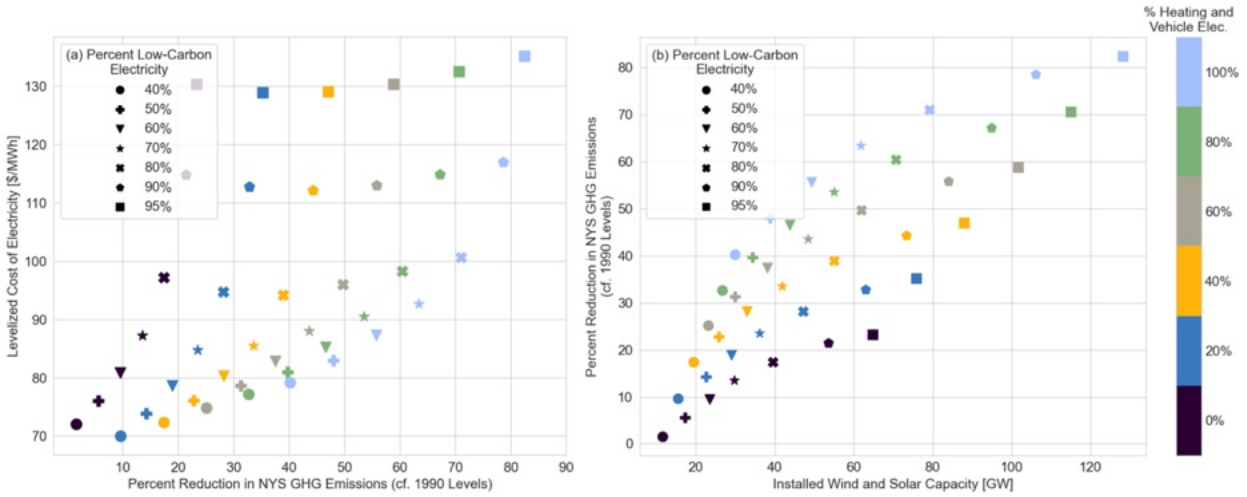

Figure I-3: (a) LCOE vs. percent emissions reduction; (b) percent emissions reduction vs. installed wind and solar capacity. All emissions reductions are compared to 1990 levels. Marker shape indicates percent low-carbon electricity (LCP), and marker color indicates heating and vehicle electrification (HVE). All points represent independently solved SECTR-NY decarbonization scenarios with specified LCP + HVEs. For scenarios shown, all low-carbon electricity generation is from wind, solar, and hydropower.

[xiii] In the scenarios presented, heating and vehicle electrification rates are equal.

Figure I-3(a) shows how computed grid supply LCOE (strictly that of the electricity utilized) rises sharply with increasing LCP for a specified HVE, whereas for a specified LCP, higher HVEs cause limited growth in LCOE. Figure I-3(b) provides a partial explanation, showing that high HVE scenarios achieve the same GHG reductions with lower installed wind and solar capacities. The results suggest that emissions reductions can be achieved with a shallower initial rise in LCOE by prioritizing electrification of heating and vehicles in conjunction with deployment of solar and wind, as opposed to the latter by itself. Added loads from HVE can even slightly reduce LCOE up to a point (20-40% HVE, depending on LCP), as the additional electricity demand decreases the per-unit cost of existing infrastructure. (The same trend holds when the system includes an average of 3 GWh/h of nuclear generation in Node 1, albeit at LCOEs approximately 10% lower; see Figure I-10.)

It is worth noting the straightforward impact of HVE on GHG emissions: In NYS, a current average emissions rate for fossil fuel-based heating of 148 kgCO2e/MMBtu$_t$ (i.e. per unit heat delivered) is computed based a recent heating model [118] and GHG emissions rate assumptions described in Section I-2.3; with electrified heating and 40% low-carbon electricity supply[xiv] in SECTR-NY, this reduces to 44 kgCO2e/MMBtu$_t$. Similar reductions occur for vehicle electrification: A current average emissions rate for fossil fuel vehicles of 543 gCO_2e/mi (per vehicle mile traveled) is computed, and 241 gCO_2e/mi for electric vehicles with 40% LCP in SECTR-NY. Therefore, even with the remaining 60% of grid power being supplied by gas-based generation, substantial reductions in overall emissions from electrification are computed.

[xiv] 40% LCP mirrors the current NYS fuel mix.

Consider two changes in system characteristics starting at the 40% LCP and 0% HVE point of Figure I-3(a). Approximately 10% GHG emissions reductions could be achieved without additional electrification and with 60% LCP at an LCOE of $80.9/MWh; this scenario represents a 3.1 GWh/h increase in average wind and solar supply. A similar emissions reduction could be achieved with a 20% HVE and no LCP increase at a cost of $70.0/MWh; the average wind and solar supply increases by 1.1 GWh/h to maintain 40% LCP with the electrification-driven increase of 2.7 GWh/h average demand. Consider now two scenarios in Figure I-3(b) with approximately 30 GW wind and solar capacity: The scenario with 50% LCP and 60% HVE has computed GHG emissions reductions of 31%, more than double the 14% reduction in the scenario containing 70% LCP and 0% HVE. Here, the computed LCOE for the first scenario ($78.7/MWh) is nearly $10/MWh less than the second scenario ($87.2/MWh).

These various trade-offs are demonstrated with four scenarios that all contain approximately 1/3 reductions in GHG, but via different combinations of LCP and HVE. For the lowest LCP scenario shown in Table I-2 (Scenario 1), GHG reductions require a high HVE that increases average load and peak load, the latter requiring larger amounts of gas turbine capacity. Comparatively, Scenario 3 contains 33 GW less gas generation capacity, accompanied by a drop in average gas generation from 15.3 GWh/h to 6.0 GWh/h. Here, higher LCP scenarios avoid increases in gas capacity and generation through additional renewable generation and battery capacity, a tradeoff that increases supply costs by $10/MWh.

Table I-13: Select scenarios achieving emissions reductions of approximately 1/3 compared to the 1990 reference quantity.

Scenario	% GHG[a]	% HVE[b]	% LCP[c]	Avg. Load [GWh/h]	Wind and Solar Cap. [GW][d]	Battery Cap. [GW]	Gas Cap. [GW][e]	Avg. Gas Gen. [GWh/h][f]	LCOE [$/MWh]
1	-32.9	80	40	29.4	26.6	4.7	63.0	15.3	77.2
2	-31.3	60	50	26.7	29.8	4.2	48.9	11.4	78.7
3	-33.6	40	70	24.0	41.8	6.9	29.9	6.0	85.5
4	-32.8	20	90	21.3	63.0	15.0	27.0	1.8	112.8

[a] '% GHG' refers to the percent change in greenhouse gas emissions compared to the 1990 reference quantity. Negative values indicate reductions.
[b] '% HVE' refers to the percent of additional heating and vehicle electrification simulated; some heating electrification (and a very small amount of vehicle electrification) currently exists in NYS.
[c] '% LCP' refers to the percent of in-state electricity supply from low-carbon sources.
[d] 'Wind and Solar Cap.' refers to installed onshore wind, offshore wind, and utility-scale solar capacity.
[e] 'Gas Cap.' contains 27.0 GW existing gas-based generation capacity and model selected new gas turbines.
[f] 'Avg. Gas Gen.' refers to the average generation over the entire 6-year simulation period from existing gas-based generation and model-selected new gas turbines.

The synergy of renewable energy generation and electrification is further explained by looking at "excess low-carbon generation": Potential electricity generation from model-selected wind and solar capacities exceeding demand. Excess low-carbon generation exists as an hourly time series of either 0 MWh (when total low-carbon generation is less than the demand) or a positive value equal to the amount of low-carbon electricity generation that exceeds demand. In model simulations, excess low carbon generation must be either 1) stored for later use, or 2) curtailed. Figure I-4(a) shows that despite significant growth in renewable energy capacity with increasing HVE, excess low-carbon electricity generation remains below 6% as long as LCP does not exceed 70%; at LCP of 50% or less, excess generation is below 1%. Figure I-4(b) shows the relationship between excess low-carbon generation and LCOE for the same scenarios in Fig. 4(a).

By combining the effects discussed thus far, three primary LCOE drivers are identified: (1) decreasing per-unit costs of existing infrastructure with increasing demand from HVE, (2) higher generation costs from wind and solar power relative to existing resources, and (3) increasing integration costs when large amounts of wind and solar power produce electricity in excess of demand. Figure I-4(b) shows a general linear trend of integration costs (curtailment and battery storage) increasing LCOE at higher percents excess low-carbon generation, but also how the effects of the three cost drivers change over the entire range of LCPs and HVEs simulated. At LCPs at or below 60%, the primary cost tradeoffs discussed earlier are observed: Higher LCOEs from more wind and solar are partially mitigated by higher utilization of existing infrastructure with HVE. In the 70-80% LCP range, a transition begins in which some spread in excess low-carbon generation affects LCOE, but the first two LCOE drivers prevail. Beyond 80%, the integration cost-driven linear relationship between increasing excess low-carbon generation and computed LCOE dominate.

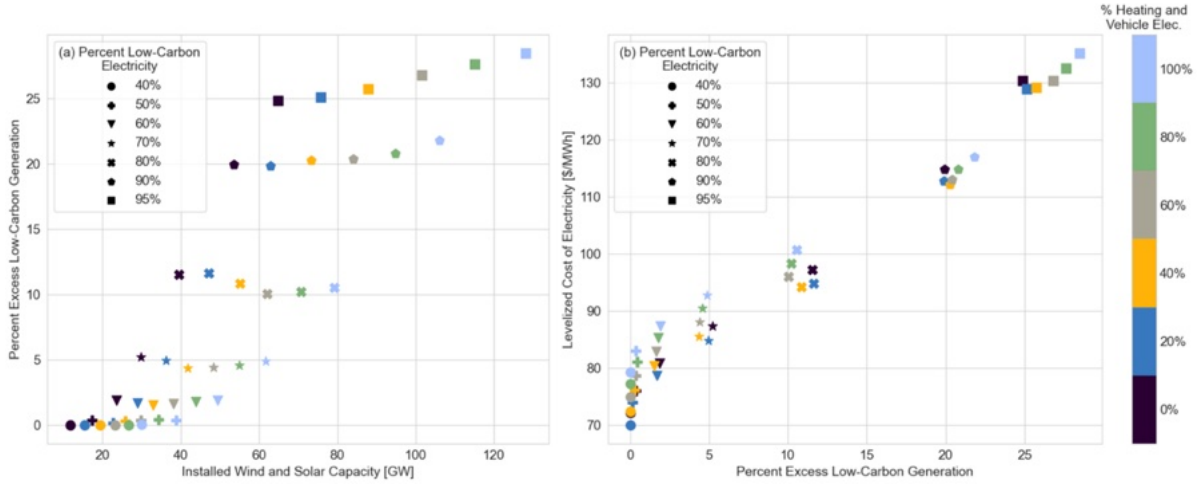

Figure I-4: (a) Average percent excess low-carbon generation for the entire 6-year simulation period vs. installed wind and solar capacity; (b) LCOE vs. percent excess low-carbon generation. Results are shown for the same independent decarbonization scenarios in Figure I-3, whereby the low-carbon electricity percent and the rate of heating and vehicle electrification are set, and SECTR-NY determines the least-cost energy system.

The results presented thus far show how electrification accompanied by a significant buildout of renewable energy can keep LCOE low. On the other hand, a focus on large LCP fractions beyond 70% represents a major cost escalation. Competing drivers and trade-offs are next examined among scenarios with increasing HVE while maintaining LCP at 60% (Figure I-5(a-d)) vs. scenarios where HVE is 40% and LCP is progressively increased (Figure I-5(e-h)). (The trends observed here hold for other combinations of HVE and LCP; see Figures I-15 and I-16.) Figure I-5(a-d) demonstrates the stable buildout of generation capacity and consistency of system behavior and costs as electrification increases. In order to meet the increased demand, low-carbon generation, gas generation, and battery capacity all increase with electrification, per Figure I-5(a); gas generation undergoes the largest capacity increase – from 27.0 GW to 67.2 GW at 100% HVE – in order to meet higher electrification-induced demand peaks. Here, additional gas capacity is selected due to its low cost relative to the model's other dispatchable generation option, battery storage. With additional policy-based constraints in place, such as a limit on additional gas turbine capacity or demand-side strategies to mitigate peak heating loads, much less new gas capacity would be built out. Electricity generation trends (Figure I-5(b)) largely mirror the expansion in generation capacity, with the ratio of solar to wind generation (combined onshore and offshore) staying consistent from 0.31 at 0% HVE to 0.34 at 100% HVE, although with an increasing amount of wind generation coming from offshore capacity. Figure I-5(d) reveals the reason for consistency in system behavior: Despite increasing average uncurtailed low-carbon electricity generation from 9.5 GWh/h at 0% HVE to 17.7 GWh/h at 100% HVE, average excess low-carbon generation only increases from 177 MWh/h to 336 MWh/h.

Electrification thus supports renewable energy integration by keeping the LCOEs of those supply resources low (Figure I-5(c)).

Conversely, optimal energy system characteristics change substantially with increasing LCPs. The previously noted inflection point at 70-80% LCPs is characterized by a large increase in battery capacity (Figure I-5(e)): Of the 33.4 GW of installed battery capacity at 95%, 26.1 GW is installed between 80% and 95%. As implied by Figure I-4, this buildout is due to the significant increase in excess low-carbon generation shown in Figure I-5(h). Furthermore, as battery capacity increases, battery energy throughput does not increase as much (Figure I-5(f)), resulting in battery LCOE growth from $117/MWh at 80% LCP to $198/MWh at 95% LCP (Figure I-5(g)). Similarly, gas-based generation capacity remains fairly steady even at very high LCPs, but the electricity generation from that capacity decreases significantly. The result is gas generation LCOE steadily increasing from $57/MWh at 40% LCP to $72/MWh at 70% LCP and accelerating to $260/MWh at 95% LCOE. It is worth noting that these results partially reflect the constraints of the model; they suggest that other technologies not included in SECTR-NY due to their non-competitive costs become beneficial in pushes to eliminate emissions from electricity generation. Regardless, these technology costs coupled with the significant increase in wind and solar LCOEs due to curtailment give a strong indication of the dominance of integration costs at high LCP.

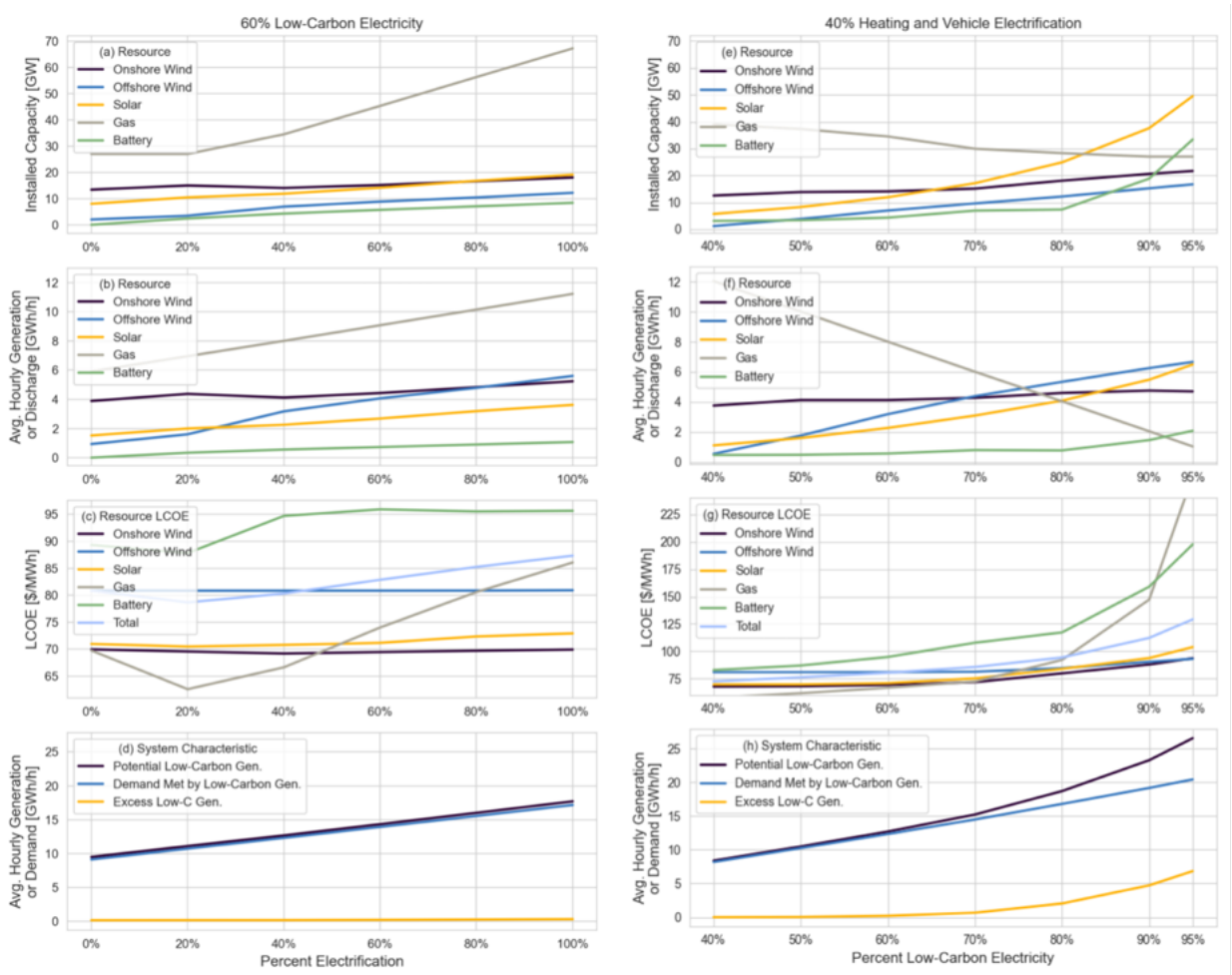

Figure I-5: System characteristics for scenarios with (a-d) increasing HVE at 60% LCP; and (e-h) increasing LCP at 40% HVE. Subplots (a, e) present installed capacity; (b, f) present average generation by resource; (c, g) present LCOE per MWh for the generation and storage resources; and (d, h) present demand and generation quantities. In (c, g), resource LCOE for onshore wind, offshore wind, and solar refers to the LCOE of generation; LCOE for battery storage is per-MWh discharge; total LCOE contains all system costs; and in (c), gas generation LCOE at 95% LCP ($260/MWh) is cropped out to preserve y-axis resolution.

Figure I-6 shows the monthly low-carbon electricity supply for (a) 60% LCP for HVEs of 0%, 40% and 80%, and (b) 40% HVE for LCPs of 60%, 80% and 95%. The seasonal low-carbon supply in Figure I-6(a) is nearly identical regardless of HVE and is largely in line with wind supply patterns shown in Figure I-2; this holds despite the low-carbon generation supply increasing 68% between HVEs of 0% and 80%. Accordingly, low-carbon electricity supply phenomena are shown to be

essentially independent of HVE, despite very significant shifts in diurnal and seasonal demand patterns with HVE. In contrast, Figure I-6(b) shows a significant shift in seasonal low-carbon supply behavior reflecting the increased share of solar shown in Figure I-5(f). (Additional system operation characteristics were investigated on this monthly timescale to inform the findings here; these are included in Figures I-11 and I-13.)

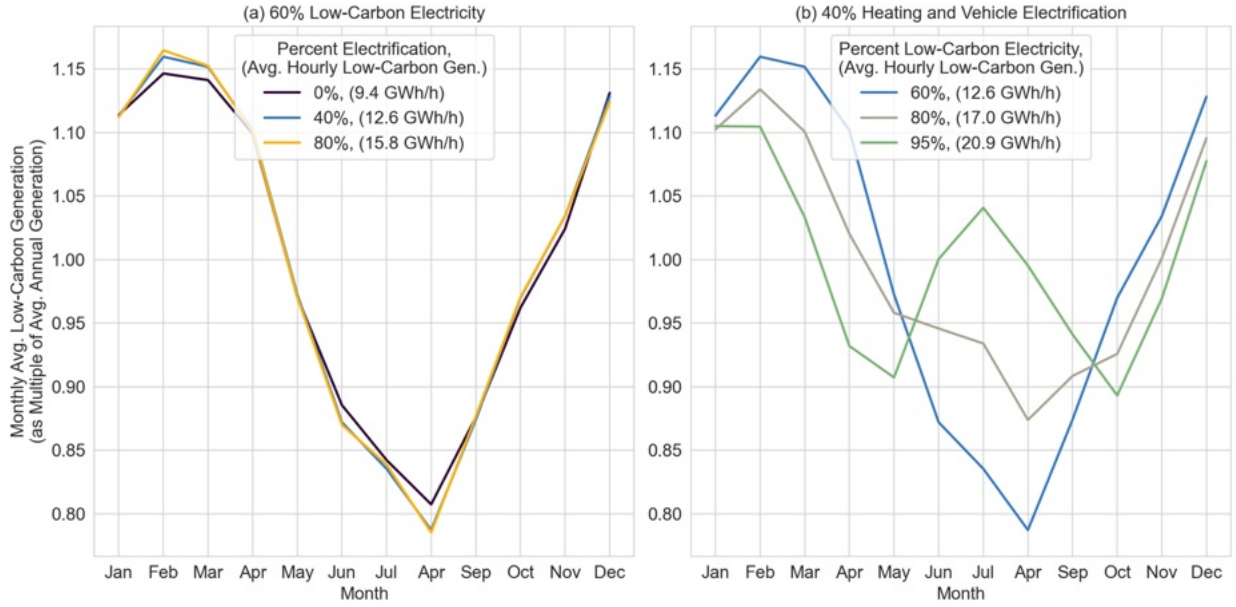

Figure I-6: Monthly average low-carbon generation as a multiple of the average annual low-carbon generation. (a) monthly averages for 0%, 40%, 80% HVEs at 60% LCP; (b) monthly averages for 60%, 80%, and 95% LCPs at 40% HVE.

Solar's contribution to the overall supply mix increases most dramatically beyond 80% LCP as battery storage increases: whereas 19.2 GW of solar capacity is installed between 40% and 80% LCP, 24.6 GW of capacity is installed just between 80% and 95% LCP (see Figure I-5(e)). This reflects complex dynamics in which overall system behavior may mask unique marginal behaviors of individual components: the operation of the same resource at lower LCP may be quite different with other resources present at higher LCPs. To this end, the paired buildout of

solar and battery capacity at very high LCPs provides the most cost-effective method of displacing the remaining gas generation, as the daily cycling of solar generation allows for regular battery charging during the day and discharging at night even as it becomes the highest LCOE renewable resource (Figure I-5(g)). Figure I-7 shows how battery behavior and its relation to wind and solar supply changes at increasing LCPs for a given 40% HVE. (See Figures I-17 and I-18 for other HVEs, which show the same trends as Figure I-7.)

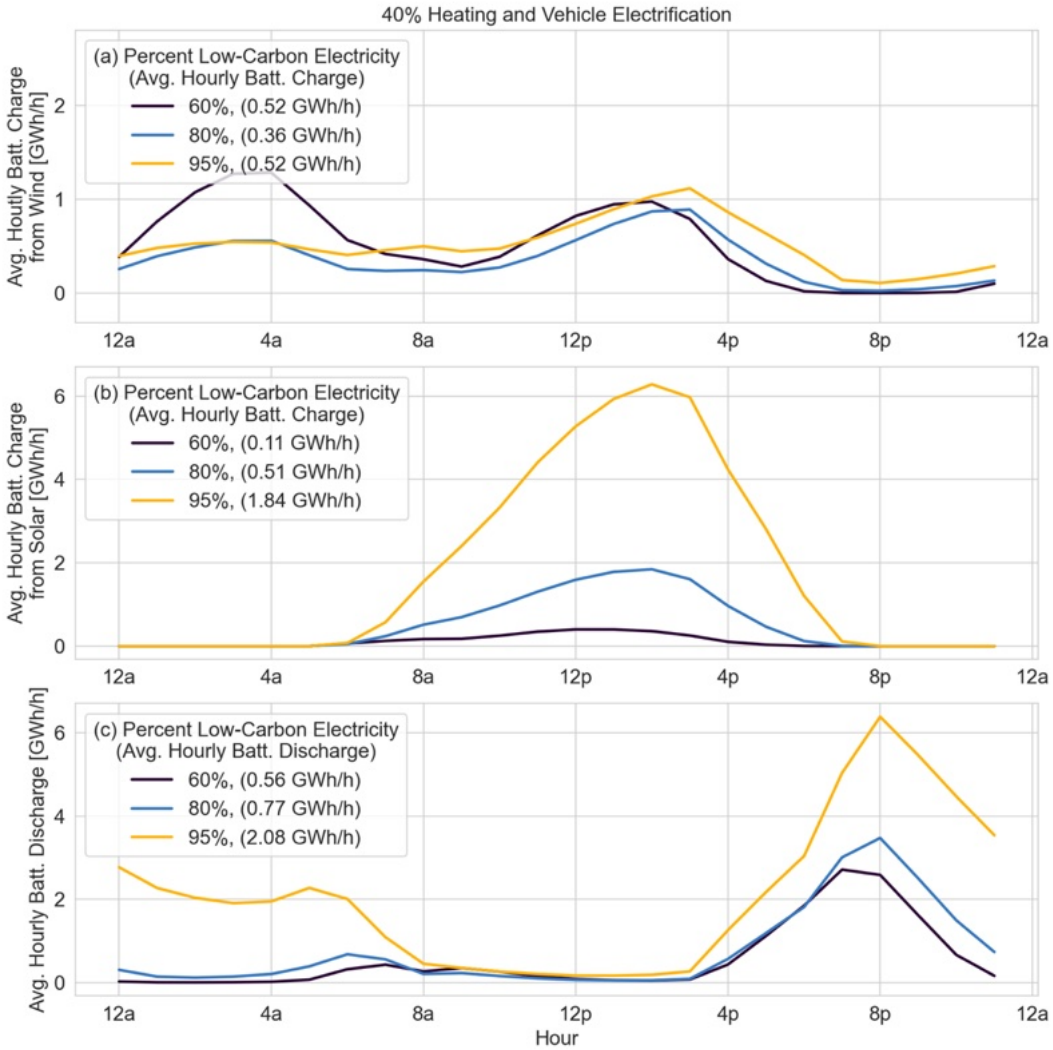

Figure I-7: Average battery operation by hour for 60%, 80%, and 95% LCPs over the entire 6-year simulation period. (a) average hourly battery charging from wind (note y-axis scale is unique from (b) and (c)); (b) average hourly battery charging from solar; and (c) average battery discharge, all in GWh/h.

At 60% LCP in Figure I-7, when total wind supply is roughly three times total solar supply, battery charging from wind is approximately 5 times higher than solar with distinct overnight and afternoon charging periods. At 80% LCP, wind's overnight charging reduces while both wind and solar charge the batteries in the afternoon; battery charging from solar becomes 1.4 times that from wind despite total wind supply being 2.4 times solar supply. Despite this shift between 60% and 80% LCP, the battery discharge remains almost entirely in the evening while total battery throughput increases by 38%. From 80% to 95% LCP, the maximum hourly average discharge in the evening doubles from 3 GWh/h to 6 GWh/h, extending throughout the night with a steady average 2-3 GWh/h supply resulting in a near tripling of the total throughput. The additional energy supply to the battery comes almost entirely from solar: While total wind supply remains 1.8 times the solar supply, battery charging from solar is 3.5 times that from wind. Here, the diurnal pattern of solar generation allows for daily battery cycling and higher battery throughput, behavior that enables the integration of more low-carbon generation.

While the average diurnal behavior shown in Figure I-7 is useful in understanding broad system behavior and the results of model decisions, decision-making is often based on complex dynamics occurring at hourly timescales over particular periods of time that set capacity and operational needs. Figures I-8 and I-9 show representative weeks in the winter and summer, respectively: The upper figures (Figures I-8(a) and I-9(a)) show scenarios of 80% LCP and 40% HVE, and the lower figures (Figures I-8(b) and I-9(b)) show scenarios of 95% LCP and 40% HVE. Figure I-8(a) shows that the lowest LCOE low-carbon option of wind provides much of the winter energy needs at 80% LCP, due to the resource's high seasonal productivity. Conversely, there are higher needs for gas-based generation in the summer (Figure I-9(a)). In both figures, curtailment

(i.e., slack in the SECTR-NY energy balance constraint) is attributed to solar and wind in proportion to their hourly generation; however, as noted in the discussion around Figure I-7, the natural pairing of solar generation and battery storage means that more wind generation is curtailed relative to solar.

As the LCP increases to 95% (Figures I-8(b) and I-9(b)), the reason for coupling more solar power with battery storage is revealed: Solar generation exceeding demand during the afternoon is used to charge battery storage, which is then discharged to meet evening demand (and overnight demand, if enough stored energy is available). In Figures I-8(b) and I-9(b), approximately 5% of demand met by gas generation occurs during extended hours of low wind production. Here, batteries are not as cost-effective in displacing gas generation: low wind generation potentials lasting a day or longer would require multi-day battery cycling periods, and accordingly, underutilization of storage capacity relative to its usage with solar. (For further exploration that reinforces this interpretation, Figures I-19 and I-20 present the same representative week and LCPs as Figures I-8 and I-9 but at 80% HVE.)

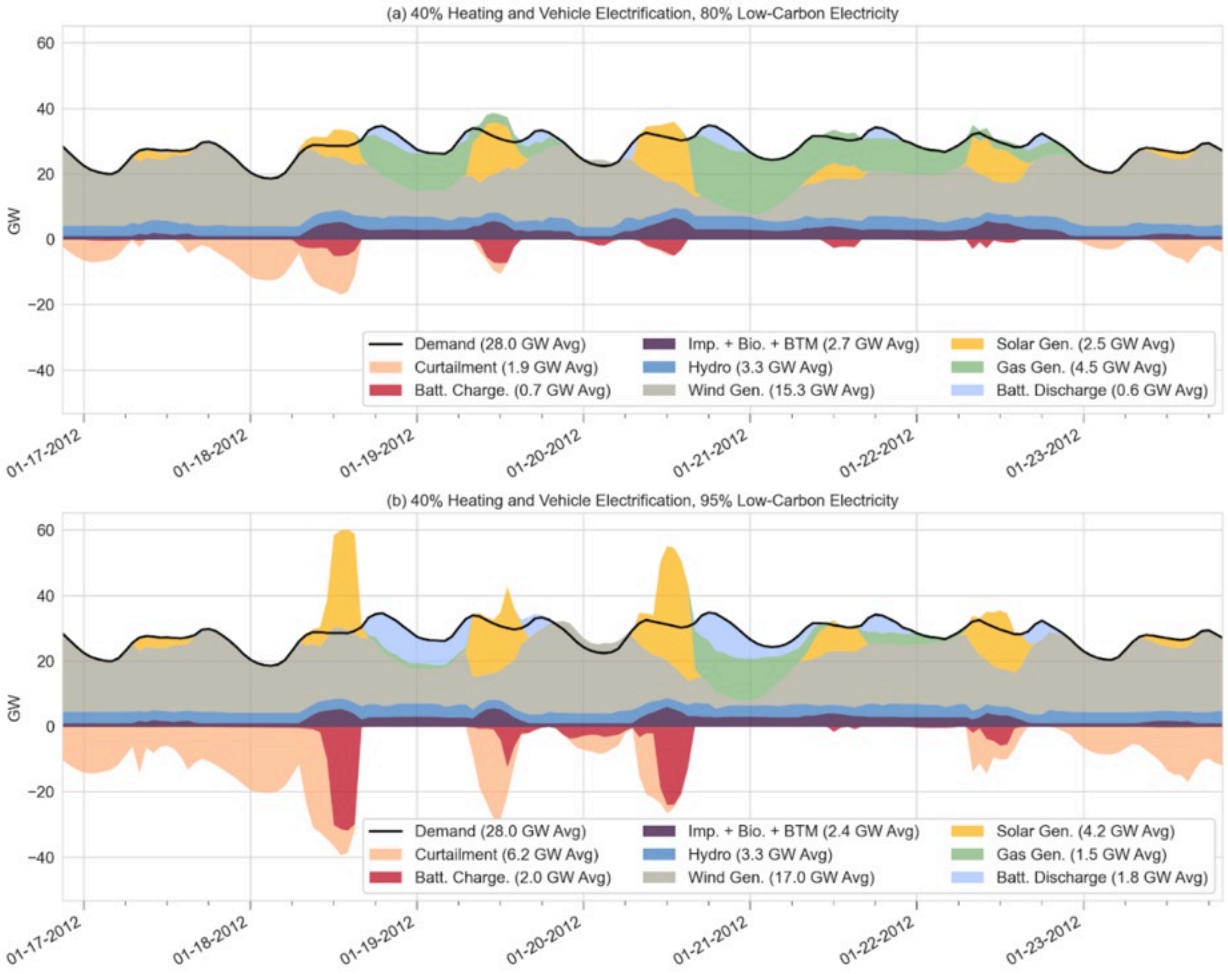

Figure I-8: Electricity generation and demand for a representative winter week with 40% HVE. (a) 80% LCP; (b) 95% LCP. 'Imp. + Bio. + BTM' represents the sum of imports, biofuel, and behind-the-meter solar generation. Average values reported in the legend are for the week shown.

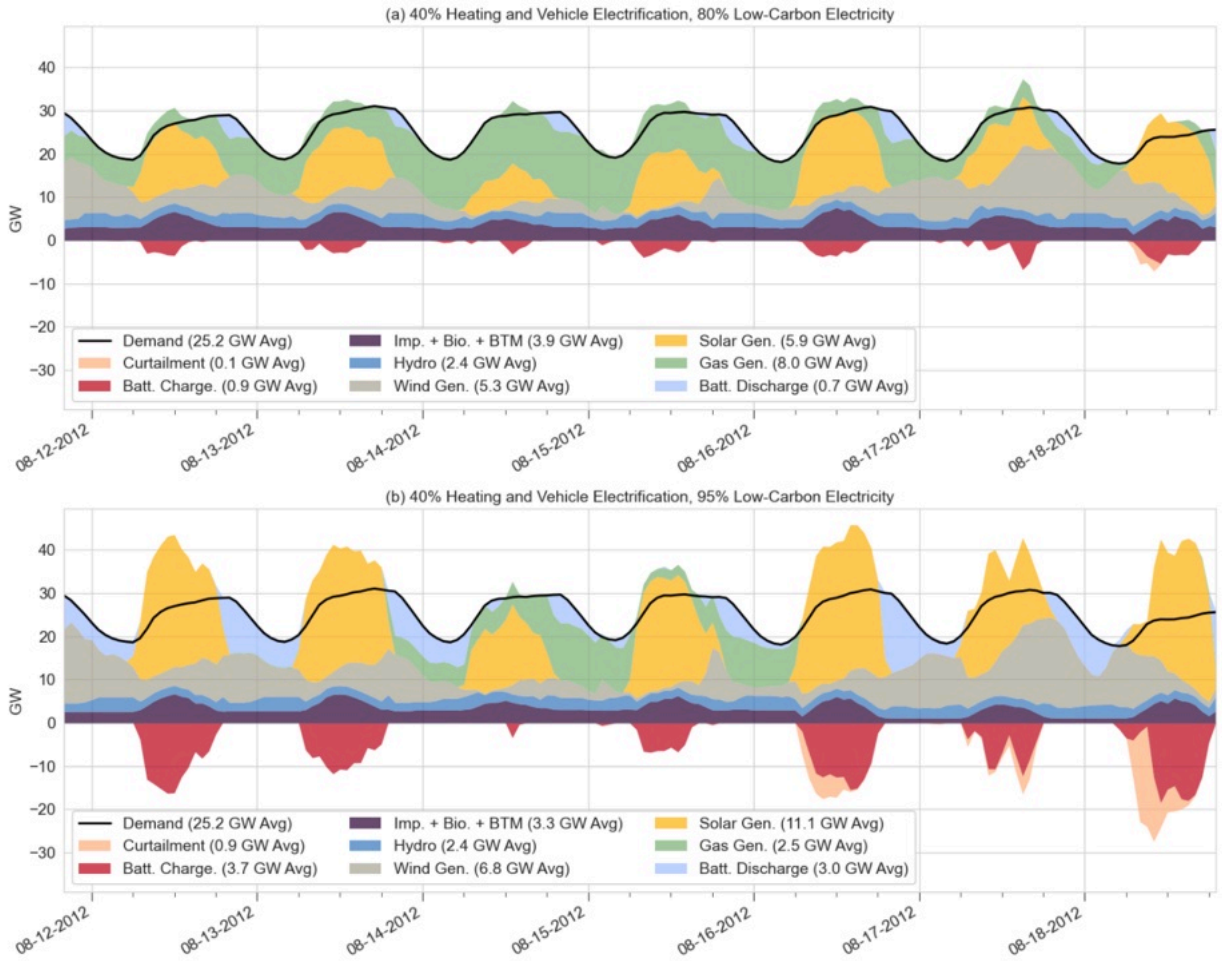

Figure I-9: Electricity generation and demand for a representative summer week with 40% HVE. (a) 80% LCP; (b) 95% LCP. 'Imp. + Bio. + BTM' represents the sum of imports, biofuel, and behind-the-meter solar generation. Average values reported in the legend are for the week shown.

3.3 Comparison to New York State policy studies

SECTR-NY model results are compared to initial analyses presented to the New York State Climate Action Council[xv], a committee preparing a scoping plan for CLCPA, both to validate SECTR-NY outputs and to evaluate the effects of different model assumptions and input data. A comparison of select characteristics of the NYS Climate Action Council Advisory Panel (AP) 2030 scenario and two SECTR-NY scenarios is shown in Table I-14. The AP 2030 scenario includes an 85% LCP and approximately 15% HVE[xvi] with a computed energy related GHG emissions reduction of 47.4% (relative to 1990, as are all GHG reductions discussed here); this scenario includes 28.4 GW of total wind and solar capacity and 3 GW battery storage capacity. For the same LCP and HVE, SECTR-NY Scenario A computes a total wind and solar capacity of 39.2 GW, 3.2 GW battery storage capacity, and GHG emissions reduction of 27.7%. There are two primary drivers for the greater SECTR-NY capacities here:

1. 14% higher average total wind and solar generation in SECTR-NY Scenario A (9.0 GW) than in AP 2030 (7.9 GW). This is due to more hydropower generation in AP 2030 than in the historical data used in SECTR-NY [100] and approximately 2.3 GW higher average statewide load in SECTR-NY Scenario A. The latter stems from a combination of SECTR-NY using historical electricity demand timeseries containing a higher existing average load (18.7 GW) than is simulated in NYS studies (18.2 GW); 15% SECTR-NY HVE likely being slightly higher than the estimate for AP 2030; AP 2030 considering combinations of

[xv] NYS published studies are available at the following link: https://climate.ny.gov/Climate-Resources. Technical analysis of initial results [32] and of key drivers and outputs [335] last updated in November and December of 2021 are of particular use in understanding the state's modeling methodology and simulated decarbonization pathways.
[xvi] The AP considered different electrification rates for different end-uses, so this estimate is not directly analogous to that of SECTR-NY presented here. See Table I-14, footnote 2 for a breakdown of the different electrification rates assumed in the AP recommendations.

population growth and efficiency savings; and SECTR-NY's more accurate representation of low-temperature effects of EHPs and EVs. (These low-temperature effects also explain the difference in fossil fuel-based generation capacity to meet the 35.4 GW peak statewide load computed in SECTR compared to the 29.6 GW statewide peak in AP 2030.)

2. 21% higher aggregate wind and solar capacity factor (CF) in AP 2030 (0.278) than in SECTR-NY Scenario A (0.230). This is primarily driven by significantly lower solar and onshore wind CFs in the latter. Model wind output in SECTR-NY is less than that of most available wind data: SECTR-NY employs a dataset that contains adjusted model data based on historical output of actual wind farms in NYS [64]. A comparison of solar data series was not performed; however, the authors believe SECTR-NY Scenario A's statewide solar CF of 0.166 represents more realistic expectations for NYS's latitude range than AP 2030's 0.194.

The difference in computed GHG reductions between AP 2030 and SECTR-NY Scenario A stems from model assumptions related to methane leakage in natural gas production and transport upstream of NYS. SECTR-NY relies on research on natural gas leakage [108,111] that estimates approximately 3.6% leakage with an associated impact on fossil fuel emissions factors [106]. AP 2030 reduces the leakage to approximately 2%, though the authors have not seen an explanation for this assumption. The implications of these assumptions can be seen in SECTR-NY Scenario B, in which more heating and vehicle electrification is needed to achieve the same percentage GHG emissions reduction as that computed for AP 2030. Here, total computed wind and solar capacity increases to 51.4 GW, 81% greater than that anticipated by the recent analyses presented to the NYS Climate Action Council.

Table I-14: Comparison of NYS Climate Action Council Advisory Panel (AP) recommendations and SECTR-NY simulation results for modeled 2030 decarbonization scenarios.

	Modeled Scenario		
	NYS AP 2030	SECTR-NY, A	SECTR-NY, B
Low-Carbon Electricity Percent (LCP)	85%	85%[1]	85%[1]
Heating and Vehicle Electrification (HVE)	15%[2]	15%[1]	50%
GHG Emissions Change (Compared to 1990)	-47.4%	-27.7%	-47.4%[1]
Electricity Demand Peak [GW] | Average [GWh/h]	29.6 |18.4	35.4 | 20.7	52.7 | 25.4
Onshore Wind Capacity [GW] | Average Generation [GWh/h]	5.2 | 1.7	11.2 | 2.6	14.2 | 3.4
Offshore Wind Capacity [GW] | Average Generation [GWh/h]	6.2 | 2.9	8.4 | 3.8	13.2 | 5.8
Solar Capacity [GW] | Average Generation [GWh/h]	17.0 | 3.3	19.6 | 2.7	24.0 | 3.9
In-State Hydropower Capacity [GW] | Average Generation [GWh/h]	4.6 | 3.5	5.3 | 3.0	5.3 | 3.0
Hydropower Imports Capacity [GW] | Average Generation [GWh/h]	2.7 | 2.2	2.8 | 2.0	2.8 | 2.0
Nuclear Capacity [GW] | Average Generation [GWh/h]	3.4 | 3.0	3.5 | 3.2	3.5 | 3.2
Battery Capacity [GW]	3.0	3.2	9.9
Fossil Fuel Capacity [GW] | Average Generation [GWh/h]	20.8 | 2.7	27.0 | 2.5	27.6 | 3.2

[1] Indicates configuration parameters specified for the SECTR-NY model scenario.

[2] Approximated from the following proportions of vehicle and building stock end-use equipment transitioning to electric alternatives in the AP 2030 scenario: 14% of light duty vehicles, 6% of heavy duty vehicles, 11% of residential space heating; 11% of commercial space heating, 25% of residential water heating, and 19% of commercial water heating.

3.4 Additional Baseline configuration results

Figure I-10 presents an analogous plot to Fig. I-3(a), but with a continuous 3.2 GWh/h of upstate nuclear generation present. Here, nuclear generation allows for approximately 10% lower LCOEs on average at the simulated scenarios, cost savings that grow larger at higher LCPs. However, the addition of nuclear generation does not change the overall shape of Figure I-3(a), and accordingly the same conclusions are reached: 1) Emissions reductions can be achieved at lower LCOEs by prioritizing electrification of heating and vehicles in conjunction with deployment of solar and wind, as opposed to the latter by itself, and 2) system costs increase substantially above 70-80% LCPs.

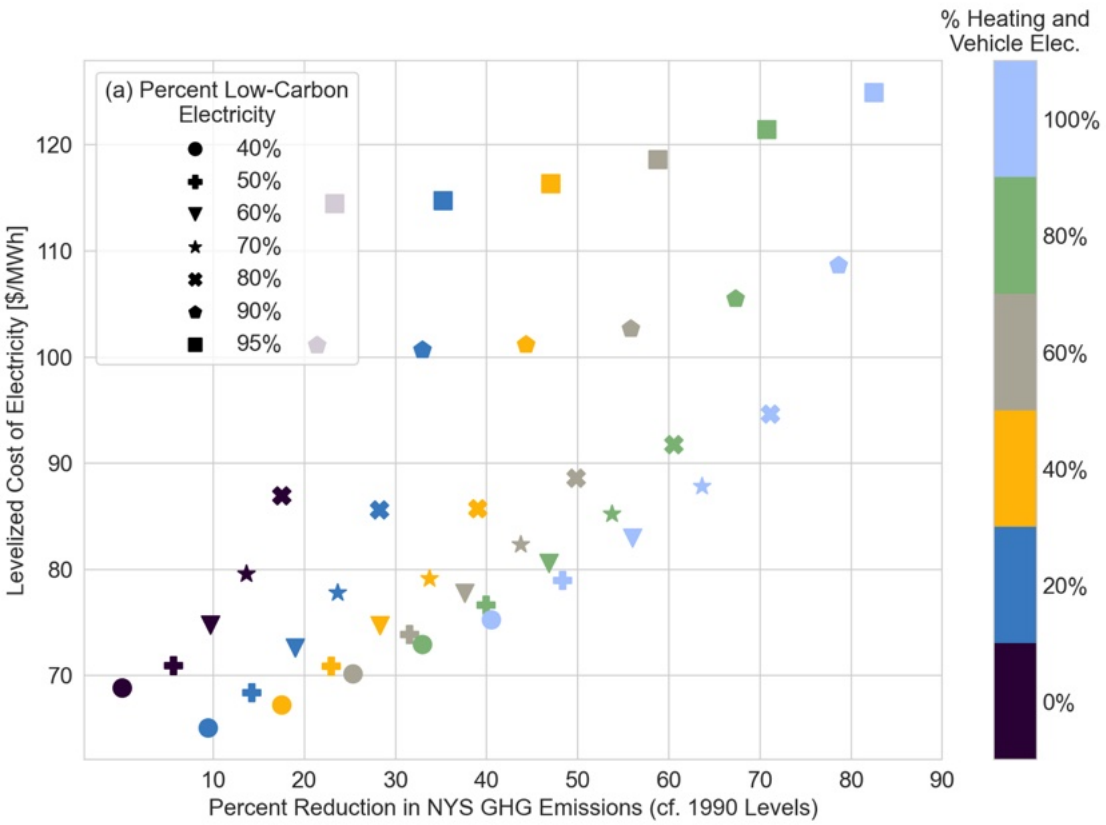

Figure I-10: LCOE vs. percent reduction in NYS GHG emissions (compared to 1990 levels). Marker shape indicates percent low-carbon electricity (LCP), and marker color indicates heating and vehicle electrification (HVE). For scenarios shown, all low-carbon electricity generation is from wind, solar, nuclear, and hydropower.

Next explored are the effects of either increased HVE or LCP on peak gas generation, average gas generation, low-carbon electricity generation, and battery storage throughput. In evaluating the peak gas generation characteristics, increasing electrification at a set LCP results in substantial winter peaks: Figure I-11(a) presents the monthly peak to annual average gas generation ratio at 60% LCP for 0%, 40% and 80% HVE. At 80% HVE, additional, peaky heating demand causes January gas generation peaks of 46.9 GWh/h, equal to 4.6 times the annual average, compared 15.9 GWh/h at 0% HVE with a peak-to-average ratio of 2.7. In contrast, the July peak only increases from 22.4 GWh/h at 0% HVE to 25.5 GW at 100% HVE. Figure I-11(b) shows that there are no equivalent seasonal effects to increasing the LCP at 40% electrification. However, increasing the LCP to 80% and 95% results in lower average gas generation (4.0 GWh/h and 1.0 GWh/h, respectively, compared to 8.0 GWh/h), quantities which result in substantial peak-to-average ratios (above 20 in December and January for the 95% LCP).

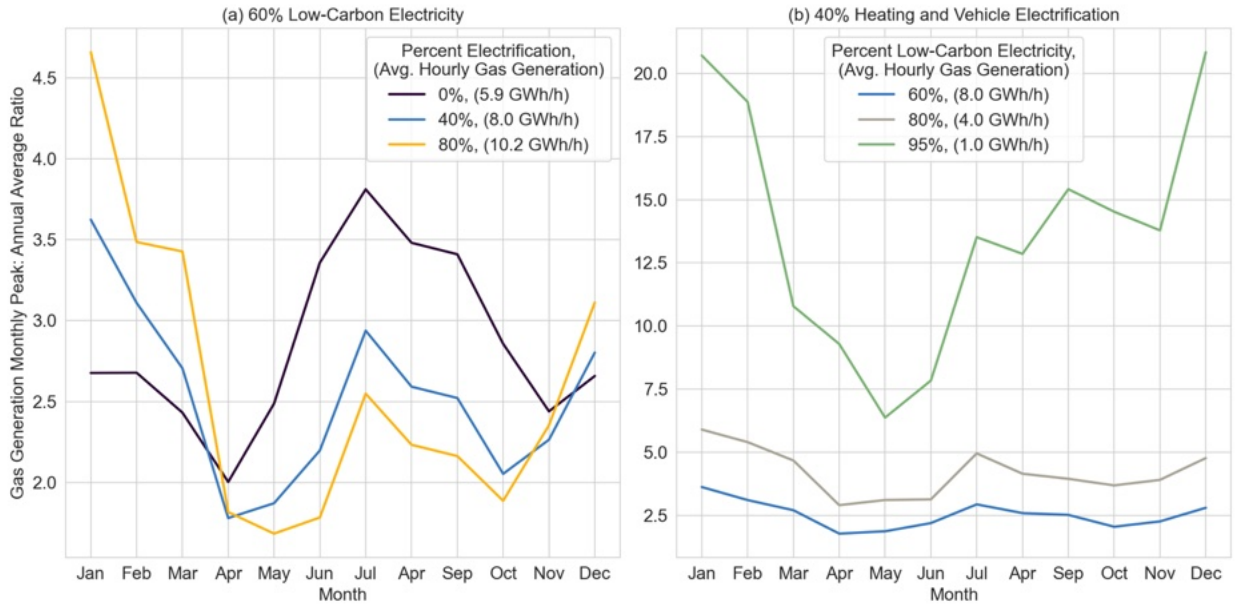

Figure I-11: Monthly peak to annual average gas generation ratios for (a) scenarios containing 60% low carbon electricity with increasing amounts of electrification; and (b) scenarios containing 40% electrification with increasing percents low-carbon electricity.

Increasing electrification at a set LCP has a similar seasonal shift on average gas generation, shown in Figure I-12(a). For the same 60% LCP and 0%, 40%, and 80% HVEs, increased electrification results in higher average winter gas generation – in absolute terms and relative to the annual average – and lower relative generation during the summer. In January, 0% HVE corresponds to an average 5.3 GWh/h of gas generation, or 0.9 times the annual average; 100% HVE increases this to 16.3 GWh/h, or 1.6 times the annual average. Again, this increase in average generation is attributable to the higher amounts of peaky heating demand on the system: Heating demand proves difficult to meet with low-carbon electricity and is accordingly satisfied by dispatchable gas generation. The suitability of gas generation in meeting electrified heating demand also explains the relative decreases in gas generation during summer months. As the same LCP needs to be achieved despite increased winter gas generation, gas generation during the summer is reduced (1.2 times the annual average with 80% HVE compared to 1.8 times at 0% HVE in the month of July), as this less-peaky demand can more easily be met by a combination of solar generation and battery storage.

Figure I-12(b) demonstrates that raising the LCP from 60% to 95% increases the January gas generation from 1.3 to 2.5 times the annual average, a shift that indicates the costliness of meeting electrified heating demand with only low-carbon generation and battery storage. In contrast, gas generation in the shoulder seasons is the first to be displaced by low-carbon generation, due to 1) the high productivity of onshore wind, offshore wind, and solar resources, and 2) the lack of peaky heating demand during these months.

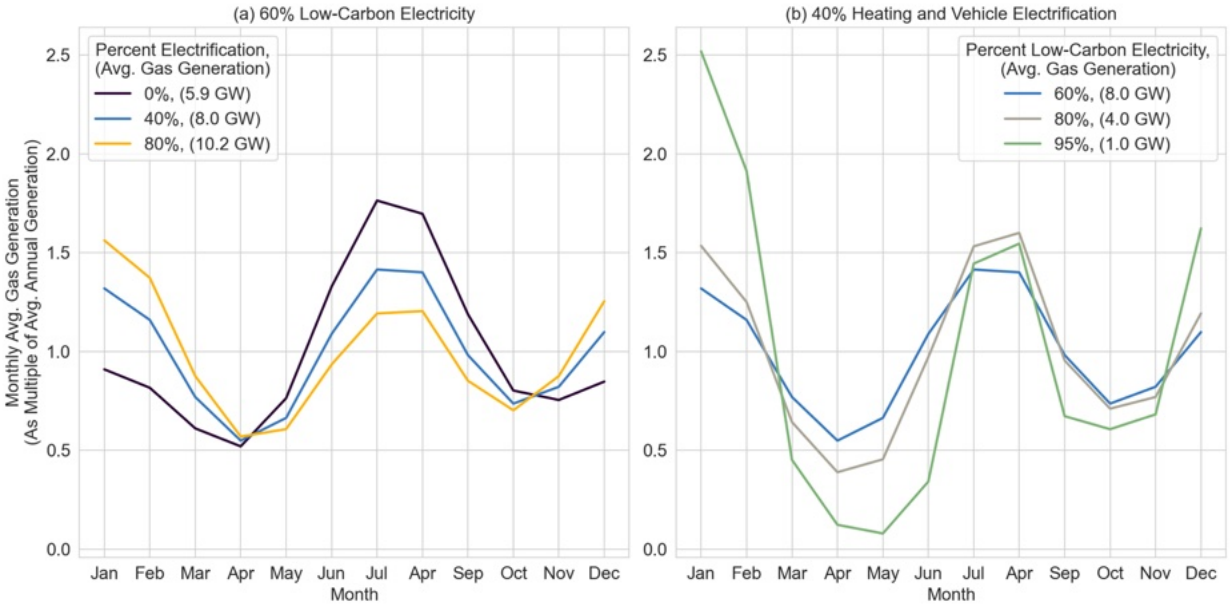

Figure I-12: Monthly gas generation as a multiple of the annual average for scenarios containing (a) 60% low-carbon electricity with increasing amounts of electrification; and (b) 40% electrification with increasing percents low-carbon electricity.

Evaluation of monthly battery storage behavior reinforces the findings presented in Section I-3.2. Increasing electrification at 60% low-carbon electricity shifts battery throughput towards summer months when battery storage is well-paired with the daily cycles of productive solar generation, per Figure I-13(a). While this relative seasonal shift is apparent in the changing shapes of the normalized throughput curves, the absolute seasonal difference in battery throughput is not as stark: Increasing HVE from 0% to 80% only raises battery throughput by an average 1.0 GWh/h, indicating that battery output is not utilized to meet a significant portion of demand at 60% LCP. In contrast, battery throughput increases substantially in the summer months – in both absolute and relative terms – and experiences a relative drop during the shoulder seasons as LCP increases from 60% to 95% at 40% HVE (Figure I-13(b)). At 95% LCP, battery throughput reaches an average of 3.1 GWh/h in August (1.4 times the annual average), a

quantity that is double the average throughput in April (1.6 GWh/h); to compare, the 60% LCP scenario contains average throughputs in August and April both roughly equal to the annual average of 0.6 GWh/h. From this figure, one concludes that pairing batteries with productive solar generation during summer months provides a cost-effective method of meeting additional load with low-carbon electricity. It is also notable that this effect is substantially greater when increasing the LCP at a given HVE, due to the greater amounts of excess low-carbon generation present in these scenarios.

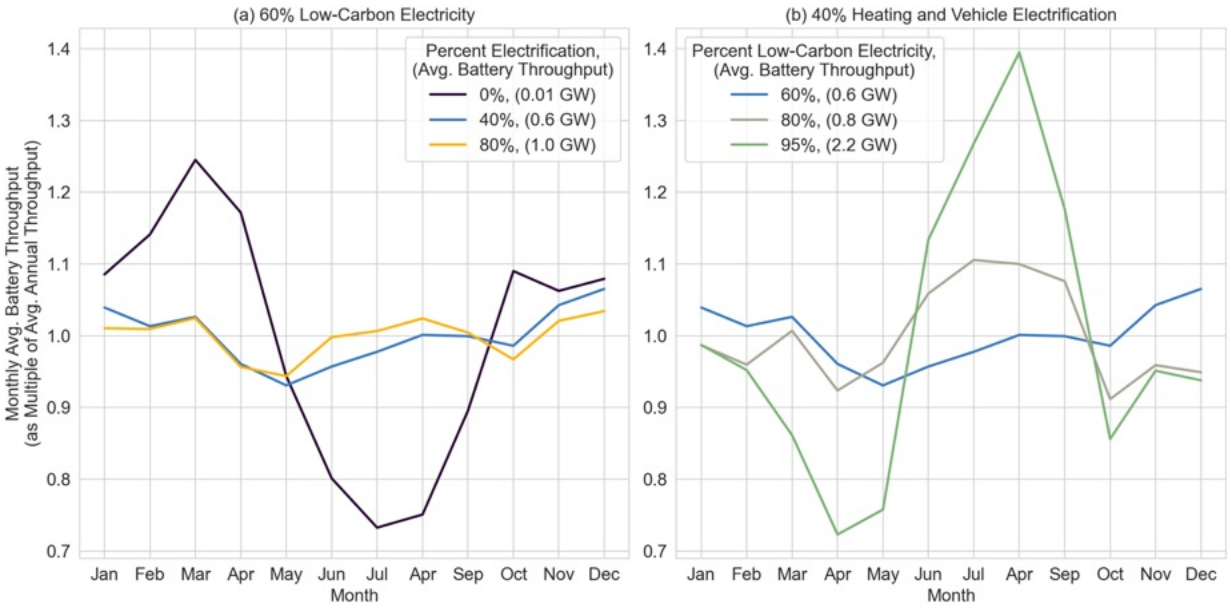

Figure I-13: Monthly battery throughput as a multiple of average annual throughput. Results are presented for scenarios containing (a) 60% low carbon electricity with increasing amounts of electrification; and (b) 40% electrification with increasing percents low-carbon electricity.

3.5 Impact of existing system parameterization

To understand the impact of SECTR baseline parameters and how different parameterizations affect model results, two additional configurations are evaluated: A 'Greenfield' configuration and a 'Greenfield with Constant Costs' configuration. The Greenfield

configuration represents a type of parameterization often seen in the capacity expansion modeling literature: This configuration includes no existing solar, wind, gas, biofuel, or transmission capacity; and no existing biofuel generation. The Greenfield with Constant Costs configuration combines the greenfield parameterization with homogenous nodal costs, which are calculated via a weighted average of the costs associated with the returned capacity and generation quantities from the Greenfield configuration model solution.

All configurations are evaluated at two scenarios, one representing a combination of a high LCP and a low HVE (referred to as the high LCP scenario), and the other representing a combination of a lower LCP and a higher HVE (referred to as the low LCP scenario). In both scenarios, the GHG reduction is set to 40%. For the high LCP scenario, electrification of heating and transport is set to 40% and the LCP is determined by the model; for the low LCP scenario, LCP is set to 60% and the HVE is determined by the model. To ensure equivalent LCPs across configurations, the efficiency of new gas turbines in a Greenfield-based configuration is set to the weighted average efficiency of existing and new generation in the corresponding Baseline scenario.

Table I-15 presents a comparison of model-selected gas, battery, transmission, and LCOE characteristics. Here, both Greenfield configurations (with and without constant costs) contain LCOEs approximately 10% lower than in the fully parameterized Baseline configuration, regardless of the combination of HVE/LCP. As the Greenfield configurations do not include any existing gas, biofuel, or transmission capacity, the fixed costs associated with maintaining this infrastructure (see 'Existing Cap. LCOE' column) drop to nearly $0/MWh, a reduction that causes the total LCOE decline. Moving from the Baseline to the Greenfield configuration, an average

60.2% decline in total installed transmission capacity is observed across both scenarios, with reverse transmission being completely eliminated; moving to the Greenfield with Constant Costs configuration causes an average 69.2% decline. Accordingly, the transmission capacity that is installed in the positive direction is utilized more frequently, a trend which is particularly pronounced in the Greenfield configuration results, due to lower amounts of installed downstate gas generation (see Figure I-14).

Table I-15: Select computed characteristics of Baseline and Greenfield configurations.

	Configuration Parameters			Model-returned Generation and Storage Capacities (Cap.) and Transmission (Tx.) Characteristics							
Configuration	% GHG[a]	% HVE[b]	% LCP[b]	Total Gas Cap. [GW]	Total Upstate Battery Cap. [GW]	Total Downstate Battery Cap. [GW]	Total Pos. Tx. Cap. [GW-mi][c]	Total Rev. Tx. Cap. [GW-mi][d]	Avg. Pos. Tx. Util. %[c]	Existing Cap. LCOE [$/MWh][e]	Total LCOE [$/MWh]
Baseline	-40	40	81.7	27.2	7.3	1.2	2646.8	2287.4	24.3	27.0	96.4
Baseline	-40	64.8	60	47.2	2.8	3.9	2646.8	2083.2	28.1	23.7	83.5
Greenfield	-40	40	81.3	26.1	8.8	3.5	2473.3	0.0	31.7	0.8	86.5
Greenfield	-40	64.3	60	48.4	2.6	3.9	1371.0	0.0	41.7	0.7	75.8
Greenfield w. Constant Costs	-40	40	81.3	25.7	10.9	2.2	2339.8	0.0	27.3	0.8	86.7
Greenfield w. Constant Costs	-40	64.4	60	48.0	5.7	2.7	629.0	5.3	29.9	0.7	75.0

[a] '% GHG' refers to the percent change in greenhouse gas emissions compared to the 1990 reference values. Negative values indicate reductions.
[b] LCPs/HVEs are not identical across configurations due to slight differences in model-computed electricity imports given the specified GHG reduction and the HVE/LCP.
[c] 'Pos.' refers to "positive" upstate-to-downstate transmission directionality, i.e. from Node 1 to 2, Node 2 to 3, and Node 3 to 4.
[d] 'Rev.' refers to "reverse" downstate-to-upstate transmission directionality, i.e. from Node 4 to 3, Node 3 to 2, and Node 2 to 1.
[e] The costs of maintaining existing gas, hydropower, biofuel, and transmission capacity constitute the cost portion of 'Existing Capacity LCOE.'

For the high LCP scenarios in the two greenfield configurations, less transmission capacity and lower amounts of installed gas generation are compensated by increased battery capacity: The Greenfield configuration contains 2.8 GW additional storage capacity (a 29.4% increase), while the Greenfield with Constant Costs configuration contains 3.6 GW additional storage capacity (a 37.9% increase). This larger quantity of installed battery capacity is less prominent in the low LCP scenarios, due to their lower need for low-carbon electricity shifting; however, the low-LCP scenario in the Greenfield with Constant Cost configuration contains 1.7 GW more battery capacity than its Baseline analogue, an increase of 25.4%

Figure I-14 displays the change in gas capacity and generation characteristics across the three configurations. Here, the spatial heterogeneity of SECTR-NY results is investigated by splitting NYS into upstate and downstate regions[xvii]. Upstate NYS contains the state's onshore wind capacity, low-cost utility-scale solar, and existing low-carbon generation, while downstate NYS contains substantial electricity demand in and around New York City and offshore wind capacity. These differences in regional characteristics results in distinct system behavior on either side of the interface between Nodes 2 and 3.

Comparing results across configurations, the top row – representing the high LCP scenario – contains a 7.1 GW shift in gas capacity from downstate to upstate nodes when changing from the Baseline to the Greenfield configuration, due to the relatively higher cost of downstate gas capacity. Adopting constant nodal costs causes a smaller shift: When all new capacity has the same cost, the high LCP scenario shifts 1.8 GW gas capacity towards downstate regions compared to its equivalent Baseline configuration. In the low LCP scenarios (bottom row of Figure I-14), a

[xvii] 'Upstate' is defined as a region containing Nodes 1 and 2; 'downstate' refers to a combination of Nodes 3 and 4.

consistent shift from upstate to downstate gas generation capacity is observed: The Greenfield configuration contains a shift of 5.4 GW, while the Greenfield with Constant Costs configuration contains a shift of 12.0 GW.

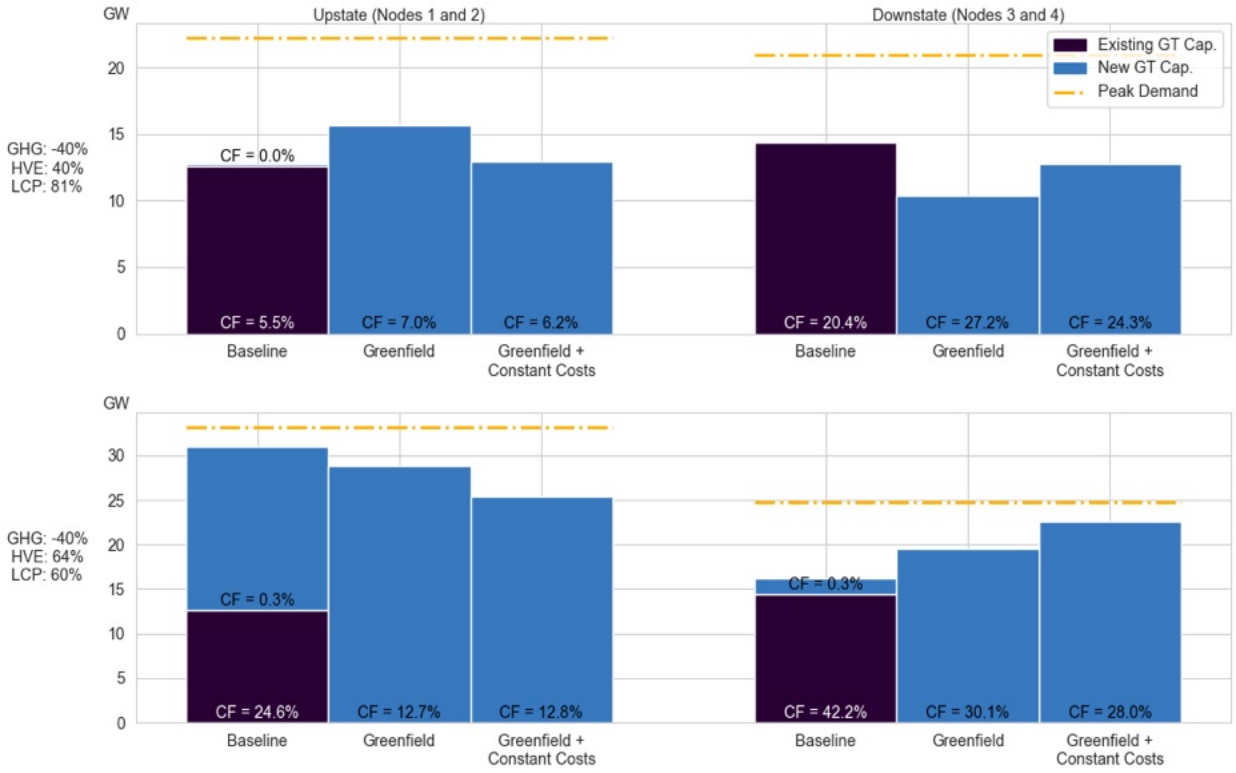

Figure I-14: Existing and new gas capacity, distribution, and capacity factors (CF), shown with peak demand, for upstate and downstate New York State regions. The top row presents results for the high low-carbon electricity percent (LCP) scenario; the bottom row presents results for the LCP scenario.

Both scenarios reveal the low capacity factors (CFs) of gas capacity in energy systems that achieve 40% GHG reduction, regardless of the configuration. In the top row, the high percent low-carbon electricity means that gas generation meets 19% of demand; this corresponds to capacity factors less than 7% upstate and less than 28% downstate. Here, CFs are lower upstate as this where the bulk of the renewable generation capacity is located. In comparison, gas

generation CFs are higher on average for the low LCP scenario despite the larger amounts of GT capacity required to meet the additional electrified load: The looser low-carbon electricity constraint means that gas generation can satisfy approximately 40% of the demand. The outlier to this trend is the new gas capacity installed upstate in the Baseline configuration. For this scenario, 18.4 GW of new upstate capacity generates an average of 62.7 MWh/h, and 1.8 GW of new downstate capacity generates an average of 5.7 MWh/h, both corresponding to rounded CFs of 0.3%.

3.6 Additional figures presented at different rates of heating and vehicle electrification and different percents low-carbon electricity

Figures I-15 and I-16 display versions of Figure I-5 at different HVEs and LCPs; Figures I-17 ad I-18 display versions of Figure I-7 at different HVEs; and Figures I-19 and I-20 display versions of Figures I-8 and I-9 at different HVEs. These figures demonstrate that the results presented in Section I-3.2 are not unique to the selected percents low-carbon electricity or electrification rates therein.

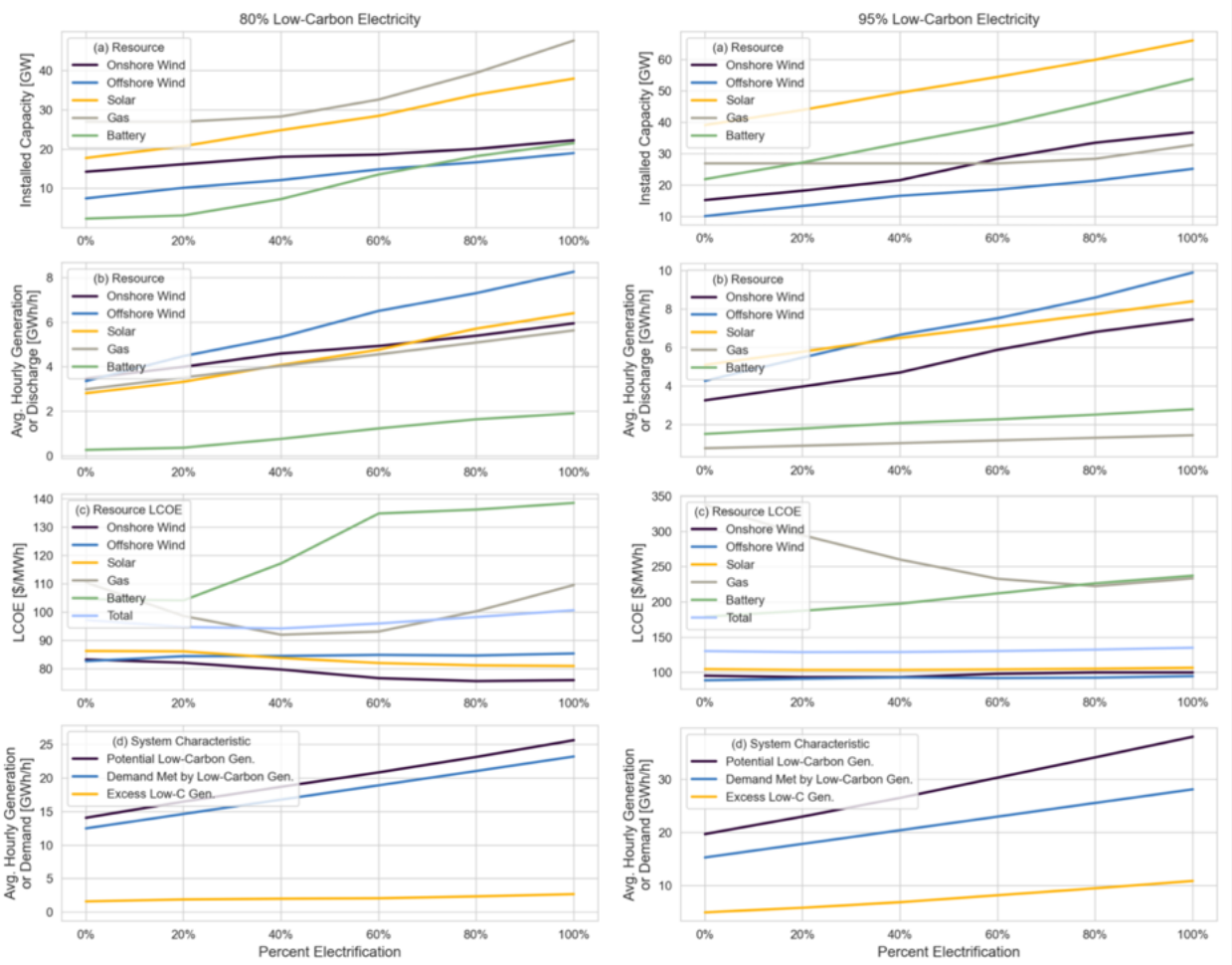

Figure I-15: System characteristics for scenarios with (a-d) increasing HVE at 80% LCP; and (b) increasing HVE at 95% LCP. Subplots (a, e) present installed capacity; (b, f) present average generation by resource; (c, g) present LCOE per MWh for the generation and storage resources; and (d, h) present demand and generation quantities. In (c, g), resource LCOE for onshore wind, offshore wind, and solar refers to the LCOE of generation; LCOE for battery storage is per-MWh discharge. Note the different y-axis ranges for side-by-side panels.

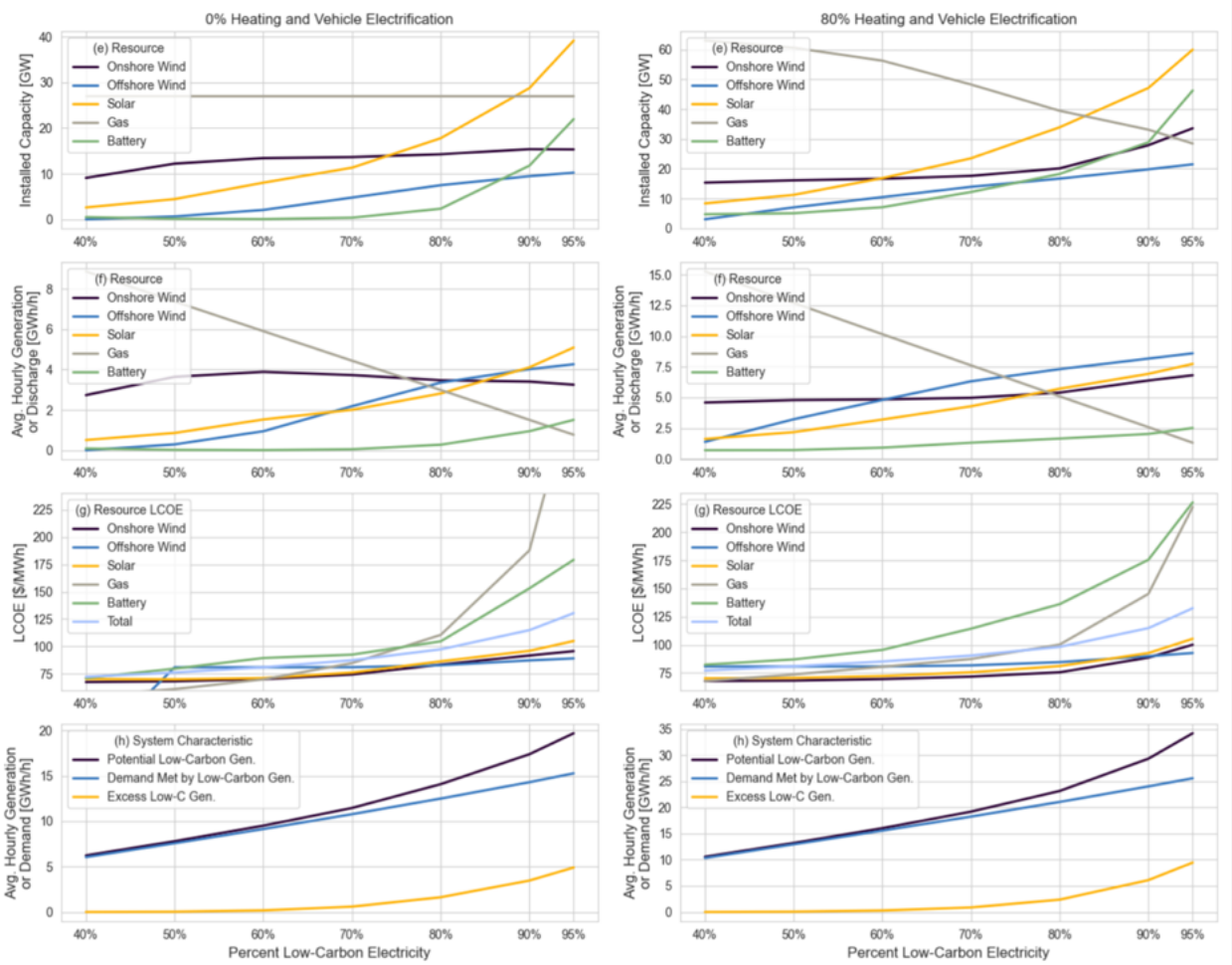

Figure I-16: System characteristics for scenarios with (a-d) increasing LCP at 0% HVE; and (b) increasing LCP at 80% HVE. Subplots (a, e) present installed capacity; (b, f) present average generation by resource; (c, g) present LCOE per MWh for the generation and storage resources; and (d, h) present demand and generation quantities. In (c, g), resource LCOE for onshore wind, offshore wind, and solar refers to the LCOE of generation; LCOE for battery storage is per-MWh discharge; and in (c), gas generation LCOE at 95% LCP ($338/MWh) is cropped out to preserve y-axis resolution at lower LCOE values. Note the different y-axis ranges for side-by-side panels.

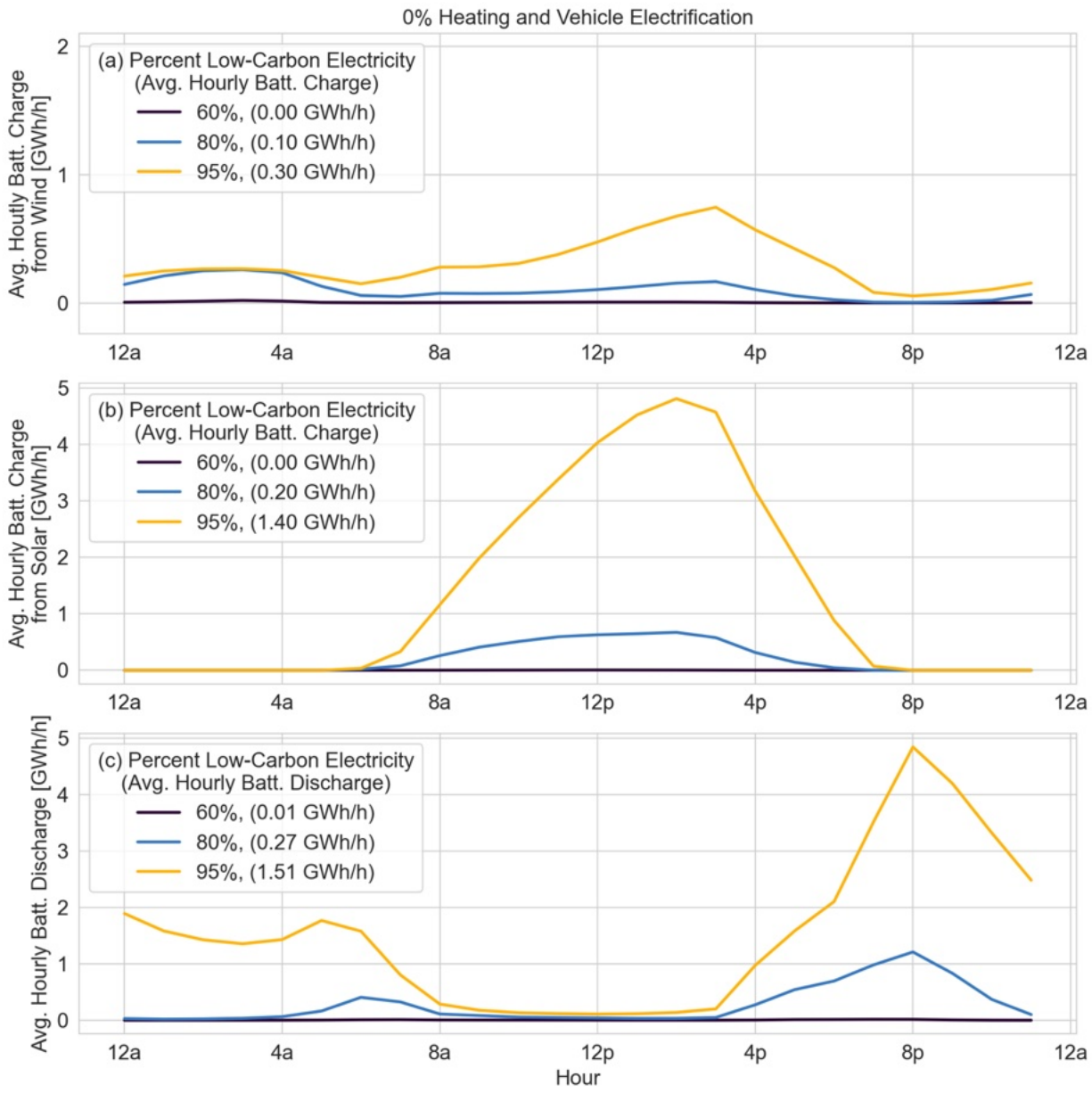

Figure I-17: Average battery operation by hour for 60%, 80%, and 95% LCPs at 0% HVE. (a) Average hourly battery charging from wind (note y-axis scale is unique from (b) and (c)); (b) average hourly battery charging from solar; and (c) average battery discharge, all in GWh/h.

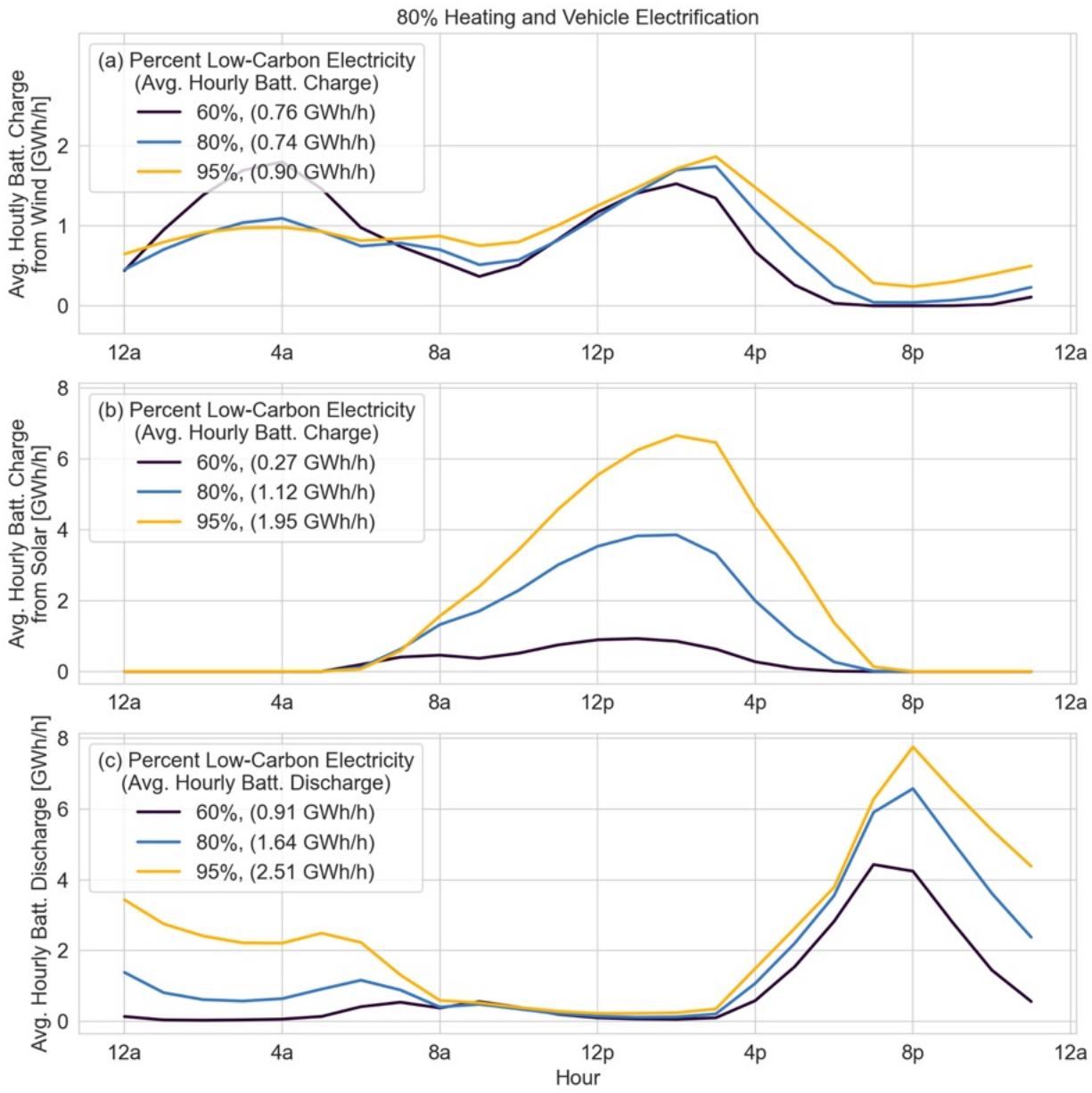

Figure I-18: Average battery operation by hour for 60%, 80%, and 95% LCPs at 80% HVE. (a) Average hourly battery charging from wind (note y-axis scale is unique from (b) and (c)); (b) average hourly battery charging from solar; and (c) average battery discharge, all in GWh/h.

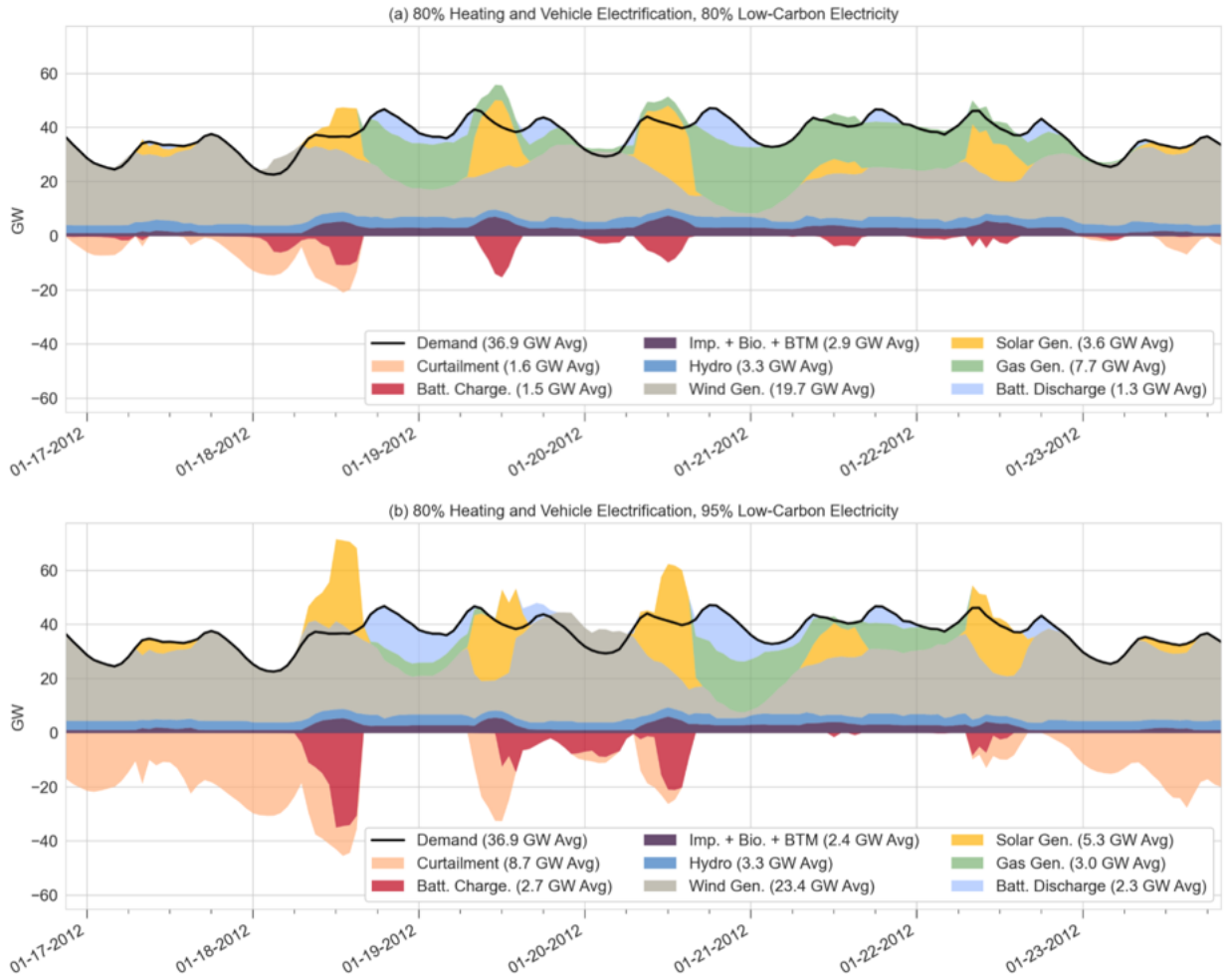

Figure I-19: Electricity generation and demand for a representative winter week with 80% HVE. (a) 80% LCP; (b) 95% LCP. 'Imp. + Bio. + BTM' represents the sum of imports, biofuel, and behind-the-meter solar generation. Average values reported in the legend are for the week shown.

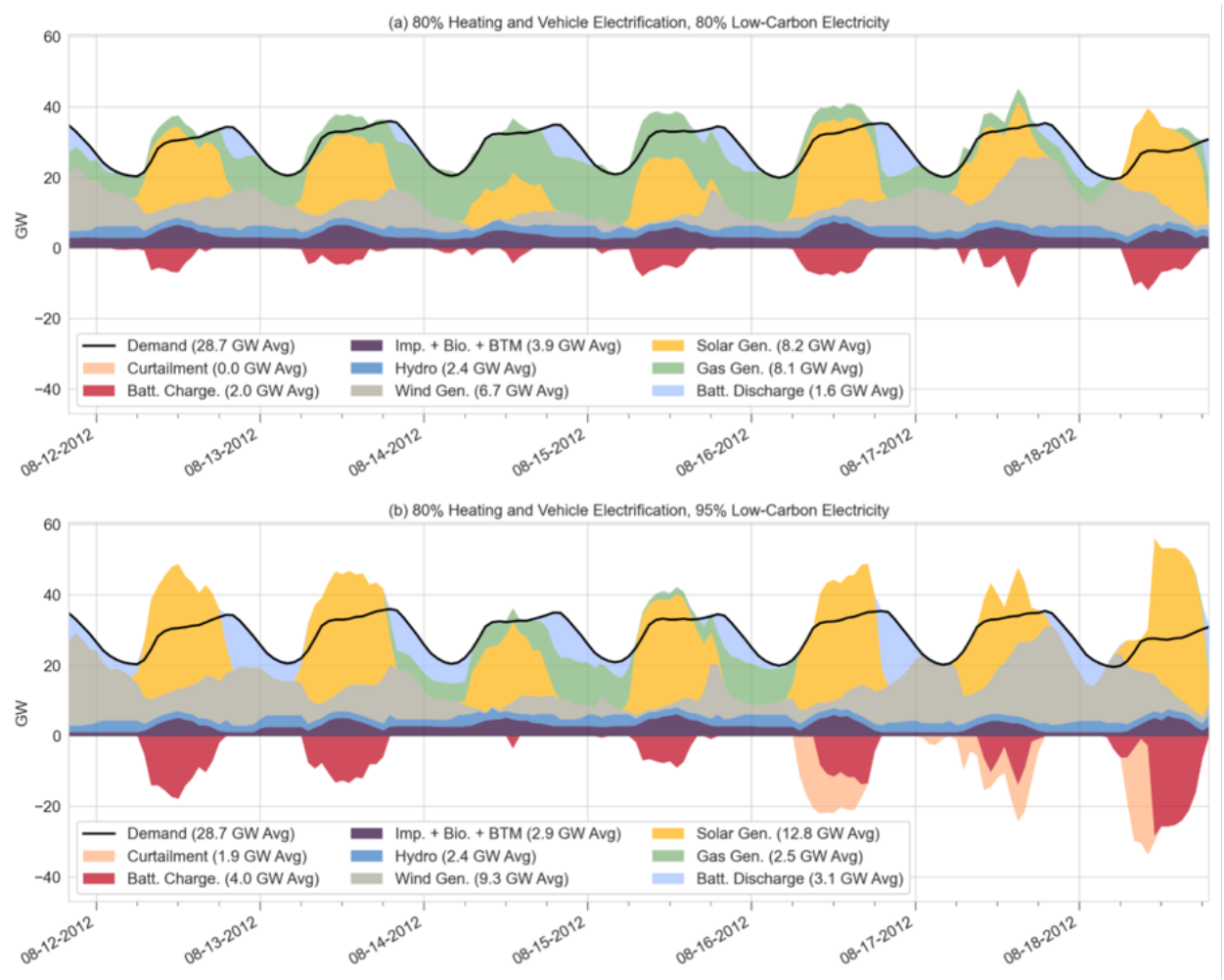

Figure I-20: Electricity generation and demand for a representative summer week with 80% HVE. (a) 80% LCP; (b) 95% LCP. 'Imp. + Bio. + BTM' represents the sum of imports, biofuel, and behind-the-meter solar generation. Average values reported in the legend are for the week shown.

4. Discussion

This study's results are broadly consistent with previously published research that deep greenhouse gas (GHG) emissions reductions require both a significant low-carbon electricity percentage (LCP) and increases in heating and vehicle electrification (HVE); however, an important finding is that by prioritizing heating and vehicle electrification in conjunction with renewable energy deployment rather than first focusing on LCP, emissions reductions can be achieved with lower electricity supply costs. Through comparative scenarios, the benefits of end-use electrification to the electricity system are emphasized: Heating and vehicle electrification allows the same amount of renewable energy to be installed with significantly lower electricity supply costs all while producing deeper reductions in GHG emissions.

First order GHG reductions from electrification occur because of improved energy efficiency compared to the direct use of fossil fuels for heating and vehicles, even when the LCP is close to 40%, i.e. that of the existing NYS electricity grid. At this LCP, average heating emissions per unit heat delivered are 70% lower with current electric technologies than existing fossil fuel-based heating; average vehicle emissions per mile traveled are 56% lower.

For LCPs at or below 60%, higher levelized costs of electricity (LCOE) of wind and solar generation are mitigated by higher utilization of existing infrastructure with increased HVE (with LCOE even decreasing at HVEs up to 20-40%). The 70-80% LCP range represents a transition phase: Beyond 80%, integration costs (e.g., curtailment and battery storage) lead to rapidly rising LCOEs. Accordingly, three primary levelized cost of electricity (LCOE) drivers are identified from the range of LCPs and HVEs investigated: (1) per-unit costs of existing infrastructure decrease with increasing demand from HVE, (2) wind and solar power generation costs rise relative to gas-

based and hydropower generation, and (3) costs of integration increase when large amounts of wind and solar power produce electricity in excess of demand.

For LCPs below 80%, wind generation meets most of the low-carbon generation requirement, as onshore wind represents the lowest LCOE renewable resource, followed by offshore wind resources near the dense load areas of New York City and Long Island. Beyond 80% LCP, paired solar generation and batteries become the most cost-effective method of displacing fossil fuel-based electricity generation. At higher LCPs, battery cycling occurs daily, making solar a more appropriate paired generation resource – at least some electricity is generated from solar daily whereas wind can drop off considerably for multi-day periods, particularly in the summer.

The marginal costs of lowering emissions from the limited set of electricity supply technologies considered here (wind, solar, battery and gas turbines) become high enough at LCPs larger than 80% to suggest that other nascent technologies (e.g., hydrogen storage) may play a role in achieving full energy sector decarbonization. Moreover, targeted deployment of other demand-side technologies not modeled – such as upgraded building envelopes, thermal storage and ground-source heat pumps – could further reduce supply costs by reducing heating-driven system peaks. Demand-side flexibility measures – like dual-fuel capabilities and grid-interactive controls – may also mitigate integration costs and reduce dispatchable capacity requirements. Lastly, breakthroughs in energy and emissions intensive industrial sectors could partially scale down emissions reductions needed in the residential, commercial, transportation, and electricity sectors.

A comparison of model results described in this chapter to initial analyses presented to the New York State Climate Action Council ("NYS study") validated SECTR-NY outputs, but also highlighted important factors in assessing the planning implications of such models. While SECTR-NY and the NYS study compute similar energy resource capacities for a scenario in line with the State's Year 2030 targets, deviations between the two can largely be attributed to differences in time series data for wind/solar potential time series and historical demand data, and to this chapter's particular attention to low-temperature effects on heat pump and electric vehicle performance. Accurately modeling the potential generation from renewable resources and new electrification-driven peak demands does thus affect the resource capacity required to meet the electric load. However, the two models do diverge significantly in the calculation of GHG emissions. SECTR-NY computes lower emission reductions than the NYS study for a given combination of LCP and HVE; SECTR-NY includes upstream natural gas leakage in line with recent research and its related quantifiable GHG effects, whereas the NYS study assumes a lower leakage rate. As detailed previously, this distinction has significant implications for the amount of electrification needed to meet the State's GHG reduction targets.

A couple of caveats surrounding this chapter's methodology and results are also worth mentioning. Foremost, SECTR does not model the electricity distribution network. As there will be a need to upgrade distribution to incorporate end-use heating and vehicle electrification, future work should investigate the scale, location, and costs of this reinforced capacity. Second, all SECTR generation is considered to be lumped. While this assumption substantially increases model tractability, it masks operating practices at the individual generator level where decisions are made. Third, LCPs are imposed on the amount of instate electricity generation, and do not

account for the carbon content of any imported electricity. Should state regulations change to allow clean, imported electricity to satisfy low-carbon generation targets, the SECTR general formulation will need to be adjusted. Lastly, this chapter presents results for a single set of cost assumptions. Should these assumptions prove inaccurate, rerunning the presented decarbonization scenarios will be required.

As the SECTR framework is an open-source, computationally efficient, capacity transition and system operation framework, the energy systems research community can adapt it in a number of ways for future work. One possibility is parameterizing SECTR for other RTO/ISO settings to explore comparative lowest cost decarbonization pathways. Moreover, within an RTO/ISO, researchers can investigate the impact of further interconnections to external generation. Lastly, researchers can build upon the SECTR framework by addressing the caveats mentioned above, such as by adding location specific costs for upgraded distribution capacity.

5. Conclusions

Chapter I introduces an open-source System Electrification and Capacity Transition (SECTR) modeling framework; the framework is then applied to the New York State (NYS) regional energy system (SECTR-NY). By characterizing existing system capacities, loads, and pricing structures, SECTR-NY reasonably approximates current electricity supply costs, establishing a reliable baseline from which to investigate different combinations of low-carbon electricity percentages (LCP) and heating and vehicle electrification rates (HVE).

Methodologically, SECTR addresses several shortcomings of traditional capacity expansion models (CEMs), including characterization of existing energy infrastructure systems, multi-year simulations with weather-dependent time series inputs, and spatially resolved end-

use electrification effects. In parameterizing the model for NYS, the model incorporates improved emissions accounting assumptions specified by recent climate legislation but previously unimplemented in state decarbonization studies. This study demonstrates that overall energy emissions reductions can be achieved at lower electricity costs by prioritizing heating and vehicle electrification ahead of complete grid decarbonization; the former approach still requires a major buildout of wind and solar power, but at lower percentage penetration into the grid because of higher demands from more electrification. Moreover, three main electricity supply cost drivers are established for a decarbonizing energy system: (1) decreasing per-unit supply costs of existing infrastructure with increasing electrification (i.e., with higher demand); (2) higher wind and solar power supply costs relative to current hydropower and fossil fuel-based generation; and (3) increasing costs of integration (due to curtailment and energy storage) as solar and wind supply in excess of demand increase with LCP.

6. Data Availability

All code and data used for the SECTR-NY model formulation can be found in the following GitHub repository: https://github.com/SEL-Columbia/sectr-ny.

7. Appendix A: Full list of cost and existing capacity assumptions

Table I-16 contains a full listing of all nodal cost assumptions in SECTR-NY. The Methodology provides a full accounting of how these assumptions were reached. Internodal transmission upgrade and O&M costs are presented in Table I-2.

Table I-16: Cost assumptions used in SECTR-NY.

Quantity	Unit	Node 1 ($)	Node 2 ($)	Node 3 ($)	Node 4 ($)
Onshore Wind Capacity Cost, High	$/kW	1992	1992	N/A	N/A
Onshore Wind Capacity Cost, Low	$/kW	1698	1698	N/A	N/A
Offshore Wind Capacity Cost, High	$/kW	N/A	N/A	3583	3583
Offshore Wind Capacity Cost, Low	$/kW	N/A	N/A	2256	2256
Utility-Scale Solar Capacity Cost, High	$/kW	1341	1341	1593	1593
Utility-Scale Solar Capacity Cost, Low	$/kW	1006	1006	1195	1195
Battery Storage Energy Cost, High	$/kWh	208	208	208	208
Battery Storage Energy Cost, High	$/kWh	144	144	144	144
Hydrogen Storage Energy Cost	$/kWh	0.35	8.29	8.29	8.29
Hydrogen Storage Power Cost	$/kW	3013	3013	4036	4036
New Fossil Fuel-Based Generation Capacity Cost	$/kW	772	772	1034	1034
Hydropower Generation Cost	$/MWh	18.47	28.02	N/A	N/A
Nuclear Generation Cost	$/MWh	37.94	N/A	26.82	N/A
Biofuel Generation Cost	$/MWh	20.66	27.41	27.05	32.39
Imported Electricity Cost	$/MWh	22.13	N/A	70	N/A
Wholesale Natural Gas Price	$/MMBTU	2.89	4.04	3.67	3.62
Existing Fossil Fuel-Based Generation Ramping Cost	$/MW-h	79	79	79	79
New Fossil Fuel-Based Generation Ramping Cost	$/MW-h	69	69	69	69
New Fossil Fuel-Based Generation Fixed O&M Cost	$/kW-yr	6.97	6.97	6.97	6.97
Onshore Wind Capacity Fixed O&M Cost	$/kW-yr	18.1	18.1	N/A	N/A
Offshore Wind Capacity Fixed O&M Cost	$/kW-yr	N/A	N/A	38	38
Utility-Scale Solar Capacity Fixed O&M Cost	$/kW-yr	10.4	10.4	10.4	10.4
Hydrogen Storage Fixed O&M Cost	$/kW-yr	48.87	48.87	48.87	48.87
New Fossil Fuel Based Generation Variable O&M Cost	$/MWh	4.48	4.48	4.48	4.48
Existing Generation Capacity Maintenance Cost	$/kW-yr	27.64	53.44	101.303	104.6
Existing Transmission Capacity Maintenance Cost	$/MWh	16.9	16.9	27.3	27.3

Table I-17 contains a full listing of existing nodal capacities modeled in SECTR-NY. The Methodology provides a full accounting of how these values were reached. Internodal existing transmission capacities are presented in Table I-2.

Table I-17: Existing capacities modeled in SECTR-NY.

Capacity Type	Node 1 (MW)	Node 2 (MW)	Node 3 (MW)	Node 4 (MW)
Onshore Wind	1985	0	0	0
Offshore Wind	0	0	0	0
Utility Scale Solar	0	0	0	56.5
Behind-the-Meter Solar	562	523	293	259
Gas-Fueled	3934.2	8622.5	10249.9	4192.7
Hydropower	4717.4	608.7	0	0
Nuclear	3536.8	0	2311	0
Biofuel	258	45	59.7	142.2
Interregional Import Limits	1500	0	1250	0
Battery Storage, Energy	5.2	80	0	65
Battery Storage, Power	3	20	0	10

Chapter II: Assessing new transmission and energy storage in achieving increasing renewable generation targets in a regional grid

Abstract

This study evaluates generation, transmission, and storage capacity needs to achieve deep renewable energy penetration in a regional electricity grid with an average load of approximately 20 GW. Increasing renewable energy targets are analyzed to evaluate the effect of realistic regional transmission upgrade and energy storage cost assumptions on the cost-optimal mix of generation, transmission, and storage capacity. Contextual data is used for New York State's grid to examine how electricity generation from renewable energy resources (wind, water, and solar power) can meet between 50% and 80% of electricity demand. A central finding of the study is that when realistic transmission upgrade costs are assumed, new interzonal transmission and battery storage are not needed to cost-effectively meet near-term renewable energy goals. In fact, New York can achieve 50% renewable energy penetration with only a buildout of new generation capacity: Onshore wind (13.7 GW), offshore wind (4.1 GW), and solar photovoltaics (3 GW). The presence of grid-scale battery storage, electric vehicles, or additional behind-the-meter solar capacity does not markedly change the model-selected generation mix. To achieve the 50% target, we compute a $52/MWh levelized cost of electricity for new renewable energy, which is in line with current generation costs.

As the renewable generation target increases beyond 50%, the model begins to select transmission upgrades and new storage capacity, the latter particularly if battery costs continue to decline as anticipated. At deeper targets, marginal generation capacity would otherwise experience high curtailment primarily due to supply-demand imbalances; we calculate the value

of energy storage at a 65% renewable energy penetration level to be 2.5-3 times higher than its value at a 50% level. However, the additional storage and generation – and transmission, to a lesser degree – needed to achieve longer-term renewable energy goals lead to a substantial rise in total investment. Between 50% and 55% targets, the computed marginal levelized cost of electricity for new variable renewable energy is $94/MWh, compared to $592/MWh between 75% and 80%, suggesting alternative integration measures are likely necessary at such high penetration rates.

1. Introduction

The use of variable renewable energy (VRE) technologies to decrease fossil fuel usage and greenhouse gas (GHG) emissions is widely accepted, e.g. [119,120]. However, the stochastic and intermittent nature of VRE supply is expected to require some suite of system integration measures at large installed capacities [121]. Such measures can include advanced grid monitoring, communication, and control [122]; expanded transmission capacity [50]; electrification of transportation and heating [123]; increased energy storage capacity [124]; and further interconnection among regional systems [125]. Two integration measures that could be achieved at large scale are the primary focus of the following chapter: expanded transmission and grid-scale battery storage.

Researchers looking at transmission dynamics have shown that increased transmission is more effective than battery storage at lowering wind power curtailment [126]; curtailment is almost entirely due to transmission constraints in some studies [127]. However, to ease computational requirements or to standardize across large geographic regions, many energy system models do not account for the full set of constraints that face new transmission projects,

instead assuming (a) costs below historical rates [50,128], (b) idealized network topologies [129], or (c) unlimited interregional transmission capacity [130,131].

Previous analyses have modeled the ability of battery storage to improve VRE integration [37]. Storage is shown to be a valuable balancing asset at high VRE penetration levels, but its deployment is often not a cost-effective method of reducing curtailment; system benefits diminish with increased adoption [132] and integration can become largely a seasonal issue with large VRE capacities, particularly in the case of wind power [133]. To date, energy storage has largely been used to provide energy system services other than VRE supply shifting (e.g. regulation services and peak load reduction) [134]. Evaluating storage adoption and transmission expansion together, researchers have found that in a transmission-constrained system, energy storage at generation sites allows for greater renewable power utilization compared to storage at load centers [135]. Yet while co-locating storage with transmission bottlenecks has shown to be an effective method of integrating VRE, such practice can also reduce the economic viability of the batteries [136]. A previous study further identified that large-capacity VRE supply variability is likely to be highest distant from the VRE resource; this finding implies distributed energy storage will have value for reliability services that may not be captured in capacity expansion models.

In this study, we evaluate the cost-effectiveness of these two integration measures (energy storage and transmission) to achieve renewable generation targets (RGTs) in New York State's (NYS) regional grid; the NYS grid aligns with the New York Independent System Operator (NYISO) control area. We perform simulations with and without electric vehicle (EV) adoption, as the presence of a sizable electric transportation load can influence how a system decarbonizes

[137]. While this analysis uses the NYS system as a case study, many states across the US have adopted RGTs [138], and all include some common characteristics: Spatially heterogeneous electricity demands, transmission line limits, potential for battery storage, and existing fossil fuel-based transportation that may shift to some proportion of electric vehicles during a larger energy transition. Therefore, the approach described here can be applied to any regional electric grid after adjusting for domain-specific topologies.

Previous work by two of the authors showed that up to 10 GW of onshore wind can be added to the NYS grid with minimal curtailment; beyond this point, curtailment is largely a seasonal issue with higher wind output and lower demand in the winter [139]. While this prior work identified transmission bottlenecks, it did not evaluate whether upgrades would be economical. A NYISO study evaluated the ability for NYS to integrate 8 GW of wind, finding that this capacity would have no adverse reliability impacts, would decrease total system costs, would result in less than 2% curtailment, and would create congestion only at local transmission facilities [140]; it remains to be seen whether these results hold at higher renewable penetrations. Similarly, large NREL studies investigated the integration of high levels of VRE in NYS as a portion of the larger Eastern Interconnection [37,131]. Yet these studies ignored intra-NYS transmission and did not quantify the cost-effectiveness of various flexibility measures. [129] considered the possibility of 100% decarbonization of all NYS energy infrastructure; however, in modeling the state's electricity grid, this study both ignored intrastate transmission limits and underestimated the costs of state-specific transmission expansion. The current chapter addresses the above-mentioned gaps in previous NYS-specific work, and in doing so, presents a modeling framework translatable to other grid topologies.

Chapter II makes two principal contributions to the literature on VRE integration. First, it examines the cost-effectiveness of transmission expansion in a regional grid with price assumptions based on historical projects. Related renewable penetration studies have understated the cost of transmission expansion, especially near high-density load zones [50]. Second, our work appraises the need for particular types of infrastructure along a pathway of near- to long-term RGTs. By evaluating cost-effective methods of meeting more immediate goals, and then by comparing these results to those for more distant renewable energy objectives, we investigate optimal energy planning decisions at various stages in the transition of an electricity grid. Other contributions include substantiating the value of offshore wind generation in a transmission-constrained system with coastal load pockets and quantifying the value of battery storage at increasing renewable targets.

The structure of Chapter II is as follows: *System Topology* (Section 2) discusses NYS RGTs and relevant characteristics of the NYS electricity grid. *Methodology* (Section 3) details the data sources utilized and the development of the Renewable Target Model, the study's primary analytical tool. *Results* (Section 4) presents relevant figures and findings. *Discussion* (Section 5) unpacks what the results mean for meeting RGTs and how policy can best support near- and long-term goals in NYS and other regional systems. *Conclusion* (Section 6) summarizes the most salient aspects of the study.

2. System Topology

NYISO manages New York State's electricity grid. The NYISO control area shares boundaries with NYS and is divided into 11 load zones. For the purposes of the present study, we

group these zones into the four regions shown in Figure I-1 based on major transmission interfaces.

Region 1 (NYISO Zones A-E) produces 91% of the state's hydropower electricity and 64% of its nuclear-generated electricity, while accounting for only 34% of statewide demand [38]; this region also contains 86% of the state's potential onshore wind power capacity [62]. Region 2 (NYISO Zones F and G) holds the remaining 9% of hydropower supply and 13% of the state's electricity demand [38]; this region contains 10% of the state's potential onshore wind capacity [13]. In contrast, the two remaining downstate regions (NYISO Zones H-J, designated Region 3, and Zone K, designated Region 4) account for 51% of NYS electricity demand [38] but offer little potential onshore renewable capacity [62]; however, there is abundant undeveloped offshore wind power potential adjacent to these areas, and NYS (and nearby states) has begun incentivizing its deployment [141]. Region 3, which includes the New York City metropolitan area plus Westchester County, does include one nuclear power plant; however, this plant is slated to be decommissioned in 2021, so we do not include it in this study.

Bulk transmission of electricity in NYS primarily follows a west-to-east pathway from Buffalo to Albany, and then changes to a north-to-south orientation, connecting Region 2 to Region 3. Approximately 1.4 GW of transmission capacity exists at the interface of Regions 3 and 4 [48]; however, because of high electricity demand in Region 3, these transmission lines are infrequently loaded. In 2016, NYS obtained 19% of its electricity from hydropower (86% from instate generation, 14% from imports from Quebec and Ontario), 31% from nuclear power, and 4% from wind and solar photovoltaic (PV), resulting in one of the lowest carbon-intensity fuel mixes in the country [38,142]. The fossil fuel-based generation fleet in NYS is composed of

primarily natural gas-fired plants that are well dispersed throughout the state, generally near the loads as required by NYISO [143]. Because of this distribution, in-region dispatchable generation generally satisfies loads not met by low-carbon resources (VRE, hydropower and nuclear power) [144].

Table II-1 summarizes relevant regional data: Average demand, average generation from fixed capacity sources (i.e., nuclear, solar PV and hydropower), maximum potential wind capacity (see Section 3.3), and average EV load assuming 25% electric vehicle adoption (see Section 3.5). The processes for determining these quantities are detailed in *Methodology*.

Table II-1: Summary of relevant NYISO regional data and model parameters.

	Average Regional Electric Demand and Nuclear-Solar-Hydro Generation (MW)					Maximum Wind Capacity (MW)		Average EV Load (MW)
Region	Demand[1]	Nuclear[2]	Solar PV[3]	Hydro[4]		Onshore Wind[5]	Offshore Wind[6]	25% Adoption[7]
				Fixed	Flexible			
1	6,382	3026	120	2395	328	32,406	0	519
2	2,495	0	167	0	270	4,376	0	249
3	7,211	0	74	0	0	0	37,572	341
4	2,566	0	85	0	0	0		292

[1]Average 2007-2012 electricity demand [24].
[2]Average 2016 nuclear generation, excluding soon-to-be decommissioned facility [24].
[3]3 GW solar capacity distribution, per [80].
[4]See Section 3.2.
[5]Maximum potential onshore wind capacity, per [62].
[6]Maximum potential offshore wind capacity for installations at water depths <60 meters [65]; see Section 3.3.
[7]Average EV load given 25% electric vehicle adoption as described in Section 3.5.

To evaluate the effect of transmission prices on future energy scenarios, we reviewed the costs of recent and proposed transmission projects in NYS, as well as the cost assumptions used in other studies. The reported costs for recent NYS projects align with the high-cost estimates used in this study [51–53,145]. Low-cost estimates come from a combination of previous

integration model assumptions and comparisons to the high-cost prices. [50] estimates a cost of $1400/MW-mi for a line length of 300 miles. [130] assumes a cost of $4173/MW-mi for 345 kV cables and supporting substation and transformer improvements. None of these references account for the unusually high costs of large-scale transmission in densely populate areas (e.g., Regions 3 and 4 in the present study). Table II-2 summarizes the results of this process and shows transmission upgrade costs in $/MW-mi for comparison.

In 2010, the NYS Public Service Commission established a 30% renewable energy (wind, water, and solar power; abbreviated as WWS) target to be met by 2015 [146]. While the state has not yet achieved the 30% RGT (28% of generation in 2017 [147]), [64] demonstrated that a build out of onshore wind power without additional integration measures would be sufficient to reach this level of renewable energy penetration. In 2016, NYS accelerated its clean energy goals, adopting a 50% RGT for 2030 [148].

Table II-2: Summary of transmission interfaces.

Regional Interface	Distance[1] (miles)	Current Limits[2]		Low Cost Estimates $/MW-mi	High Cost Estimates $/MW-mi
		W->E (MW)	E->W (MW)		
1: Region 1 -> 2 (Buffalo to Albany)	300	4,925	3,400	1,600[3]	3,200[4]
2: Region 2 -> 3 (Albany to NYC)	150	5,750	2,000	3,200[5]	6,400[6]
3: Region 3 -> 4 (NYC to Long Island)	60	1,424	120	8,000[7]	16,000[8]

[1] Interregional distance ($d_{rr'}$) calculated via Google Maps and rounded to the nearest 10 miles.
[2] Current line limits ($L_{rr'}$) ascertained from [149].
[3] From [50], for line length of 300 miles, approximately $1,400/MW-mi (aboveground HVDC). After conversion, [130] assumes a cost of $2056/MW-mi for a 300-mile line.
[4] From [51], after subtracting upgraded substation costs, approximately $3,614/MW-mi (underground HVDC).
[5] From [130], approximately $4173/MW-mi for overhead 345 kV and supporting transformer and substation reinforcements.

[6] From [52], approximately $6567/MW-mi (underground HVDC). Because of the unique challenges facing the referenced project, we assume ½ the cost of [52] for our low-cost scenario, a quantity more closely in line with the cost estimate from [130] after subtracting the latter estimate's transformer and substation reinforcement costs.
[7] Because of the unique challenges facing transmission projects in densely populated areas and the lack of low cost estimate for such projects, we assume ½ the cost for the Region 3-4 transmission high cost scenario
[8] From [53,145], approximately $13,986/MW-mi (underground HVDC).

3. Methodology

This section describes the formulation and assumptions of the Renewable Target Model (RTM), a mixed integer linear program (MILP) that minimizes the total system capital investment necessary to meet an RGT. The model is optimized over a 6-year time period (2007-2012) selected based on data availability; unless otherwise noted, all data described here applies to this time period. The RTM is formulated in Python and solved in Gurobi on a 16C GeForce GTX Titan Black GPU with 32 GB of RAM.

3.1 Objective function and net load constraints

The objective function for the RTM (Eq. II-1) minimizes total system investment based on the per-unit capacity cost, *CAP*, and installed capacity, *X*, of onshore wind generation ("on"), offshore wind generation ("off"), grid-scale battery storage ("batt"), and new interregional transmission capacity ("trans") between each region *i* and adjacent region *ii'*. The distance between each region *i* and *i'* is defined by $d_{ii'}$. All costs, *CAP*, are annualized over with their respective capital annualization rates and then summed over the 6-year optimization period.

$$minimize \sum_{i \in I} \left[CAP_{on,i} * X_{on,i} + CAP_{off,i} * X_{off,i} + CAP_{batt-e} * X_{batt-e,i} + \sum_{r'} CAP_{tx,ii'} * d_{ii'} * X_{tx,ii'} \right] \quad \text{(II-1)}$$

Model constraints for regional net load, onshore and offshore wind utilization, and interregional transmission are shown in equations (II-2) – (II-10). Unless otherwise stated, all constraints apply for all time steps, *t,* in *T* and for all regions, *i*, in I.

The metric of "net load" is used throughout the analysis; the net load is the remaining electricity demand after using low-carbon resources (i.e., VRE, hydropower and nuclear power). Eq. (2), below, defines net load, NL_r^t, for each region, *i*, at each time step, *t*. Eq. (2) includes five exogenously-defined variables: Existing electricity demand, $\mathrm{D}_{elec,i}^t$; nuclear generation, N_i^t; solar generation, computed as the product of behind-the-meter (BTM) solar capacity, $X_{btm-solar,i}$, and generation potential, $W_{btm-solar,i}^t$; fixed hydropower generation, $H_{fix,i}^t$; and fixed electric vehicle charging, $D_{veh-fix,i}^t$. The RTM also uses the following decision variables: Flexible hydropower generation, $H_{flex,i}^t$; onshore wind generation, computed as the product of onshore wind capacity, $X_{on,s,i}$, and generation potential, $W_{on,s,i}^t$ at all sites *s*; offshore wind generation, computed as the product of onshore wind capacity, $X_{off,i}$, and generation potential, $W_{off,i}^t$; increase in battery state of charge (i.e. battery charge), $\gamma_{batt,i}^t$; decrease in battery state of charge (i.e. battery discharge), $\delta_{batt,i}^t$; flexible electric vehicle charging, $D_{veh-flex,i}^t$; renewable generation curtailment, ζ_i^t;electricity exported to an adjacent region, Z_{ii}^t, and electricity imported from an adjacent region, $Z_{i'i}^t$. The model assumes transmission losses, l, of 3% between adjacent regions and applies them to the imports; this is not meant to provide a definitive transmission loss model, but it ensures that electricity is first used to meet demand nearest the region in which it is generated.

$$NL_i^t = \mathrm{D}_{elec,i}^t + D_{veh-fix,i}^t + D_{veh-flex,i}^t - N_i^t - X_{btm-solar,i} * W_{btm-solar,i}^t - H_{fixed,i}^t - H_{flex,i}^t - \sum_{s \in S_i} x_{on,s,i} * W_{on,s,i}^t - X_{off,i} * W_{off,i}^t + \gamma_{batt-e,i}^t - \delta_{batt-e,i}^t + \zeta_i^t + \sum_{r'} [Z_{rr'}^t - (1-l) * Z_{r'r}^t]$$

(II-2)

Eqs. (II-3) – (II-5) impose electricity utilization limits and site capacity constraints for the onshore wind generation. Onshore wind power capacity, $x_{on,s,i}$, is installed at individual sites, s, selected from all sites within a region, S_i.; the aggregate regional onshore wind power capacity is defined as $X_{on,i}$. Each onshore wind site is defined by a potential wind-generated electricity output at each time step, $W_{on,j}^t$.

$$x_{on,s} \leq x_{on,s}^{max}$$

(II-4)

$$\sum_{s \in S_i} x_{on,s} = X_{on,i}$$

(II-5)

Eqs. (II-6) – (II-7) constrains offshore wind capacity. The total offshore wind power capacity across all regions is limited by the maximum potential offshore wind capacity for the state[xviii], X_{off}^{max}.

$$\sum_{i \in I} X_{off,i} \leq X_{off}^{max}$$

(II-7)

[xviii] For all potential installations at water depths <60 meters [65]. Note that offshore wind power is only available for Regions 3 and 4; the total capacity is used here as it is not yet clear where offshore transmission lines will make landfall.

Eq. (II-8) restricts the amount of electricity transmitted between regions, $Z_{ii'}^{t}$, to the sum of existing transmission limits, $X_{tx,ii'}^{existing}$, and new transmission capacity, $X_{tx,ii'}$:

$$Z_{ii'}^{t} \leq X_{tx,ii'}^{existing} + X_{tx,ii'} \tag{II-8}$$

Eq. (II-9) is the domain constraint:

$$All\ variables \geq 0 \tag{II-9}$$

The RTM characterizes renewable generation as that from wind, water, and solar (WWS), consistent with NYS policy targets as well as generally accepted definitions of renewable energy. For a given renewable generation share of statewide electricity (*RGT)*, the model requires that this fraction of the demand be met by WWS (accounting for transmission losses):

$$\sum_{t\in T}\sum_{i\in I}\left[S_i^t + H_{fixed,i}^t + H_{flex,i}^t + \sum_{s\in S_i} x_{on,s,i} * W_{on,s,i}^t + X_{off,i} * W_{off,i}^t - \zeta_i^t - \sum_{i'} l * Z_{i'i}^t\right] \geq RGT * \sum_{t\in T}\sum_{i\in I}\left[\mathrm{D}_{elec,i}^t + D_{veh,i}^t\right] \tag{II-10}$$

The model does not include capital and operational costs for fossil fuel-based electricity generation. While these costs may constitute a substantial portion of overall system expenditures, we do not expect that they would significantly affect the renewable energy generation mix for a particular RGT. There is a body of work, including by the present study's authors, that investigates reliability services in addition to energy services; for example, our earlier analysis of large capacities of wind power in NYS indicated that operating reserve and regulation requirements could increase in magnitude and become more concentrated near load

centers where large dispatchable thermal generation capacity already exists [139]. By tailoring the focus of the RTM, we explore the primary costs and infrastructure planning challenges of renewable energy only, filtering out downstream concerns about how other grid actors will respond to system change. We therefore do not consider how the internal economics of individual market participants will influence future bids or retirements; we do not expect this to be significant for transmission or storage expansion in the context of deep VRE penetration.

3.2 Hydropower

We do not explicitly model hydropower reservoirs, as NYS has over 300 individual hydropower stations. Instead, we rely on a method we investigated in detail in a previous study [150]. Actual monthly hydropower output by facility [151] is aggregated at the regional level to produce total regional monthly hydropower output, $H_{tot,i}^{monthly}$. As shown in Eq. II-2, the RTM includes both fixed hydropower (defined exogenously, it varies in time and by region) and flexible hydropower (limited regional daily energy, but with intraday flexible output limited by regional maxima). This basic formulation is adaptable to many approaches to hydropower; here, we assume that in each region some fraction of the total regional hydropower output, is fixed, $H_{fix,i}^{monthly}$. In New York, the two largest hydroelectric plants (both in Region 1) operate near their maximum possible outputs given available stream flows. As such, we treat these facilities as fixed hydro. The resulting proportion of hydropower considered fixed is 88% of $H_{tot,1}^{monthly}$ (Region 1) and 0% of $H_{tot,2}^{monthly}$ (Region 2); Regions 3 and 4 contain no hydropower at all. We fit a cubic spline function to $H_{fix,i}^{monthly}$ to determine the regional hourly fixed hydropower output, $H_{fix,i}^{t}$. The cubic spline ensures continuity and smoothness between time steps.

As NYS hydropower exhibits a degree of load-following behavior that is generally diurnal with some additional storage capabilities, we designate the remaining portion of hydropower output as the total regional monthly flexible hydropower generation, $H_{flex,i}^{monthly}$. The resulting proportion of hydropower considered flexible is 12% of $H_{tot,1}^{monthly}$ (Region 1) and 100% of $H_{tot,2}^{monthly}$ (Region 2). We fit a cubic spline function to $H_{flex,i}^{monthly}$ to determine regional daily flexible hydropower, $H_{flex,i}^{daily}$. As above, the cubic spline ensures continuity and smoothness across the time series.

For the purposes of the present study – and as recently investigated by the authors in more detail in [150] – the hourly flexible hydropower output by region, $H_{flex,r}^{t}$, is subject to the constraint that it must meet a daily total regional flexible hydropower output, $H_{flex,r}^{daily}$, without exceeding aggregate regional flexible hydropower capacity, $H_{flex,r}^{max}$. Eqs. (II-11) – (II-12) impose these limitations.

$$\sum_{t=1+24m}^{24*(m+1)} H_{flex,i}^{t} = H_{flex,i}^{daily}\,, \qquad m = 0..\frac{T}{24} - 1 \tag{II-11}$$

$$H_{flex,i}^{t} \leq H_{flex,i}^{max} \tag{II-12}$$

The resulting regional fixed and flexible hydropower generation averages are shown in Table II-1 (Section 2).

3.3 Variable renewable energy potentials

In simulating capacities of onshore wind power far exceeding current levels, the authors rely on model wind power data for 126,000 potential U.S. wind sites developed by National Renewable Energy Laboratory (NREL) [62,63]. In a previous study, two of the authors found the NREL model consistently to over-predict the electricity generation at existing onshore wind power sites in NYS and developed a procedure to adjust the time series to reflect actual output [64]; the resulting hourly potential onshore wind power output by site (normalized by installed capacity), $W_{on,s}^{t}$, is used in the current study.

The same NREL database provides hourly model offshore wind power outputs for two sites off near Regions 3 and 4. Because no method of independently verifying the offshore wind generation estimates exists and because the model capacity factors at those locations more closely reflect the performance of global offshore wind installations, we make no additional modifications to these time series for the hourly potential offshore wind power output by region, $W_{off,i}^{t}$. We assume that generation from the westernmost site makes landfall in Region 3 and that generation from the easternmost site makes landfall in Region 4. As existing literature has only placed an upper bound on the total potential offshore wind capacity in NYS waters, the RTM limits the combined offshore capacity of both regions to 37.6 GW for water depths less than 60 meters, per [65].

To determine the potential solar resource in each region, we first select a representative city for each NYISO zone from those in the NREL National Solar Radiation Database [82]. We then compute hourly potential solar PV output, normalized by capacity, using the NREL System Advisor Model (SAM) [75]. SAM simulates the performance of commercially available equipment and

realistic system configurations; we select PV panels with 22% efficiency, installed with fixed tilt equal to location latitude and with an inverter with 95.3% weighted efficiency (as determined by the protocols of the California Energy Commission [76]). Because of its cost and potential production relative to wind power in NYS, model simulations do not select any solar PV capacity in cost-optimal infrastructure mixes.

To account for the reality that solar installations will almost certainly continue to grow in NYS, we impose behind-the-meter solar PV capacities of either 3 GW or 6 GW in all model optimizations. These capacities are distributed to NYISO zones in proportion to projected zonal capacity distributions from a recent NYISO study [80]. The products of the non-dimensional hourly zonal potential solar-generated electricity and the zonal solar capacity are then aggregated at the regional level to produce the regional hourly solar-generated electricity output, S_r^t. Table II-1 (Section 2) presents the average regional generation for 3 GW of installed capacity.

3.4 Battery storage treatment

Standalone battery storage in the RTM is assumed to have 95% efficiency in charging and discharging. The battery power-to-energy ratio is (1 kW):(4.2 kWh), a specification equivalent to that of an available commercial-scale battery storage product [85]. As this specification is based on usable storage capacity, we place no limitations on the battery depth of discharge; we also do not constrain cycling behavior aside from the charge/discharge limits imposed by the power-to-energy ratio and formalized in Eq. (II-13). The RTM treats regional storage as a lumped capacity, and at all times limits the energy stored in a standalone battery, $E_{batt,i}^t$, to less than its capacity,

$X_{batt-e,i}$. After accounting for battery efficiencies, the hourly change in battery energy level is equal to the battery charge or discharge. Eqs. (II-14) and (II-15) govern these relationships.

$$\gamma_{batt,i}^{t} + \delta_{batt,i}^{t} \leq \frac{1}{4.2} * X_{batt,i} \tag{II-13}$$

$$E_{batt,i}^{t} \leq X_{batt-e,i} \tag{II-14}$$

$$\frac{\delta_{batt,i}^{t}}{\eta_{batt}} - \eta_{batt} * \gamma_{batt,i}^{t} = E_{batt,i}^{t-1} - E_{batt,i}^{t} \tag{II-15}$$

In addition to including battery storage capacity as a decision variable in the cost minimization, we also impose different amounts of battery storage capacity in certain model scenarios to investigate its distribution and value in reducing VRE capacity needs.

3.5 Electric vehicle charging constraints

In the RTM, EV load is specified as a percent of statewide automobile use, as determined by reported gasoline consumption quantities. We use 2015 NYS annual gasoline sales by county, aggregated by region, to determine the annual quantity of energy used for automobile transport [152]. After accounting for standard electric and gasoline-engine vehicle efficiencies (24.7 MPG [153], 0.36 kWh/mile [154]) and a charging efficiency (η_{veh} = 95%), we convert the annual quantity of gasoline sold to average hourly regional electric loads, $D_{veh,i}$. $D_{veh,i}$ for the 25% EV adoption scenario is shown in Table II-1 (Section 2).

Many studies have investigated the system benefits of flexible EV charging in a variety of domains, each with particular constraints. Here, we look at one fairly straightforward charging regime, in which we investigate both fixed and flexible charging. The purpose of the current study

is not to determine an optimal EV charging strategy, but to evaluate how charging flexibility impacts VRE capacity selection and the value of other large-scale integration measures.

In fixed charging scenarios, the RTM distributes daily EV load, $D_{veh.i}^{daily}$ ($D_{veh.i}^{daily} = 24 * D_{veh,i}$), evenly over the hours between 7pm and 6am (inclusive), leaving twice the hourly EV load at each. In flexible charging scenarios, the model meets the daily EV load over these same hours. By assuming that EV charging operates on a daily cycle, and that the model needs to meet charging requirements by 7am, this methodology aligns with others found in the literature, e.g. [155]. The RTM imposes fixed and flexible charging conditions through Eqs. (16) and (17)-(18), respectively. Maximum charging capacity is limited to one quarter the daily EV load.

$$D_{veh-fix,i}^{t} = \frac{D_{veh,i}^{daily}}{12}, \qquad for\ t = 19 + 24m \ \ to \ \ 6 + 24(m+1), \qquad where\ m = 0..\frac{T}{24} - 1 \tag{II-16}$$

$$\sum_{t=19+24m}^{6+24(m+1)} D_{veh-flex,i}^{t} = D_{veh,i}^{daily}, \qquad for\ m = 0..\frac{T}{24} - 1 \tag{II-17}$$

$$D_{veh-flex,i}^{t} \leq \frac{D_{veh,i}^{daily}}{4} \tag{II-18}$$

3.6 Capital cost assumptions

While renewable energy infrastructure cost projections abound, to limit computational requirements and the model solution set, we restrict the analysis to two capital cost scenarios. The high-cost scenario simulates prices similar to those currently available [66,156–158]. The

low-cost scenario assumes the presence of an accelerated learning curve for offshore wind power and standalone battery storage. Cost assumptions for both scenarios are shown in Table II-3.

In both cost scenarios, the price of onshore wind power remains constant; further cost declines in onshore wind power are likely to be less significant than those for offshore wind power and batteries [53, 54]. We do not expect this assumption to alter our conclusions as these cost assumptions largely only influence tradeoffs between onshore and offshore wind (this may be analogous to onshore wind vs. solar tradeoffs in other regions of the United States). After first exploring optimization sensitivity to transmission costs (Section 4.1.1), we take the price of new transmission capacity to be the average of the low-cost and high-cost estimates presented in Table II-2 with an annualization period of 20 years.

Table II-3: Wind and Storage RTM Capital Costs.

	Unit Cost ($)		Unit	Capital Annualization Period (Years)
	High Cost Scen.	Low Cost Scen.		
Onshore Wind[1]	1,588	1,588	kW^{-1}	20
Offshore Wind[2]	4,644	3,754	kW^{-1}	20
Battery Storage[3]	250	100	kWh^{-1}	10

[1] Onshore wind costs assembled from [66,156].
[2] Offshore wind costs assembled from [66,156].
[3] Battery costs assembled from [157,158].

The RTM annualizes capital costs for technology x based on a capital annualization rate, $A_{P_x,j}$. $A_{P_x,j}$ is computed as follows, with a 5% interest rate, j, over a technology-specific annualization period, P_x:

$$A_{P_x,j} = \frac{j * (1+j)^{P_x}}{((1+j)^{P_x} - 1)} \quad \text{(II-19)}$$

3.7 Treatment of system costs, savings, and marginal LCOEs

Total system costs are calculated over a 20-year period. For the types of infrastructure with a 20-year annualization period (wind generation and transmission), we multiply the costs given in Table II-3 by the capacities returned by the model. For battery storage, with an assumed 10-year lifetime, we include twice the quantity of batteries selected by the optimization in order to calculate the storage costs over a 20-year horizon. To determine the marginal levelized cost of electricity (LCOE) for a given RGT, we take the marginal total system cost compared to the previous RGT and annualize this difference per Eq. (II-19). This annual value is then divided by the difference between the average annual renewable electricity generated at the given RGT and at the previous RGT to compute the marginal LCOE.

In certain simulations, we impose a no-cost, system-wide battery storage capacity and compare the computed total system cost to that computed for the same RGT without battery storage. The difference between the two quantities, the "avoided capital cost," is considered the total battery storage value. However, the battery storage lifetime is shorter than that of other infrastructure investments considered. To determine the value of a unit of energy storage ($/kWh) for a given RGT, the avoided capital cost is standardized with the following equation:

$$Storage\ Value = \frac{\{Avoided\ Capital\ Cost\}}{\{Storage\ Capacity\} * \left(\frac{20\ years}{\{Storage\ Lifetime\}}\right)} \quad \text{(II-20)}$$

In addition to the 10-year battery lifetime described above, we also considered a battery lifetime based on 3000 cycles equivalent usage. We present system savings in both scenarios as future battery lifetimes are not yet known.

4. Results

This section begins with a detailed investigation of the computed mix of energy infrastructure to meet the 50% RGT (Section 4.1). We then extend the analysis to evaluate continuing trends up to an 80% RGT (Section 4.2), the upper limit of electricity sector decarbonization that researchers believe achievable in the US without notable increases in LCOE [50].

4.1 The 50% renewable generation target

Regional resource potentials and interregional transmission limits heavily influence the cost-optimal mix of energy infrastructure needed to meet the 50% RGT. As such, it is helpful first to understand regional generation and interregional flow dynamics under two divergent scenarios: (a) one with existing transmission limits, and (b) one with nearly unlimited transmission. (Nearly unlimited transmission is simulated by solving the RTM with an assumed upgrade cost of 1/20th the low-cost transmission estimates.) Figure II-1 summarizes the average NYS electricity load, utilized low-carbon generation (WWS plus nuclear), and interregional electricity flow for these two cases.

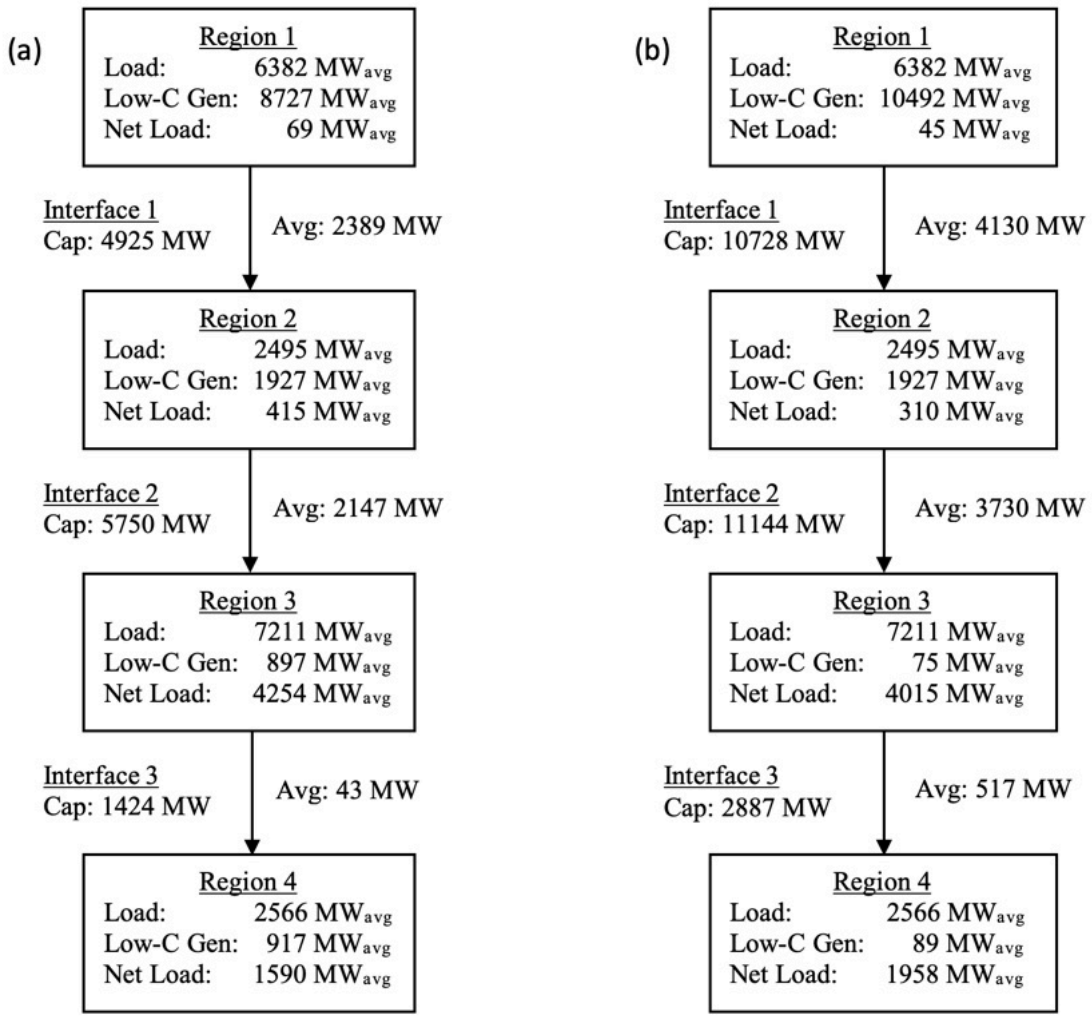

Figure II-1: Select simulated system characteristics under a 50% renewable generation target: (a) existing transmission limits; (b) nearly unlimited transmission. Low-carbon generation averages include 3 GW of solar capacity. All values in regional boxes are averages.

Figure II-1 shows that, regardless of transmission assumptions, the bulk of NYS low-carbon energy generation occurs in Region 1. As presented in Table II-1, Region 1 contains the entirety of the state's simulated nuclear generation, 91% of the state's hydropower generation, and 81% of the state's potential onshore wind power capacity. The model-selected onshore wind capacity in Region 2 is that region's total potential capacity and remains constant in both transmission scenarios. Two factors drive this result: (1) Potential wind power in Region 2 has a

higher capacity factor than in Region 1, and (2) Region 2 is nearer to load centers in NYC and Long Island, so new generation does not congest the upstream Region 1-2 transmission interface.

The principal difference between scenarios is the tradeoff between Region 1 onshore wind and offshore wind in Regions 3 and 4. Without transmission upgrades, 5076 MW of offshore wind capacity is installed in Regions 3 and 4, providing 91% of aggregate low-carbon electricity generated in those two regions. Relaxing transmission constraints results in approximately double the total existing statewide transmission capacity, an additional 8486 MW onshore wind capacity in Region 1, and no offshore wind power. The sensitivity of the onshore vs. offshore wind capacity selection to transmission upgrades motivates a more detailed analysis of model cost assumptions.

4.1.1 Transmission expansion

Increasing transmission costs result in the model selecting less additional transmission capacity to accommodate (less expensive) onshore wind power, instead calling for increasing amounts of (more expensive) offshore wind power. Figure II-2 shows the computed optimal mix of offshore wind, onshore wind, and new transmission capacity for the 50% RGT for a range of transmission costs and two offshore wind cost scenarios. Transition capacity expansions at each interface are added (in MW) even though distances between regions vary. Although transmission costs are generally considered in units of $/MW-mi, the assumed values in Table II-2 result in equivalent transmission cost upgrades in $/MW at each interface. Therefore, relative comparisons of new transmission capacities at different transmission cost estimates are equivalent to relative comparisons of total investment in transmission upgrades.

In all simulations, transmission capacity additions are selected only in the West-to-Southeast direction; specifically, this means transmission capacity additions are limited to (a) Region 1 to 2, (b) Region 2 to 3 and (c) Region 3 to 4. For comparison, we include very low transmission costs, shaded gray, which are not realistic in NYS but are on the order of assumptions used in other studies.

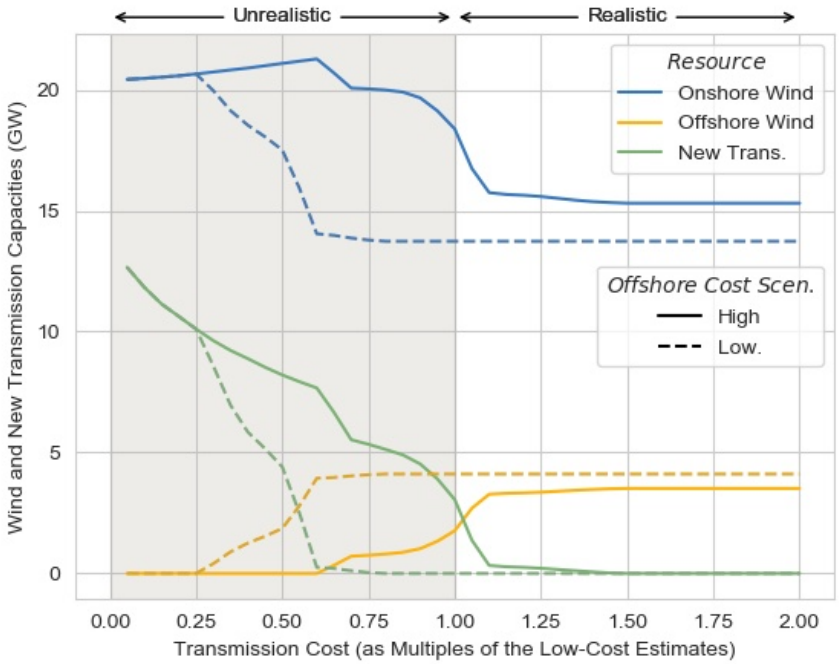

Figure II-2: Optimal wind power and new transmission capacities vs. transmission upgrade costs for the 50% RGT. All simulations include 3 GW of solar capacity. Transmission costs are presented as multiples of the low-cost estimates presented in Table II-2.

Figure II-2 shows that the cost-optimal generation mix is highly dependent on the price of transmission in the unrealistic cost range (until 1x the low-cost estimates). After this point, when transmission costs reach more realistic levels, the computed amounts of onshore and offshore wind capacity plateau, as new transmission is no longer selected as a cost-effective integration measure. Moreover, Figure II-2 demonstrates the influence of offshore wind costs on

the mix of wind capacity selected to meet the 50% RGT. In the low offshore wind cost scenario, offshore wind capacity increases more quickly at lower transmission prices and levels out at a higher quantity than in the high-cost scenario. Here, an increase in offshore capacity displaces a larger capacity of onshore wind, due to the higher relative capacity factor of offshore wind turbines.

Given that (1) we estimate new transmission in NYS to cost *at least* 1x the low-cost estimates presented in Table II-2 and (2) that offshore wind capacity costs will likely decrease to lower, internationally competitive rates, the model behavior shown in Figure II-2 implies that the most cost-effective method of meeting the NYS 50% renewable generation target may include no new interregional transmission.

The analyses described below investigate several model scenarios in detail. Having established the cost-sensitive behavior of new transmission, all following analyses assume transmission costs to be the average of the low and high transmission cost estimates; this is equivalent to a multiple of 1.5 of the low-cost estimates per Figure II-2.

4.1.2 Case-based optimal infrastructure mixes

To evaluate wind power capacity and transmission expansion sensitivity to other energy infrastructure measures, we exogenously impose battery storage, EV adoption, or increased solar PV capacity in certain model simulations, as shown in Table II-4. These capacities would not be selected in the RTM to meet the 50% RGT due to limitations of the model's formulation as a capacity expansion model. Battery storage and solar PV have quantifiable local system benefits beyond the scope of the model, and EV adoption is not included as a system planning-level decision variable with an associated cost assumption. In all cases, more qualitative considerations

of individuals or policymakers may also influence adoption and they are widely expected in future energy systems with deep VRE penetration. Table II-4 presents a summary of results for relevant high and low-cost scenarios.

Table II-4: Meeting the 50% RGT with different amounts of imposed energy infrastructure.

	Imposed Capacities[1]			High-Cost Scenario Results			Low-Cost Scenario Results		
				Optimal Cap. (GW)[2]			Optimal Cap. (GW)[2]		
Case	Batt. (GWh)	EVs (%)[3]	Solar (GW)	On.	Off.	Trans.	On.	Off.	Trans.
Base	0	0	3	15.31	3.52	0	13.74	4.11	0
1	33.6	0	3	15.93	2.72	0	14.22	3.37	0
2	0	25	3	16.94	3.93	0.26	14.87	4.79	0
3	0	0	6	14.73	2.76	0	13.21	3.33	0

Note: "On.", "Off.", and "Trans." correspond to the optimal computed amounts of onshore wind, offshore wind, and new transmission capacity, respectively.
[1] "Imposed Capacities" refers to capacities given as inputs to the RTM.
[2] "Optimal Cap." contains the quantities of energy infrastructure selected by the RTM.
[3] EV adoption rate refers to the percentage electrification of annual light-vehicle energy usage.

Simulation results indicate that energy storage reduces curtailment of onshore wind power, thus reducing the need for more expensive offshore wind capacity. Storage is most effective when co-located with low-carbon generation, as shown in Figure II-3. Co-located storage allows electricity to be stored either for later in-region use or for export when transmission lines out of the region are no longer congested. Because this behavior can be driven by both renewable energy and nuclear power, Figure II-3 presents the optimal distribution of this storage and how it relates to the mix of regional all low-carbon generation.

In Case 2, an EV adoption rate of 25% (flexible charging assumed) leads to computed capacities of onshore and offshore wind larger than those in the base case (see Table II-4). This increase in electricity load also makes transmission upgrades cost-effective in the high-cost

scenario. Installing another 3 GW of solar PV capacity (Case 3) reduces the total wind power capacity to meet the 50% RGT but affects onshore and offshore wind power differently. The computed decrease in offshore wind capacity is larger than the decrease in onshore wind capacity in both absolute and relative terms, as the added solar first displaces this more expensive generation.

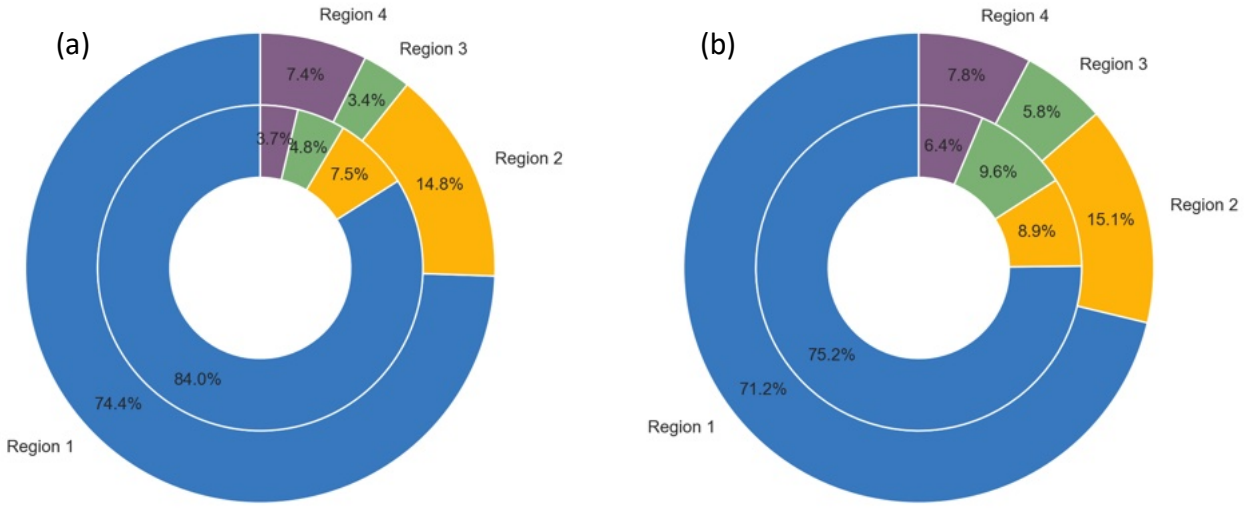

Figure II-3: Optimal standalone storage location and the regional, average low-carbon generation for the 50% RGT: (a) high-cost scenario; (b) low-cost scenario. The inner pie chart shows storage location (33.6 GWh total); the outer pie chart shows average uncurtailed low-carbon generation by region (13.03 GW in Figure II-3(a); 12.74 GW in Figure II-3(b)). Results include 3 GW of solar PV capacity present and no new transmission. In both cost scenarios, storage is most valuable when spatially paired with low-carbon generation.

Two key takeaways emerge from the results shown in Table II-4: First, at realistic prices, transmission expansion will not play an important role in meeting the 50% RGT, even with additional EV load present. Second, the optimal generation mixes in all cases, regardless of what background energy infrastructure is imposed, share a common characteristic: a considerable buildout of onshore wind capacity and a multiple-GW installation of offshore wind power. Our

analysis indicates that the pathways for meeting the 50% RGT vary little if storage, EVs, or additional solar generation is present at the scales modeled here.

4.2 Renewable generation growth beyond 50% energy penetration

To this point, our analysis has established that NYS may not need two planning-level actions to achieve a 50% RGT: Transmission capacity upgrades and standalone energy storage. This is largely driven by a more cost-effective strategy of utilizing local renewable generation (i.e., offshore wind near downstate load centers). In this section, we extend the model to deeper renewable energy penetration scenarios in order to investigate how longer-term considerations may affect planning decisions.

Figure II-4 presents the optimal capacities of wind power, new transmission, and battery storage under RGTs ranging from 50% to 80% for (a) the high-cost scenario and (b) the low-cost scenario; 0% and 25% EV adoption rates are considered. It is worth briefly noting that the 50% RGT points in Figure II-4 are consistent with the values presented for the base and 25% EV cases in Table II-4.

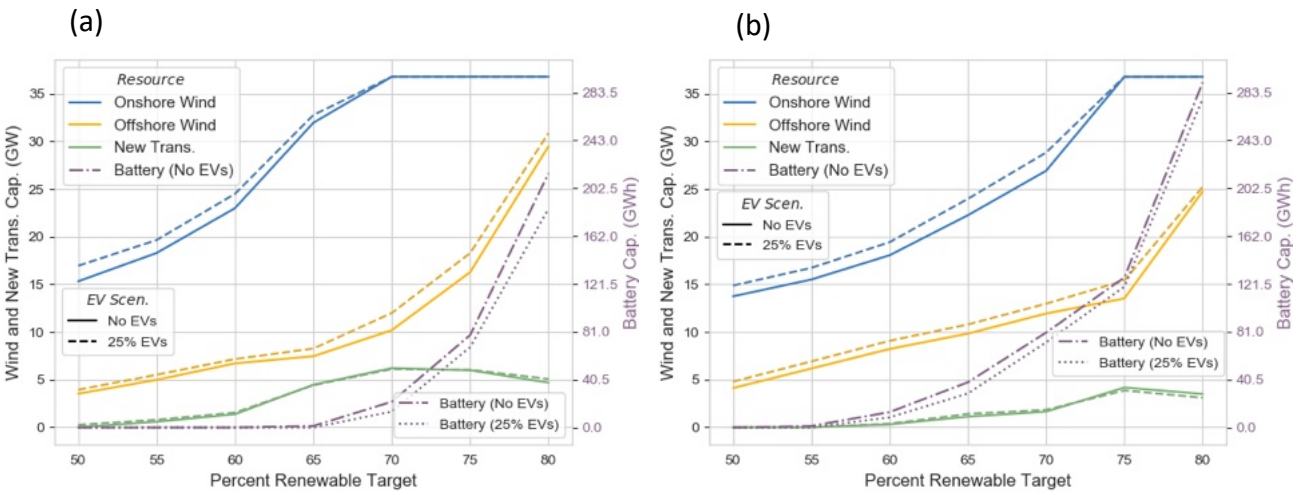

Figure II-4: Optimal mix of energy infrastructure for RGTs between 50% and 80% and for 0% and 25% EV adoption cases with flexible charging: (a) high-cost scenario; (b) low-cost scenario. Wind and new transmission capacities are shown in GW on left-hand axes; battery storage capacity is shown in GWh on right-hand axes. Results include 3 GW solar capacity.

The initial observation is that, as the scenarios extend beyond the 50% target, some level of system flexibility in the form of storage or transmission buildout is necessary to meet RGTs cost-effectively. The interplay observed in Section 4.1 continues between (1) onshore wind with transmission upgrades and (2) offshore wind power: With an accelerated offshore wind learning curve leading to lower costs, more installed offshore wind power reduces the computed amounts of onshore wind power and new transmission capacity.

Comparing Figures II-4(a) and II-4(b), up to approximately 70% renewable energy penetration, there is a tradeoff between investments in storage or transmission that appears highly dependent on the offshore wind and battery storage costs. Beyond 70%, the simulations predict sizable storage value regardless of the cost assumptions. For example, under the low-cost scenario, no storage is built to meet the 50% target; however, at the 65% target with 0% EV adoption, the model selects storage equivalent to 2.1 hours average demand, a quantity that

jumps to 15.7 hours average demand to meet the 80% target. At the same time, transmission upgrades remain relatively flat, indicating limited additional cost-effectiveness at deep renewable energy penetrations.

The effect of EVs on the overall results is relatively small beyond requiring some additional generation to meet a portion of the new electricity demand. Electric vehicle adoption decreases the amount of storage needed to meet all renewable generation targets, a result which implies that battery storage and flexible EV charging provide a similar service to the system – time shifting demand in order to aid integration – and thus act somewhat competitively. However, the limited scale of this reduction in computed standalone storage indicates a significant role for batteries to play in shifting supply even with EV flexibility present. As a point of comparison, at the 65% RGT in the low-cost scenario, the addition of EVs with a daily load equivalent to 33.6 GWh reduces the cost-optimal standalone storage by 9.47 GWh (24.7% of battery capacity). These results motivate additional investigation of the energy storage value.

4.2.1 Value to system of standalone energy storage

To simplify the discussion of results in this section, we consider only the low offshore wind and battery cost scenario based on the following:

1. The analysis presented in Section 4.1 and corroborated in Section 4.2 indicates that, regardless of cost assumptions, some substantial buildout of offshore wind power capacity is needed to meet a 50% RGT, suggesting offshore wind costs are likely to decrease in that time period.
2. Battery storage costs have rapidly decreased in recent years with projections expecting further reduction, particularly if EV adoption accelerates.

The value of a range of storage capacities – defined here as the avoided capital costs of renewable generation capacity and transmission upgrades – is computed for two representative RGTs: 50% (NYS 2030 goal) and 64% (together with 16% nuclear generation by energy, this yields 80% low-carbon electricity generation). We consider two scenarios for computing the system's avoided capital cost per kWh of storage: An assumed 3000-cycle battery life (Figure II-5(a)) and an assumed 10-year battery life (Figure II-5(b)). We offer these comparisons as it is unclear, at present, how long such batteries will operate; this is not meant to be a detailed analysis of battery operational effects or chemistries.

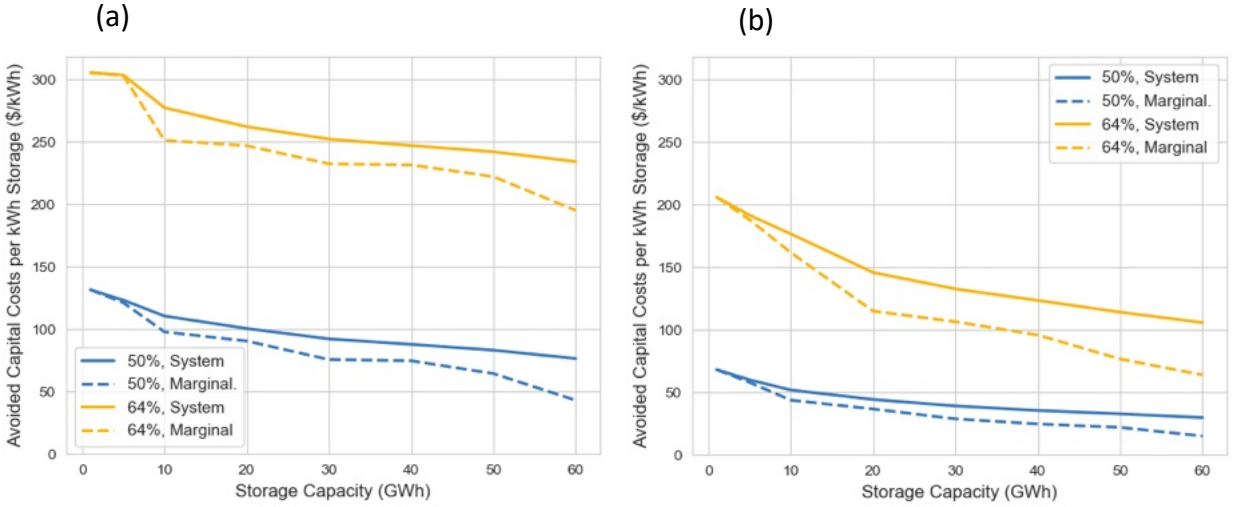

Figure II-5: Avoided capital costs per kWh storage ($/kWh) for storage capacities between 1 and 60 GWh. Results shown for 50% and 64% RGTs under the low-cost offshore wind assumption: (a) battery lifetime of 3000 cycles; (b) battery lifetime of 10 years. No electric vehicle adoption. Results include 3 GW of solar capacity.

In no simulations did a battery go through more than 3000 cycles in a 10-year period; as such, all simulations indicate higher battery value over 3000 cycles than over 10 years. We note a general trend of the computed storage values in the 64% target scenario being 2.5-3 times the computed values in the 50% target scenario, a difference which explains the large-scale storage

buildout observed at higher renewable energy targets: For a 10-year lifetime, the marginal value of energy storage exceeds its $100/kWh cost in the 64% target scenario up to approximately 35 GWh, whereas the energy storage value is less than its cost in the 50% target scenario. This general trend holds when EVs are introduced, but we can observe additional effects when including both EVs and 5 GWh of battery storage in our simulations, as shown in Figure II-6.

By comparing the results of flexible and inflexible EV charging for a 10-year standalone battery lifespan (Figure II-6), we note that flexible charging displaces some of the benefits from standalone storage for a 64% RGT, but there is a point of diminishing effect: The reduction in standalone storage value caused by flexible EV charging remains relatively stable beyond approximately 20% EV adoption. Perhaps counterintuitively, the value of standalone energy storage increases slightly in the 50% renewable target scenario. These effects imply that there is value in shifting supply to the time periods in which EV charging occurs *even if* that charging schedule is flexible.

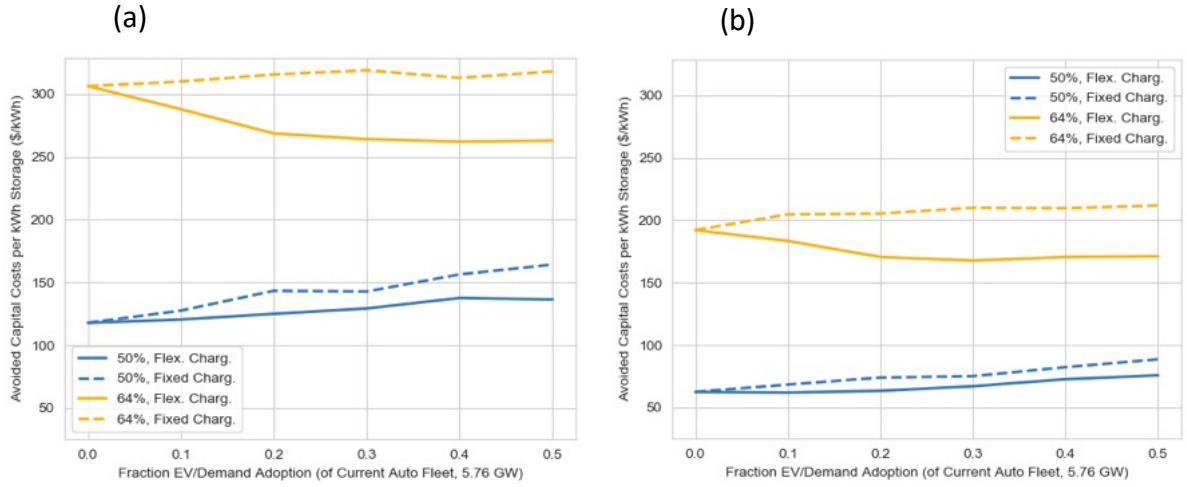

Figure II-6: Avoided capital costs per kWh storage ($/kWh) for EV adoption rates between 0% and 50% due to 5 GWh of storage. Results are shown for flexible and inflexible EV charging scenarios for 50% and 64% RGTs under the low-cost offshore wind assumption: (a) battery lifetime of 3000 cycles; (b) battery lifetime of 10 years. Results include 3 GW of solar capacity.

4.2.2 Total system cost breakdown

Figure II-7 presents the total capital cost breakdown for a NYS electricity system that achieves 50-80% RGTs for (a) high-cost and (b) low-cost scenarios. Comparing the two figures side-by-side, it is helpful to divide the RGTs into two sections: medium-term goals, consisting of targets between 50 and 65%, and long-term goals, represented by targets between 65% and 80%. In meeting the medium-term goals, the scale and distribution of investments are similar regardless of the cost scenario. At these levels of VRE integration, the RTM selects simultaneous, near-equal investments in onshore and offshore wind power, investments which make up the bulk of all system costs; computed expenditures on new transmission and battery storage are minor. To achieve the 65% RGT under a low-cost scenario, the state needs $84.3B (annualized cost of $6.0B, 0.39% of NYS 2017 GDP [159]); high cost assumptions increase this quantity to $91.7B (annualized cost of 6.5B, 0.42% of NYS 2017 GDP).

In both cost scenarios, wind power capacity (both onshore and offshore) contributes the vast majority of the total system cost until the deepest penetration rates, when marginal investments in both offshore wind power and battery storage are similar (and all onshore wind power sites have already been utilized fully). By comparing the cost breakdown results in Figure II-7 to the capacity optimizations in Figure II-4, we can see that while more investment is made in batteries in the high-cost scenario, the state would install more battery capacity in the low cost-scenario. If battery prices reach the lower estimate of $100/kWh, the increased amount of storage present allows the system to install less offshore wind power capacity and new transmission at all RGTs modeled; the model also selects less onshore wind power capacity at RGTs less than 75%. To meet the 80% RGT under low cost assumptions, the state needs $209.5B

(annualized cost of $14.9B, 0.96% of NYS 2017 GDP); high cost assumptions increase this quantity to $295.4B (annualized cost of $21.0B, 1.36% of NYS 2017 GDP).

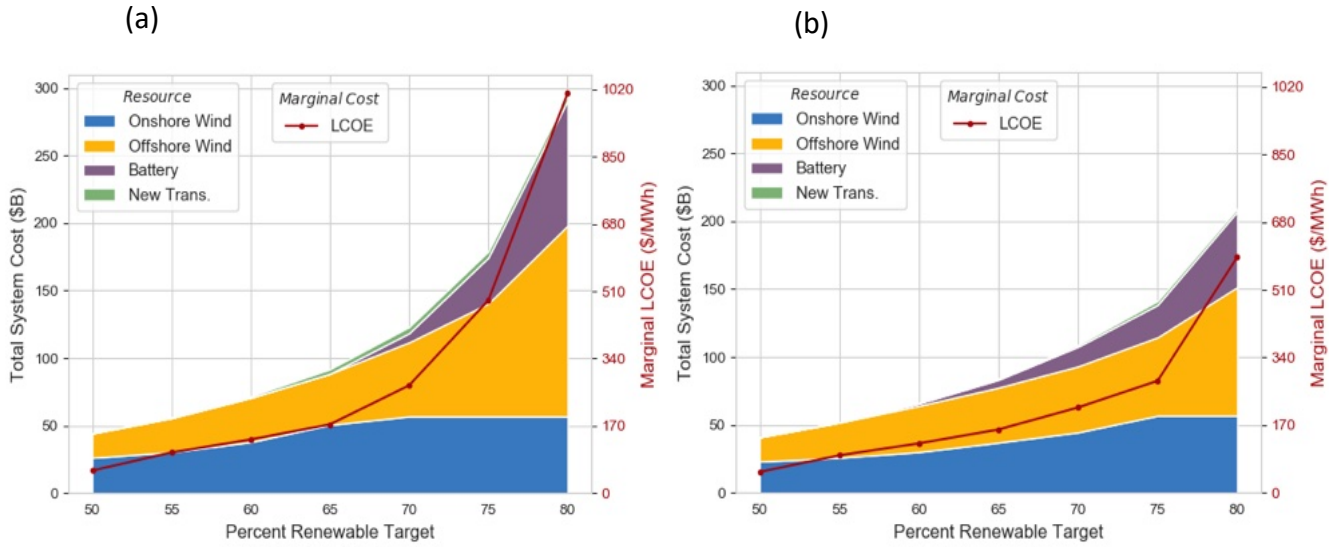

Figure II-7: Total system costs ($Billion) and marginal LCOEs ($/MWh) for RGTs between 50-80%: (a) high-cost scenario; (b) low-cost scenario. Both scenarios include 3 GW of solar PV capacity and 25% EV adoption with flexible charging. Total system costs are calculated over a 20-year lifespan; battery costs are accordingly adjusted based on an assumed 10-year lifespan.

By comparing the marginal LCOEs at a range of RGTs, we see that the price of additional utilized renewable energy accelerates as VRE penetration increases. For the low-cost scenario, our calculations indicate that a 50% RGT can be achieved with LCOE of $52/MWh for new VRE. Between the 50-55% RGTs, the marginal LCOE of utilized VRE increases to $94/MWh; between 75-80%, the marginal LCOE rises sharply to $592/MWh. For RGTs through 65%, the high-cost scenario marginal LCOE is computed to be less than 10% greater than the low-cost scenario marginal LCOE. Beyond this point, the computed marginal LCOE values diverge: The high-cost scenario LCOE is 27% higher at the 70% RGT and 70% higher at the 80% RGT.

5. Discussion

At different points along a pathway to deep energy decarbonization, the value of particular resources is likely to vary. Because such considerations have not permeated renewable energy integration studies, Chapter II investigates several likely large-scale energy planning decisions in the context of near-to-long-term renewable generation targets (RGTs): Whether to build new intrastate transmission to connect high wind areas to load centers, whether to invest in dedicated energy storage to align supply and demand, and which generation resources to prioritize in a geographically heterogeneous region. The effects of wide-scale adoption of EVs on these decisions is also investigated. Electricity grid regions at the state or multiple-state level can generally contain large distances between the most economical variable renewable energy (VRE) resources and the largest load centers; more expensive local or nearby resources may make up a sizable part of a pathway to deep VRE penetration. To complete an analysis at this scale, we simulate New York State's (NYS) regional electricity system, which is representative of regions with transmission-linked zones and heterogeneous demand and potential supply.

When we adopt existing interzonal transmission constraints and realistic transmission cost assumptions, we find that the computed cost-optimal buildout of new transmission capacity is less than what other analyses propose. In fact, we demonstrate that NYS can most cost-effectively meet a 50% renewable generation target (RGT) with no new interzonal transmission capacity. (We do not investigate smaller-scale transmission that may be needed to connect wind power sites themselves to the larger grid.) We attribute the difference between our results and those of other studies to the tendency for other models to underestimate transmission costs or assume deployment that does not consider that new infrastructure will handle only the marginal

increases in transmission; if this new transmission comes solely at times of very high VRE production, the low utilization rate may render new capacity uneconomical even if it appears inexpensive compared to other renewable integration measures. It is not necessarily given that considerations of dispatchable generation to meet net loads would not affect the results of this type of analysis; however, NYS's existing regime of well-dispersed gas-fired generation is unlikely to transform into one in which net loads are met by distant fossil fuel-based generation.

A central result of the analysis is a cost-minimal solution to meeting a 50% renewable energy penetration level that includes only a large buildout of onshore wind generation (~15 GW) and a multi-GW expansion of offshore wind capacity; the inclusion of energy storage, electric vehicles, or additional solar capacity do not meaningfully change this infrastructure mix. At this renewable penetration level, the value of battery storage remains below even the most optimistic cost assumptions. Similarly, the scale of electric vehicles modelled here – 25% adoption – does not provide enough system flexibility to substantively change the computed optimal generation mix. Alongside our transmission findings, this suggests that a 50% RGT can be achieved solely through a cost-effective buildout of VRE generation capacity; further, our computed LCOE of new renewables to achieve this target ($52/MWh) is in line with reasonable current generation costs.

While co-locating dedicated energy storage with variable supply improves VRE integration, the computed optimal resource mix to achieve the 50% RGT includes no standalone energy storage, even with costs as low as $100/kWh. Though storage is already proving to have other value (e.g., grid frequency regulation services and peak demand reductions), it is unlikely to be cost-effective for shifting energy supply to times of higher demand in achieving near-term

RGTs. Since battery storage capacities are aggregated at the regional level, we note that further study is needed to investigate how the intraregional distribution of these resources and local conditions would affect operation.

In exogenously doubling the amount of solar PV capacity in our model (6 GW vs. 3 GW), we find limited effects on our overall findings. The additional solar tends to displace local resources: Onshore wind power in less densely populated western regions and offshore wind power in high load centers along the coast. That said, even though the regional distribution of solar PV capacity is similar to the regional distribution of demand, the absolute and relative reductions of offshore wind power capacity are greater than the reductions of onshore wind power capacity.

These findings are likely generalizable to other regional grids, particularly those along the U.S. Atlantic coast that contain high population density areas with limited local or nearby VRE resources other than offshore wind power. Similar findings may also apply to inland areas with access to distant onshore wind resources and more expensive local solar potential. A useful metric from our analysis is that, although 3-4 times more onshore wind than offshore wind capacity is built to reach the 50% RGT, the total financial investments of onshore and offshore wind are roughly equal.

Moving next to RGTs in the 50-65% target range, we compute growth in both transmission and storage capacity, although the infrastructure mix is dependent on assumed offshore wind and battery costs. If battery and offshore wind costs remain high, additional transmission to better utilize onshore wind power will become necessary; however, if battery and offshore wind costs drop as predicted, battery storage will provide the system flexibility necessary to integrate

additional VRE. The increased reliance on energy storage as NYS approaches the 65% target is primarily a result of the computed value of storage increasing to 2.5-3 times greater than its value at the 50% target. As near-term targets transition to longer-term goals, model-selected energy storage capacities increase from no storage at the 50% RGT to storage equivalent to approximately 2 hours average demand at the 65% RGT, and finally to 16 hours-equivalent storage at the 80% RGT. Such scales indicate that some portion of the storage requirement may be met by alternatives to batteries, a hypothesis that deserves further study.

At a 64% target (with 16% nuclear power, a total of 80% electricity generation by low-carbon sources), the value of dedicated energy storage holds relatively steady even with up to 50% adoption of flexible charging EVs. While flexible charging does decrease the value of the energy storage, all reduction in value occurs within the first 20% of EV adoption. These effects imply that there is value in shifting VRE supply to the time periods in which EV charging occurs even if that charging schedule is itself flexible. This explains why our analysis shows that large-scale adoption of EVs, even with flexible charging operation, is unlikely to alter the overall approach to meet RGTs, aside from the self-evident need for additional capacity to meet the demand.

Overall, our calculations indicate that a 50% RGT can be achieved with a LCOE largely in line with current reasonable generation prices that gradually increases at deeper penetration rates. Between the 50-55% RGTs, the marginal LCOE nearly doubles, but this reflects an increased share of offshore wind-generated electricity utilization and would not necessarily represent a large electricity price increase for the urban areas making most use of the offshore wind resource. Between the 75% and 80% RGTs, the marginal LCOE of utilized VRE rises sharply. This surge in

price can be attributed to the large amounts of wind and storage capacity needed to meet long-term targets. As these cost estimates would prove prohibitive from an investment standpoint, NYS will likely turn to other integration measures or future technologies not analyzed here to achieve long-term renewable energy goals. These measures could also include further connections to neighboring grids and large-scale electrification of heating.

In general, this study demonstrates that near-term renewable energy goals can be achieved most cost-effectively through VRE capacity buildout alone. Beyond this point, policy and investments that bring down the costs of nascent energy technologies – here, offshore wind and battery storage – will be particularly important. Even at high costs, significant shares of such technologies would be required to achieve deep energy decarbonization; an approach that incorporates them into near-term planning may make the longer-term transition more affordable.

6. Conclusion

Chapter II presents an optimization model for a regional electricity grid to assess cost-minimal generation, transmission and storage capacities required to meet a series of renewable generation targets. We compare results for a range of transmission costs and for two broader technology cost assumption scenarios. Additional insight is gleaned from exogenously varying standalone energy storage capacity, adoption levels of electric vehicles, and solar photovoltaic capacities. The first half of this chapter investigates a 50% target in detail; the second half extends the analysis to targets up to 80%.

It is first demonstrated that New York can most cost-effectively realize a 50% renewable generation target with no new transmission or large-scale storage. Assuming offshore wind generation costs decrease to internationally competitive levels, the optimal generation mix includes 13.7 GW of onshore and 4.1 GW of offshore wind power capacity; despite its high capital costs, offshore wind's high capacity factor, proximity to load centers and lack of reliance on long-distance transmission upgrades result in its selection. Here, the contribution of 4.1 GW of offshore wind represents 28% of new potential renewable energy production. Even with 33 GWh of storage or 6 GW of solar exogenously imposed, optimal offshore wind capacity does not fall below 3.3 GW.

As renewable energy penetration increases from 50% to 80%, the model builds out all of New York's available 37 GW of onshore wind capacity and dramatically increases offshore wind capacity to 25 GW. At the 80% target, offshore wind contributes 48% of new potential renewable energy generation in the state, and as this generation connects directly to downstate load centers, additional statewide transmission buildout is limited to 5 GW and 1% of the total financial investments in generation, storage and transmission. This overall picture is largely insensitive to the cost of offshore wind capacity, or the presence of electric vehicles.

Storage plays an increasingly important role at higher renewable energy targets. In our results, we compute no storage at the 50% target, storage equivalent to about 2 hours average demand at the 65% target (7% of investment), and 16 hours average demand at the 80% target (27% of investment). The rising value of storage drives this growth: the avoided capital cost per kWh of storage is nearly 3 times greater at 65% renewable penetration than at the 50% penetration level. At deeper renewable energy targets, large computed battery capacities with

relatively low annual cycling suggest that some proportion of alternative forms of storage may be effective.

Evaluating the marginal costs of utilized renewable electricity between the 50% and 80% targets, we observe a marked increase in price as New York requires more capital to build out and integrate additional generation capacity. To achieve a 50% target, we compute a levelized cost of electricity of $52/MWh for new renewable energy, a quantity in line with current reasonable generation prices. Between 50-55%, the marginal cost of electricity nearly doubles to $94/MWh, which may be reasonable for the urban areas considered in this study. To reach the 80% target, this marginal cost increases sharply to $592/MWh. In the 70-80% target range, the rapidly increasing marginal costs are largely driven by energy storage costs.

It is important to understand a couple limitations of the study. We have not modelled interconnections to adjacent grids, and these may play an important role in lowering costs at higher renewable energy penetration levels. We have also not simulated unit commitment or dispatch with deep penetration of renewables; transmission and storage may affect the internal economics of dispatchable generators, market prices and dispatchable capacity needs, warranting further study. Lastly, we do not model anticipated heating electrification, which is likely to reduce wind power curtailment in winter months.

Chapter III: The value of energy flexibility: Integrating wind resources in New York State

Abstract

The following chapter explores the effects of energy system flexibility on the contribution of wind to the New York State (NYS) electricity generation mix. First, the benefits of NYS-specific flexible hydropower are investigated. For all simulations, a mixed integer linear program minimizes net load to determine the maximum aggregate capacity factor for the installed wind power. The benefits of three different types of energy flexibility are then explored: flexible supply, flexible demand, and bidirectional flexibility (i.e., energy storage). To compare across technologies, a novel method of standardizing flexibility inputs, Potential Flexible Energy (PFE), is introduced.

With 30 GW wind capacity in NYS (average electricity demand of 18.7 GW), introducing electric vehicles with an average load of 1.44 GW and daily available battery capacity of 34.5 GWh (roughly equivalent to the daily use of 3.4 million passenger EVs) increases statewide wind utilization by 840 MW (9.0% of wind potential and 4.5% of average load). Added flexibility in the form of energy storage yields similar results: with 3.2 GW charge/discharge capability and 76.8 GWh storage capacity, statewide wind utilization increases by an average of 660 MW (7.0% of wind potential and 3.5% of average load).

Because of transmission constraints and the geographic distribution of high-potential wind resources, increased wind utilization is only achieved when flexibility is added in the region where 86% of the 30 GW simulated wind capacity is located.

1. Introduction

The use of renewable energy technologies to decrease fossil fuel emissions and mitigate the effects of climate change are well known [119]. Driven by decreasing costs and growing societal awareness of the hazards of burning carbon fuels, solar and wind power penetration will increase throughout next century, as will the challenges of integrating these inflexible, low carbon resources [121,160]. Among energy system planners, there is much debate over how to meet these challenges, which arise from the intermittent and stochastic nature of wind and solar generation [50,161,162]. The technologies and methods under discussion enable renewable energy integration in various ways, and include: carbon pricing; a well dispersed portfolio of energy sources [12]; advanced grid monitoring and communication [122]; expanded transmission capabilities [163]; interconnection between regional systems [124]; and increased system flexibility [125]. Previous research has focused on the benefits and challenges of deep penetration of Variable Renewable Energy (VRE) resources in New York State by significantly increasing the modeled capacity of wind and solar installations [64,129]. This chapter will investigate the effects of expanded wind capacities with varied amounts of complementary flexibility. Here, flexibility is defined as "the extent to which a power system can modify its electricity production and consumption in response to variability, expected or otherwise". First, the potential of hydropower to offer system flexibility and to increase consumption of wind generated energy is explored, as flexible use of hydroelectric resources offers clear benefits to power systems [164–166]. The authors conducted a comprehensive review of the state's hydropower resources in order to create a mixed-integer linear program (MILP), the *Flexible Hydropower Model* (FHM), which optimizes NYS system response to up to 30 GW of wind capacity

by minimizing net load. The simulated hydropower flexibility reflects current load-following capabilities of NYS hydro resources. To accurately assess the benefits of hydropower flexibility, solutions from the MILP are compared to a baseline case in which the state's hydroelectric generation is the result of smoothing reported generation values over two weeks. Expanding the scope of study, this chapter next evaluates the impact of three different types of power system flexibility: supply-side flexibility, demand-side flexibility, and bidirectional flexibility. A similar MILP, the Comparative Flexibility Model (CFM), was created to complete this analysis, where results are again compared to a baseline case without flexibility.

The structure of Chapter III is as follows: Section 2 presents background information on hydropower resources in NYS. Section 3 details the FHM and CFM methodologies and the data inputs used to parameterize and run the models. Section 4 discusses the FHM and CFM results. Section 5 restates the chapter's most salient conclusions and offers directions for future work.

2. Background

2.1 NYS hydropower

In New York State, four hydroelectric power plants constitute 80% of the state's total hydropower capacity and 83.2% of all hydro generation, providing 13.3% of total statewide electricity generation. The largest of these generation facilities is the Robert Moses Niagara Hydroelectric Power Station. The Moses-Niagara plant has a rated capacity of 2460 MW, and in 2016, generated 54% of the state's hydropower [151]. In accordance with the 1950 Niagara Treaty between the United States and Canada, a portion of the Niagara River is diverted into the Moses-Niagara forebay after allowing for the necessary flow of water over Niagara Falls and into

the Sir Adam Beck Hydroelectric Generation Facility, a hydro-plant on the Canadian side of the river [167].

The Moses-Niagara forebay also serves as the lower reservoir of the Lewiston Pumped Hydro Energy Storage (PHES) plant. During periods of low statewide electricity consumption, the Lewiston Plant pumps water from the Moses-Niagara forebay into its upper reservoir. During periods of high statewide demand, the Lewiston Plant generates electricity by allowing water to flow from its upper reservoir back into the Moses-Niagara forebay. The generating capacity of Lewiston is 240 MW. In 2016, the Lewiston plant generated 1.7% of NYS hydropower [151].

The Moses-Saunders Power Dam provides the second most hydropower to the state. The dam straddles the St. Lawrence River and diverts water to two adjacent power stations, the American St. Lawrence-Franklin D. Roosevelt Power Project and the Canadian R.H. Saunders Generating Station. Both facilities operate as run-of-the-river plants with limited storage capability. The St. Lawrence-FDR facility has a total rated power capacity of 912 MW. Because of reliable flow, St. Lawrence-FDR operates near its maximum generation capacity nearly all year long. In 2016, St. Lawrence-FDR generated 26.0% of NYS hydropower [151].

The last of the four large hydroelectric power plants is the Blenheim-Gilboa PHES facility, which has the second largest total rated turbine capacity in the state at 1160 MW [168]. Its upper reservoir has a capacity of 18 million cubic meters, which corresponds to approximately 14 hours of peak generating capability. In effect, Blenheim-Gilboa operates as a closed system with 73% efficiency, as the Schoharie Creek water replenishes water lost or evaporated. The Blenheim-Gilboa plant supplies the New York Independent System Operator (NYISO) with black-start capability; in 2016, it supplied 1.4% of NYS hydropower [151].

Smaller hydropower facilities constitute the remaining 20% of hydroelectric resources in New York, supplying 16.8% of the state's hydroelectric power [151]. Throughout the following analysis, these plants are lumped together to yield one representative facility, subsequently referred to as `Small Hydro', with a rated power capacity of 1230 MW.

2.2 Moses-Niagara energy inflow

This study computes energy inflow to the Moses-Niagara forebay based on flow measurements of the Niagara River [169] and the specifics of the 1950 Niagara Treaty between the US and Canada [167]. The Niagara Treaty establishes that 100,000 cubic feet per second of water must pass over Niagara Falls, downstream of the hydropower facility intakes, between the hours of 8am and 10pm EST from April 1st to September 15th, inclusive; and each day between 8am and 8pm EST from September 16th to October 31st, inclusive. The flow over the falls should never drop below 50,000 cubic feet per second at any other time. On the American side, the drawing capacity of the Moses-Niagara plant is limited to 109,000 cubic feet per second by the size of the canal that diverts water from the Niagara River to the forebay.

After applying these constraints to the river flow readings, the authors determine the time series for water inflow to the Niagara-Moses facility. With assumed turbine efficiency of 0.90, generator efficiency of 0.96, and a 91.44m head, the yearly energy inflow to Moses-Niagara was calculated to be 14.409 TWh in 2016, nearly equal to the EIA-reported energy generation of 14.410 TWh that year [151]. As such, the authors feel justified using this method to determine the Niagara-Moses energy inflow time series.

2.3 St. Lawrence-FDR energy inflow

The Ontario ISO, IESO, publishes hourly generation and generation capacity data for every plant larger than 20 MW in the region [170]. Because of the twin nature of the St. Lawrence-FDR and the R.H. Saunders hydroelectric facilities – both facilities are supplied by the Moses-Saunders Power Dam and are therefore subject to the same water inflows and weather conditions – the authors assume that the IESO-reported hourly generation capacity for 2016 applied similarly to the St. Lawrence-FDR plant after scaling for differences in rated power capacity (912 MW for St. Lawrence-FDR and 1045 MW for R.H. Saunders). On account of the St. Lawrence-FDR facility's operation as a run-of-the-river hydroelectric plant, no storage capacity is assumed to exist. With a generator efficiency of 0.96, the yearly energy inflow to St. Lawrence-FDR is calculated to be 7.05 TWh in 2016, slightly less than the EIA-reported generation of 7.10 TWh that year.

2.4 Small Hydro energy inflow and storage

The energy inflow to the representative Small Hydro facility, $H_{in,sh}^{t}$, is determined by smoothing the difference of the NYISO-reported hydropower generation (H_{nyiso}^{k}) and the calculated energy inflow to both the Moses-Niagara ($H_{in,nia}^{k}$) and St. Lawrence-FDR ($H_{in,stl}^{k}$) generation facilities with a smoothing factor τ = 336 hours (2 weeks):

$$H_{in,sh}^{t} = \frac{1}{\tau + 1} \sum_{k=t-\frac{\tau}{2}}^{t+\frac{\tau}{2}} H_{nyiso}^{k} - H_{in,nia}^{k} - H_{in,stl}^{k} \tag{III-1}$$

We believe the approximation is justified: the large generation facilities not accounted for in this equation (Lewiston, Blenheim-Gilboa) are pumped storage plants and accordingly have

no significant energy influxes. Therefore, all other hydro inflows must be captured by the remaining Small Hydro plant. To ensure that the Small Hydro inflow is not a direct response to the Moses-Niagara and St. Lawrence-FDR inflows, a smoothing factor τ-value of 2 weeks was used to smooth the difference of the hydro production and hydro inflows without allowing for the possibility of long-term storage. Smoothing the difference between hydropower production and known inflows ensures that the scale of the Small Hydro energy inflow is appropriate. Such an approach also ensures that $H_{in,sh}^{t}$ responds to macro-trends in water availability but not to small fluctuations in the flow of disconnected rivers. For the Small Hydro representative facility, the energy from 24 hours of maximum generation was assumed as storage capacity.

2.5 No flexibility baseline

To compare the results of flexible hydropower operation, we analyze a baseline, fixed hydropower simulation. For this control scenario, fixed hydropower generation at time *t*, H_{fix}^{t}, was set equal to the amount of hydropower reported by NYISO for that time step, H_{nyiso}^{k}, smoothed over τ = 336 hours:

$$H_{fix}^{t} = \frac{1}{\tau + 1} \sum_{k=t-\frac{\tau}{2}}^{t+\frac{\tau}{2}} H_{nyiso}^{k} \quad \text{(III-2)}$$

3. Methodology

3.1 Flexible hydropower model overview

The Flexible Hydropower Model uses the parameters of NYS hydropower facilities described in Section 2 to produce a MILP that minimizes net load in New York given capacities of

installed wind generation and available PHES. Minimizing the sum of statewide net load achieves the highest degree of wind energy utilization for the assumed set of constraints. The MILP is solved at an hourly time resolution and ignores interzonal transmission constraints. The objective function for the FHM is presented below:

$$minimize \qquad \sum_{t \in T} NL^t \tag{III-3}$$

where,

$$NL^t = D^t_{elec} - \left(H^t_{nia} + H^t_{lew} + H^t_{bg} + H^t_{sh} + H^t_{nc}\right) - \left(X_{on} * W^t_{on} + X_{off} * W^t_{off}\right) - \left(X_{btm-solar} * W^t_{btm-solar} + N^t + H^t_{in,stl}\right) \tag{III-4}$$

Here, the net load, NL^t, is defined by the exogenously determined statewide electricity demand, D^t_{elec}; and baseload renewable generation, equal to the sum of (1) nuclear generation, N^t; (2) behind-the-meter solar generation, $X_{btm-solar} * W^t_{btm-solar}$; and (3) generation at the St. Lawrence-FDR facility, $H^t_{in,stl}$. It is also defined by the following state variables: generation at Moses-Niagara, H^t_{nia}; generation at Lewiston, H^t_{lew}; generation at Blenheim-Gilboa, H^t_{bg}; generation from the Small Hydro facility, H^t_{sh}; generation from any new, simulated PHES, H^t_{nc}; and utilized wind energy, $X_{on} * W^t_{on} + X_{off} * W^t_{off}$. A full list of constraints can be found in Appendix B. The model is formulated in MATLAB [171] and solved in Gurobi, a commercial optimization solver [172].

Beginning December 2015, NYISO has published fuel mix data for the NYS electricity grid at 5-minute intervals [38]. The fuel mix data present the amount of total power supplied by generators classified as: nuclear, hydro, natural gas, dual fuel, wind, other renewables, and other

fossil fuels. Five minute reported hydro data for 2016 – along with Niagara River streamflow and St. Lawrence rated generation capacities – are used to determine the Small Hydro energy inflow. After parameterizing the state's hydro resources, we use 6 years of NYISO-reported hourly demand data (2007-2012) to run the FHM. Nuclear generation in the state was taken as a constant 3026 MW, based on the annual nuclear energy reported by NYISO after removing the contribution from Indian Point Energy Center, as this plant is slated to be decommissioned as soon as 2021.

For some scenarios, the FHM simulates additional PHES in NYS. This additional storage is scaled in reference to the Blenheim-Gilboa facility: a scenario denoted “Flex Hydro + 3 PS” indicates statewide flexible hydro generation with additional PHES three times the size of Blenheim-Gilboa in both generation and reservoir capacity. Energy generated by this simulated plant, H_{nc}^{t}, allows the FHM to interrogate the benefits to an energy system when supplementary flexibility is present.

To simulate capacities of wind power far exceeding current levels, the FHM relies on model wind power data for 126,000 potential wind sites developed by the National Renewable Energy Laboratory (NREL) [62,63]. A previous study found the NREL model to significantly overpredict the electricity generated at actual sites in NYS and developed a procedure to adjust the time series to reflect actual output [64]. For the current study, we employ the modified wind generation time series from this earlier study. In simulating increased wind power capacity, it was assumed that wind turbines are first installed at locations with the highest potential electricity generation, with additional turbines installed at progressively less productive sites.

The NYSERDA Distributed Generation Integrated Data System is used in order to determine solar potential in NYS [173]. This resource reports the hourly generation time series of solar facilities in New York. With the rated capacities of these facilities, the hourly solar generation potential is calculated for the 10 largest plants in New York State that were operational for the entirety of 2016. These time series were averaged to yield the generation potential of a representative plant, which is then scaled by the installed solar capacity. We assume that 600 MW of solar capacity exists in NYS when running the FHM.

3.2 Comparative flexibility model overview

The CFM also utilizes a MILP to minimize the sum of statewide net load given the presence of varying levels of VRE capacity and system flexibility. The objective function for the CFM is given as:

$$minimize \qquad \sum_{t \in T} \sum_{i \in I} NL_i^t \tag{III-5}$$

where,

$$NL_i^t = D_{elec,i}^t + D_{flex,i}^t - H_{flex,i}^t - \delta_{ps,i}^t + \gamma_{ps,i}^t - \left(X_{on,i} * W_{on,i}^t + X_{off,i} * W_{off,i}^t\right) - \left(X_{btm-solar,i} * W_{btm-solar,i}^t + N_i^t + H_{fix}^t\right) + \sum_{i' \in I} [\,(1-l) * Z_{i'i}^t - Z_{ii'}^t] \tag{III-6}$$

For every region *i*, at each time step *t*, the net load NL_i^t is partly defined by two exogenous variables: electricity demand, $D_{elec,i}^t$; and baseload renewable generation, equal to the sum of (1) nuclear generation, N_i^t; (2) solar generation, $X_{btm-solar,i} * W_{btm-solar,i}^t$; and (3) fixed hydropower generation, $H_{fix,i}^t$. The following decision variables are also used: flexible demand, $D_{flex,i}^t$; generation from flexible hydropower, $H_{flex,i}^t$; pumped storage generation, $\delta_{ps,i}^t$; pumped

storage pumping, $\gamma_{ps,i}^{t}$; utilized wind energy, $X_{on,i} * W_{on,i}^{t} + X_{off,i} * W_{off,i}^{t}$; and net imports from adjacent regions after accounting for transmission losses, $\sum_{i' \in I}[(1-l) * Z_{i'i}^{t} - Z_{ii'}^{t}]$. A list of model constraints can be found in Appendix B. By minimizing the sum of the statewide net load, the CFM minimizes the amount of electricity generated by fossil fuel-based sources. The CFM is also formulated in MATLAB and solved using Gurobi.

To simulate the spatial characteristics of load and generation in New York, this study condenses the 11 NYISO load zones (A-K) into four regions (1-4) with interregional transmission limits equivalent to those given in [48]. The boundaries of these regions align with the main NYISO transmission system interfaces. Figure III-1 shows the 11 NYISO load zones. The load zones, wind capacity (in the 30 GW wind scenario), and average electricity demand of each region are presented in Table III-1. Transmission losses are assumed to be 3% between adjacent regions, ensuring that wind electricity is first used to meet demand nearest the region in which it was generated.

In the CFM, nuclear generation in the state is taken as a constant 3026 MW (all in Region 1), and solar and wind generation are calculated as described for the FHM (150 MW solar in each region). The daily hydro output is estimated based on the actual 2007-2012 monthly output for facilities in each zone [38] and a cubic spline function constrained to be continuous and smooth from month to month. In the flexible scenarios, a fraction x_{flex} of the total hydro generation, $H_{tot,i}$, is assumed flexible, able to be dispatched with a degree of control; the balance of the hydro generation is simulated as non-flexible and needs to be consumed during the time-step it becomes available.

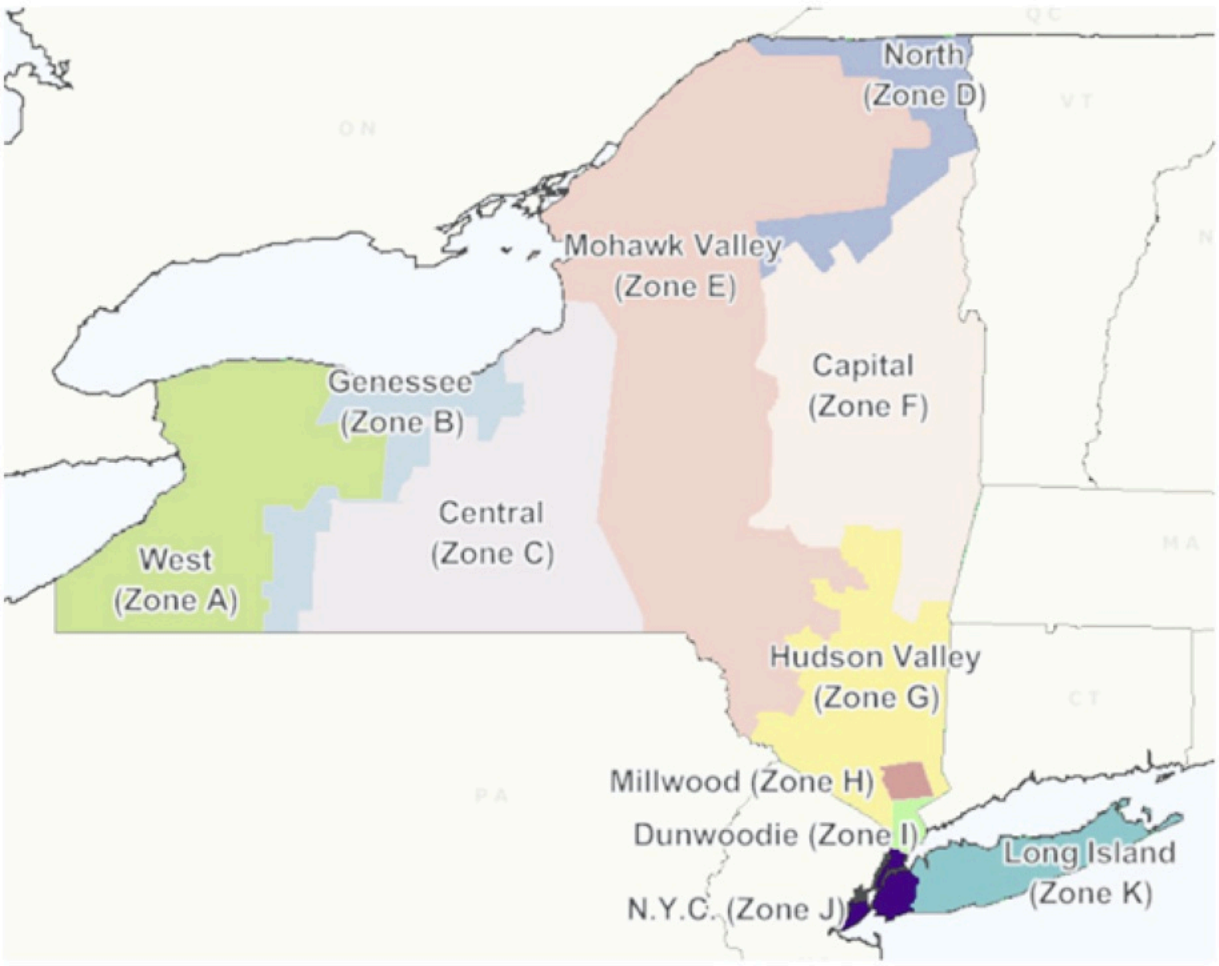

Figure III-1: NYISO control area load zones [174].

Table III-1: Spatial grouping of NYISO load zones; Distribution of wind capacity in 30 GW scenario [64] ; Average 2007 -- 2012 electricity demand [38].

Region	NYISO Zones	Wind Cap. (30 GW)	Avg. Demand
1	A, B, C, D, E	25814 MW	6382 MW
2	F, G	3358 MW	2495 MW
3	H, I, J	16 MW	7211 MW
4	K	812 MW	2567 MW

3.3 Modes of flexibility

In the Comparative Flexibility Model, three different modes of flexibility are analyzed. The first is supply-side flexibility, of which a traditional hydroelectric plant is an example. Such a facility maintains a degree of control over when it generates electricity based on its ability to

store water in its reservoir. The second is demand-side flexibility (e.g., a fleet of electric vehicles with some discretion over when it can charge); this type of flexibility offers a degree of control over when certain electric loads can be met. The third type of flexibility is bidirectional energy storage; a plant with bidirectional flexibility can both absorb and deliver power, and thus has one greater degree of control than either flexible supply or flexible demand facilities.

3.4 Potential flexible energy

In order to compare the effects of the three types of flexibility, this study proposes a metric called *Potential Flexible Energy* (PFE). PFE is defined as the potential amount of energy a flexible resource can generate or absorb over a period of time. By equating PFE in all three flexibility cases, the authors ensure that no mode of flexibility allows the system to utilize more flexible energy over the analyzed time period. Because the analysis is performed hourly over 6 years (T=52608) and the different modes of flexibility operate on different time scales, PFE is presented as an average hourly quantity. This has the units MWh/h; for simplicity, the authors use units MW. In this analysis, PFE is calculated before efficiency losses (η_{hydro} = 0.894; η_{veh} = 0.95).

PFE is determined in the following ways for the three modes of flexibility:

1. For flexible supply (traditional hydropower), PFE is equated to the average amount of energy that flows into the reservoir, H_{in}^{t}, over the analyzed time period:

$$PFE_{fh} = \frac{1}{T}\sum_{t \in T} H_{in}^{t} \quad \text{(III-7)}$$

The power of the supply-side facility is determined by dividing the average inflow power by an assumed capacity factor (CF) equal to 0.45; the reservoir is sized to hold 24 hours of max generation.

2. For flexible demand (electric vehicles), PFE is equated to the constant hourly EV load, D_{veh}:

$$PFE_{veh} = D_{veh} \quad \text{(III-8)}$$

It is assumed that the daily EV load – 24*D_{veh} – can be met anytime between 7pm and 7am. The power absorption capacity (i.e., maximum rate of charge) for the EVs is set to 6*D_{veh}.

3. For energy storage (i.e., PHES), PFE is computed from the facility's power generation capacity, X_{ps}, and a maximum capacity factor:

$$PFE_{ps} = CF_{max} * X_{ps} \quad \text{(III-9)}$$

If we were to ignore charging and discharging efficiencies, the maximum possible capacity factor would be 0.5, since all energy provided in discharge must be stored by charging; however, we consider these efficiencies and arrive at CF_{max} = 0.45. The energy storage capacity of the resource is 24 hours of peak discharge capability. Therefore, with CF_{max} = 0.45, the energy storage capacity and discharge capability equal those of the traditional hydropower model for a given quantity of PFE.

For PFE = 1440 MW, the power and energy characteristics for the three different types of flexibility are given in Table III-2. PFE equal to 1440 MW represents approximately one-half the average state-wide hydropower production over the six-year time horizon (2.99 GW), and in the opinion of the authors, offers a degree of flexibility reasonable for an energy system the size of New York State's.

Table III-2: Generation/absorption and energy storage capacities for different flexibility types for 1440 MW PFE.

Flexibility Type	Power Capacity [GW]	Energy Capacity [GWh]
Flexible Supply	3.2	76.8
Pumped Hydro Energy Storage	3.2	76.8
EV Flexibility	8.64	34.56

4. Results

4.1 Flexible Hydropower Model

Results from the FHM indicate that flexible hydropower allows for greater utilization of wind-generated electricity at deep penetrations; additional PHES capacity further increases utilization. Figure III-2 presents the statewide energy mix with wind capacities of 10 GW and 30 GW, each for the following flexibility scenarios: no flexibility (“No Flex”), hydropower supply flexibility (“Flex Hydro”), and hydropower supply flexibility plus energy storage equivalent to three times the size of Blenheim-Gilboa (“Flex Hydro + 3PS”).

With 10 GW wind capacity, the “Flex Hydro" and “Flex Hydro + 3PS" scenarios have minimal effect on the statewide energy mix (<0.5%); because curtailment is low at this wind capacity, flexibility measures have little impact. However, with 30 GW wind capacity, the simulated flexibility allows for a higher degree of wind power utilization. At this capacity, the

“Flex Hydro + 3PS" scenario increases wind utilization by 1.2 GW (12.7% of potential wind generation and 6.4% of average load) compared to the scenario without flexibility.

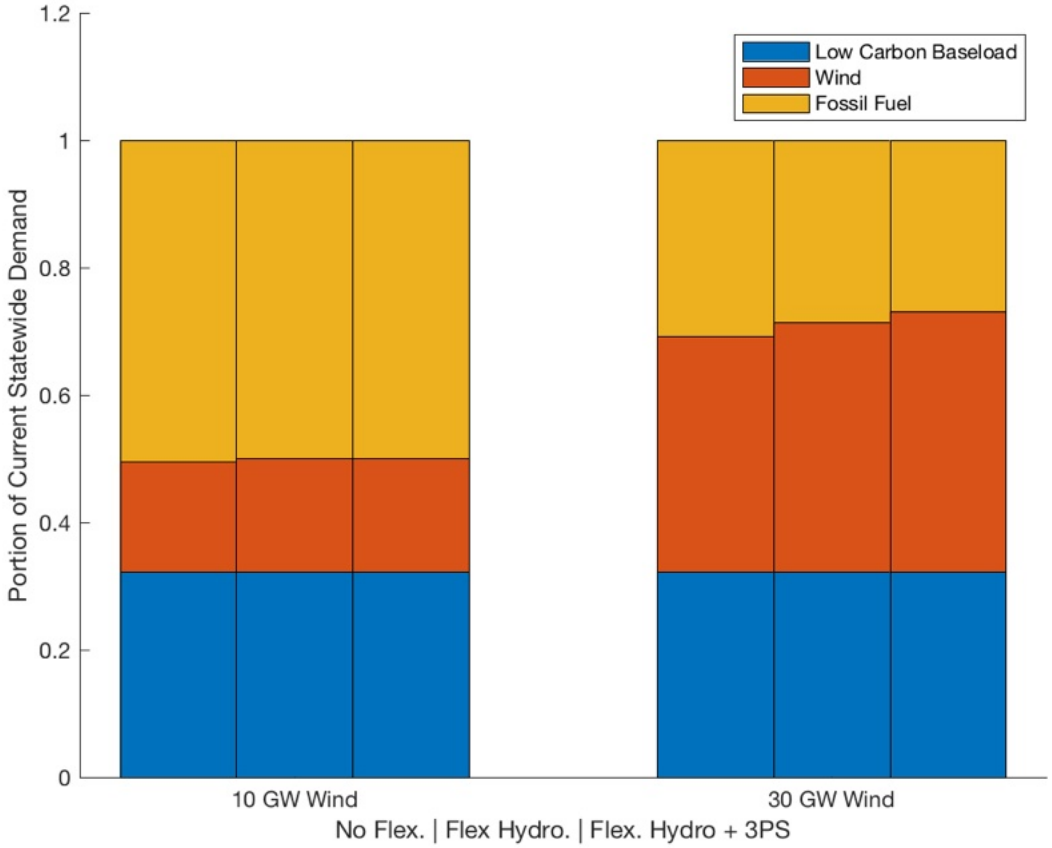

Figure III-2: FHM-computed statewide energy mix for flexibility type for 10 GW and 30 GW wind scenarios; each grouping of columns corresponds to implementations of flexibility for the given wind capacity.

4.2 Comparative Flexibility Model

In the CFM, we expand on the FHM by (a) investigating additional degrees of flexibility and (b) including the effects of interregional transmission limits. Five scenarios are considered:

1. No flexibility (“No Flex.”).
2. Flexible hydropower supply (“Flex. Hydro”).

3. Energy storage simulated as PHES.
4. Demand flexibility simulated as electric vehicles (“EV Demand Flex.”).
5. A comparison to an equivalent system with no transmission limits (“No Trans. Lim.”).

For each of these scenarios, the different types of flexibility are (a) simulated alone without other flexibility measures and (b) within single geographical areas corresponding to the Regions defined in Table III-1. To clearly illustrate our findings, only the 30 GW statewide wind capacity scenario is presented, and a single Potential Flexible Energy (PFE) amount of 1440 MW was used for all simulations. We retain the metric of computing the share of statewide electricity demand met by wind.

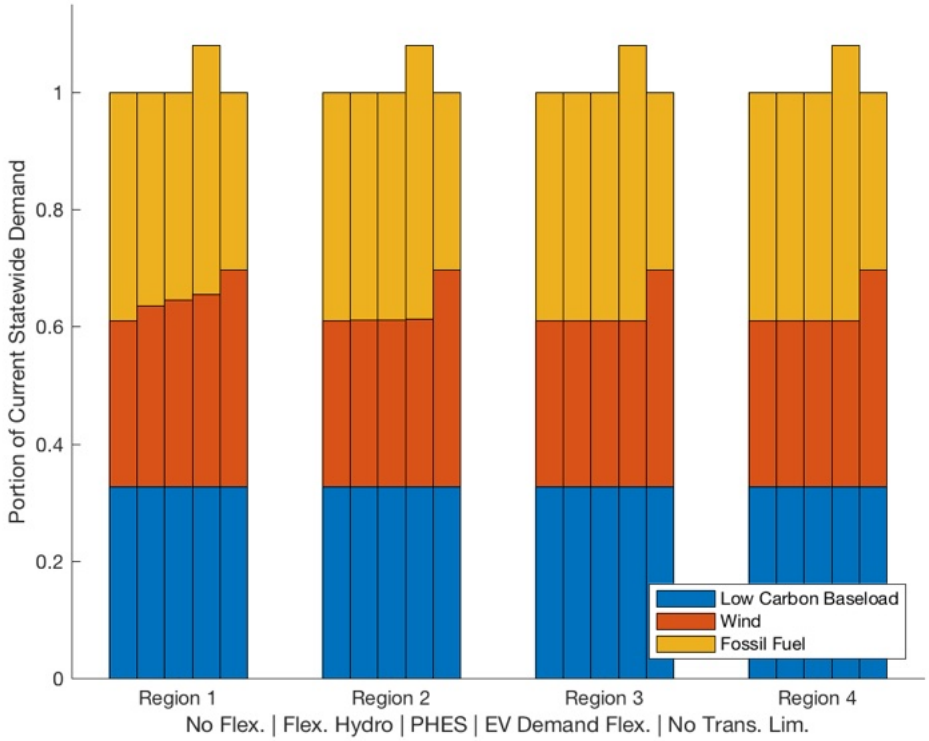

Figure III-3: CFM-computed statewide energy mix for flexibility type by region; each grouping of columns corresponds to implementations of flexibility in the noted region; 30 GW of statewide wind capacity.

Flexibility in Region 1 has the largest effect on statewide consumption of wind energy, while flexibility in Region 2 has only marginal effects, and flexibility in Regions 3 and 4 has negligible impact on the amount of wind power utilized. Figure III-3 – with columns grouped by region showing each of the five flexibility scenarios simulated in the indicated region – shows the share of statewide demand met by low carbon baseload generation (nuclear and hydropower) and wind power, with the balance assumed to be met by fossil fuel-based generation.

Without flexibility, wind-generated electricity contributes a computed average 5300 MW (28.4% of total electricity demand). Flexibility measures placed in Region 1 increase the contribution from wind by an average 470 MW (5.0% of potential wind generation and 2.5% of average load) with hydropower supply flexibility, an average 660 MW (7.0% of potential wind generation and 3.5% of average load) with PHES, and an average 840 MW (9.0% of potential wind generation and 4.5% of average load) with EV demand flexibility. It should be noted that the increase in wind-generated electricity utilization with EVs is a result of both additional demand and the flexibility of that demand: An average 720 MW increase in wind utilization is the result of the additional demand and an average 120 MW is due to its flexibility. Removing all transmission limits results in 37.0% of demand being met by wind, a result which is consistent with the FHM results in Figure III-2.

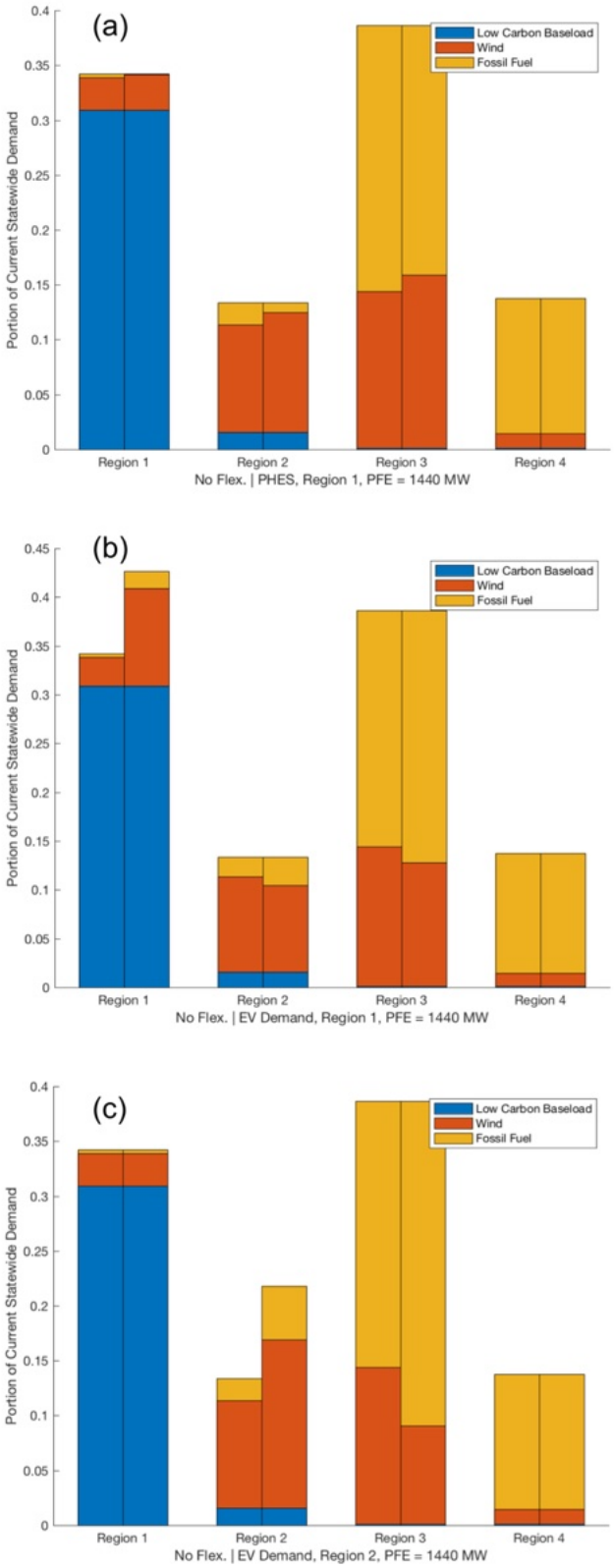

Figure III-4: Regional energy mixes for 3 different flexibility scenarios, normalized by state load. Each breakdown is juxtaposed with the energy mix breakdown from the no flexibility scenario.

As noted, flexibility in Regions 2, 3, and 4 has little effect on wind utilization. This outcome indicates that flexibility in New York's downstate regions would have little impact on reducing curtailment of wind-generated electricity – predominantly located at high-potential sites in Region 1 – due to the existing transmission limits. A wider implication beyond this specific case study is that VRE integration measures in a transmission-constrained energy system are likely to have the most impact if located near the VRE resource. By investigating a few select scenarios in detail, the overall effects described above become clearer. Figure III-4 shows the regional contribution of low carbon baseload generation, wind, and fossil fuels for three scenarios compared to the no-flexibility baseline: (a) energy storage, PHES, in Region 1; (b) demand flexibility, EVs, in Region 1; and (c) demand flexibility, EVs, in Region 2.

Outlining the results of the no-flexibility scenario defines the general topology of energy utilization in the system: Region 1, which contains 86% of the state's 30 GW wind power capacity, 91% of the state's hydropower generation, and 100% of the state's simulated nuclear generation, meets nearly its entire electricity load with low-carbon energy. Region 2, which benefits from its proximity to Region 1, meets a majority of its statewide demand from wind energy generated in the two regions. Region 3, which includes much of the New York City metropolitan area, has no low-carbon baseload energy generation within its boundaries; in the 30 GW wind power scenario, a significant portion of Region 3's load is met by wind generation from the west, but more than half of the region's demand is provided by fossil fuel generators. Region 4, at the grid “edge” and distant from wind-rich regions, has the highest portion of load met by high-carbon sources.

Figure III-4(a) compares the computed regional energy mix for the “PHES, Region 1” scenario to that of the no flexibility scenario. As PHES increases the wind energy utilized overall,

increases in demand met by wind are computed for Regions 1-3 with no effect on Region 4. This suggests that storage is effective at retaining the electricity generated by wind for discharge to other regions when transmission capacity becomes available.

Figure III-4(b) shows that with EV demand flexibility in Region 1, the increase in electricity demand (and the flexibility of that demand) allows for significantly more wind to be used to meet loads in that region. However, some decrease in wind-generated electricity used to meet loads in Regions 2 and 3 results from this electricity being used nearer the wind resource in Region 1. On balance, as is shown in Figure III-3, EVs in this region significantly increase overall wind utilization.

Figure III-4(c) shows that increased EV demand in Region 2 results in (a) an increase in wind utilization in Region 2, and (b) a similarly sized reduction in wind utilization in Region 3. Therefore, increased wind utilization in Region 2 displaces wind utilization in Region 3. The aggregate effect is a negligible increase in overall wind utilization shown in Figure III-3, above.

The inference related to the transmission effects on the results shown in Figure III-4(a) is further supported by computing the loading of interregional transmission interfaces. Figure III-5 shows the amount of electricity transmitted at the interregional interfaces with and without PHES in Region 1. With an uptick in wind-generated electricity utilization in Regions 2 and 3, increased utilization of transmission capacity at Region 1-2 and Region 2-3 interfaces are also computed; the Region 3-4 interface displays no such effect. Without flexibility, the Region 1-2 transmission line is full more than half the simulation time period. Some increase in times of full loading is computed with PHES; however, most of the increased energy flow (entirely constituted of

additional wind-generated electricity) occurs at times of lower transmission loading, supporting our earlier conclusion.

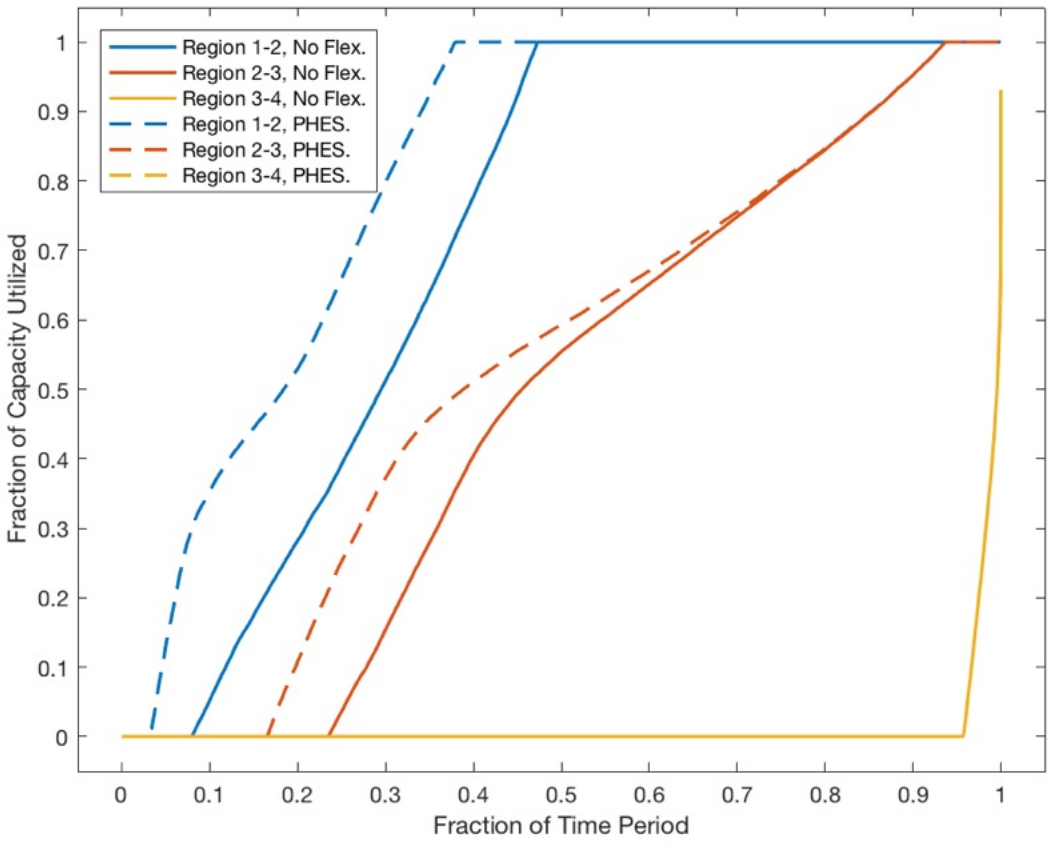

Figure III-5: Normalized transmission loading duration curve vs. time; PHES in Region 1, PFE = 1440 MW.

5. Conclusion

Chapter III presents two models, both of which seek to quantify the benefits of power system flexibility to New York State under different considerations. The Flexible Hydropower Model (FHM) investigates different levels of hydropower flexibility compared to a fixed hydropower baseline to simulate the state's current load-following hydropower capability in the context of deep VRE penetration. Simulated flexibility offers system benefits only at installed

wind capacities above 10 GW; at 30 GW wind capacity, flexible hydropower with additional PHES increases average wind utilization by 1.2 GW (12.7% of wind potential and 6.4% of average load).

The Comparative Flexibility Model (CFM) explores different types of system flexibility. Results show that energy storage (in the form of PHES) and demand-side flexibility (in the form of electric vehicles) have the largest effect on integrating substantial capacities of renewable generation when located near the VRE resource. The adoption of electric vehicles in Region 1 with an average load of 1.44 GW and daily available battery capacity of 34.5 GWh increases average statewide wind utilization by 840 MW (9.0% of potential wind generation and 4.5% of average load) in a 30 GW wind capacity scenario; PHES in Region 1 with 3.2 GW charge/discharge capability and 76.8 GWh storage capacity expands statewide wind utilization by an average of 660 MW (7.0% of wind potential and 3.5% of average load).

A central finding of this study is that in the transmission-constrained system simulated, flexibility only increases wind-generated electricity utilization when located near the wind resource, which is primarily in Region 1. All types of flexibility in Regions 2-4 have negligible impact on increasing wind-generated electricity utilization. Therefore, energy systems planners should prioritize siting large-scale flexibility measures near renewable generation.

Further research is needed to generalize the results of this analysis for other VRE resources, namely solar photovoltaics and offshore wind power. Offshore wind power is often nearer transmission-constrained load centers that are distant from onshore wind resources. Tradeoffs between these effects and the costs and opportunities for expanded transmission capacity also require investigation.

Appendix A: NYS Flexible Hydropower Model

Additional Model Constraints

All equations hold $\forall t$ unless otherwise mentioned.

Generation at each hydroelectric facility is limited to the max generation capacity at the facility. Reservoir level is similarly limited to reservoir capacity at every facility. All variables are greater than zero. For all PHES facilities, pumping and generation do not both occur in a single time step. In all equations, γ refers to pumped storage charging, δ refers to hydropower generation discharge, and E refers to aggregate storage state of charge; all variables are indexed to the hydroelectric facility in question.

$$\frac{H_{nia}^{t}}{\eta_{hydro}} - \frac{\delta_{lew}^{t}}{\eta_{hydro}} + \eta_{hydro} * \gamma_{lew}^{t} = E_{nia}^{t-1} - E_{nia}^{t} + H_{in,nia}^{t} \quad \text{(III-10)}$$

$$\frac{\delta_{lew}^{t}}{\eta_{hydro}} - \eta_{hydro} * \gamma_{lew}^{t} = E_{lew}^{t-1} - E_{lew}^{t} \quad \text{(III-11)}$$

$$\frac{\delta_{bg}^{t}}{\eta_{hydro}} - \eta_{hydro} * \gamma_{bg}^{t} = E_{bg}^{t-1} - E_{bg}^{t} \quad \text{(III-12)}$$

$$\frac{\delta_{nc}^{t}}{\eta_{hydro}} - \eta_{hydro} * \gamma_{nc}^{t} = E_{nc}^{t-1} - E_{nc}^{t} \quad \text{(III-13)}$$

$$\frac{\delta_{sh}^{t}}{\eta_{hydro}} = E_{sh}^{t-1} - E_{sh}^{t} + H_{in,sh}^{t} \quad \text{(III-14)}$$

Appendix B: NYS Comparative Flexibility Model

Additional Model Constraints

All equations hold ∀*t,* ∀*r* unless otherwise mentioned.

Generation at each hydroelectric facility is limited to the max generation capacity at the facility. Reservoir level is similarly limited to reservoir capacity at every facility. All variables are greater than zero. For the PHES facility, pumping and generation do not both occur in a single time step. In all equations, γ refers to pumped storage charging, δ refers to hydropower generation discharge, and E refers to aggregate storage state of charge; all variables are indexed to the hydroelectric facility in question.

$$\frac{\delta_{ps,i}^{t}}{\eta_h} - \eta_h * \gamma_{ps,i}^{t} = E_{ps,i}^{t-1} - E_{ps,i}^{t} \tag{III-15}$$

$$\frac{H_{flex,i}^{t}}{\eta_h} = E_{fh,i}^{t-1} - E_{fh,i}^{t} + H_{in,i}^{t} \tag{III-16}$$

$$H_{in,i}^{t} + H_{fix,i}^{t} = H_{tot,i}^{t} \tag{III-17}$$

$$\sum_{t=19+24m}^{7+24(m+1)} \eta_{veh} * D_{flex,i}^{t} = 24 * D_{veh,i}, \quad for\ m = 1..\frac{T}{24} \tag{III-18}$$

$$Z_{ii'}^{t} \leq X_{tx,ii'}^{existing} \tag{III-19}$$

Part II: Irrigation and Energy System Modeling in Ethiopia

Introduction

The second half of this dissertation focuses on detecting smallholder irrigation presence, ultimately connecting predicted irrigated area back to its energy system planning implications. The geographical context of this effort is Ethiopia, a country of 110m where current use of irrigation is very low – approximately 2.1% of all agricultural land [175] – and 51% of the population has access to electricity [176].

While electrification of households is nearly complete in Asia, approximately 700 million people in sub-Saharan Africa still live without residential access to electricity [177]. Addressing household level access requires planners to know where households are located; fortunately, given advances in satellite imagery-based techniques of locating buildings and improved administrative capacity in many countries, this challenge has been resolved in the last 5 years. Previous work from the Sustainable Engineering Lab (SEL) has estimated moderate levels of residential electricity consumption in sub-Saharan Africa to be between 20 and 50 kWh/month [178]. At prices approximately equal to $0.10/kWh, this level of consumption returns between $24 and $60 per year in electricity payments. Given average per-connection costs of $1000, the economics of the new consumers means that expanding electricity access is often a loss-making proposition for utilities [179].

In comparison to residential demand, electricity demand for productive uses – i.e., for agricultural, commercial, or industrial activities – can be much larger in scale. For example, given a one-hectare farm plot, extracting 7mm of water per day from a groundwater reservoir 25m

below the surface results in a monthly electricity demand of 286 kWh for irrigation, assuming 50% pump efficiency. Given their scale, such levels of demand must be affordable on a per-unit basis, or else the cost requirements will subsume all benefits of the irrigation itself. In Asia, grid extensions provided this affordable power. In Africa, with emerging levels of irrigation and low densities of demand, energy is often provided manually or via liquid fuels – both expensive options.

Moreover, unlike residential electricity demands, demand for irrigation has the potential for flexibility. Most crops need approximately 35mm of water in a 5-day period to maintain healthy growth; these water requirements can be supplied at any time during the 5-day period [180]. Accordingly, electricity demands for irrigation can be met when energy is particularly cheap, such as during the middle of a sunny day for a system with solar generation capacity. This flexibility lowers overall costs, as batteries or dispatchable generation are not needed to meet immediate power needs, as is the case with residential loads.

Many researchers have demonstrated the economic benefits of productive electricity demand. In Ethiopia, one study found that the average income of irrigating households was double that of non-irrigating households [181]. Revenues from secondary crops during non-rainy portions of the year provide supplementary sources of income that stimulate local economies and fully pay off the costs of a grid connection or minigrid installation, in turn generating capital for further electrification efforts [182]. Based on numbers gleaned from previous SEL experience, the incorporation of productive electricity demand can reduce the time required to pay off an investment in a grid connection from 20 to 8.5 years.

National planners recognize that providing electricity services to farmers who are currently using manual or diesel power to satisfy energy demands for irrigation first requires knowing where these farmers are. In high income countries where farms are often on the scale of tens to hundreds of hectares, this is not an issue; however, in sub-Saharan Africa, farmers are likely to be irrigating plots between a fraction of a hectare and a couple hectares in size. Locating farmers who are currently irrigating also provides information about nearby areas where irrigation can expand: If irrigation is possible on one plot, there are no barriers to expansion – e.g., excessively deep water tables, lack of local know-how, no market access – on a neighboring plot.

With a basis established for the importance of locating and promoting smallholder irrigation, Part II of this dissertation introduces methods for irrigation detection; these methods are followed by an exploration of the impacts of productive electricity demands on energy system planning. To build upon existing efforts in the agricultural monitoring space, the literature on land cover mapping and irrigation detection is first surveyed, a discussion that leads to the development of an irrigation detection methodology for the Ethiopian Highlands. After demonstrating the applicability and accuracy of this methodology for Tigray and Amhara, two Ethiopian states in the Northwest portion of the country, extensions of this approach are explored. These extensions include investigating the spatial limits of trained irrigation detection models; a method of filtering out false positive predictions; a longitudinal assessment of irrigation growth within Tigray; and an application of the irrigation detection model in a separate region of Ethiopia that demonstrates a similar vegetation characterization to that of the Ethiopian Highlands. After a thorough exploration of the irrigation detection model in Ethiopia, predicted

irrigated areas are used in the development of a new optimization model that finds the lowest cost method of providing electricity for various systems configurations with simulated productive electricity demand. Accordingly, a common through-line exists throughout Chapters IV-VII, with the following paragraphs giving a more complete, chapter-by-chapter summary of the second part of this dissertation.

For the task of irrigation detection in sub-Saharan Africa, Chapter IV summarizes relevant literature on land cover and irrigation mapping. Here, a 2012 paper by Professor Christopher Small is particularly germane, as it introduces a method of characterizing dominant regional phenologies in an area of interest, a characterization which provides the spatial boundaries for developed irrigation detection models. This method of spatiotemporal vegetation characterization is then implemented for a small subset of Amhara to visualize areas in the state that contain vegetation cycles in-phase and out-of-phase with rainfall. A separate method of characterizing the alignment of rainfall and vegetation cycles is also presented, whereby MODIS imagery and Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) precipitation estimates are used to identify areas where vegetation growth cannot be attributed to rainfall. Here, locations with rainfall-vegetation misalignment offer insight into where irrigation is occurring in Ethiopia, as the water necessary to sustain crop growth must be provided via some other mechanism; these locations also provide a point of comparison for subsequent predictions from irrigation detection models.

Chapter V contains a novel, multiscale spatiotemporal approach for smallholder irrigation detection. A fundamental problem with irrigation detection in sub-Saharan Africa generally and Ethiopia specifically is a lack of labeled data – there are no publicly available datasets that

demarcate irrigated areas. To remedy this gap, Chapter V details a process to supplement limited ground-collected labels and ensure classifier applicability in an area of interest. Adapting the spatiotemporal vegetation characterization methodology introduced by Professor Small, the first portion of this chapter presents a continuous phenology map that guides supplementary label collection and irrigation prediction model applicability. After acquisition of labeled data from seven distinct regions in the Ethiopian Highlands, timeseries of Sentinel-2 imagery are collected and cleaned to provide features for classifier training. Here, the efficacy of different training strategies and model architectures are assessed. Results indicate that enhanced vegetation index (EVI) timeseries – created using the Sentinel-2 blue, red, and near-infrared spectral bands – randomly shifted by up to ±30 days yield the highest prediction accuracies. After training over visually collected samples, the irrigation detection model is evaluated over reserved ground-collected labels, achieving 96.7% accuracy over non-irrigated samples and 95.9% accuracy over irrigated samples. Similar performance is realized over a series of samples independently collected via the introduced method of label supplementation. Lastly, the prediction model is deployed over the entirety of Tigray and Amhara, finding that irrigated area in these two states has decreased by 40% from 2020 to 2021, likely due to ongoing civil conflict in the area.

Chapter VI extends the methodology contained in Chapter V to explore other irrigation prediction dynamics in Ethiopia. First, the size distribution of predicted irrigated areas in Tigray and Amhara is explored, finding significant similarities between irrigated segments in the two regions except for the largest 1% of plots, which are substantially larger in Amhara. Next, misclassified predictions along the Amhara/Afar border are investigated, revealing that the 2020 dry season contained precipitation and vegetation cycles anomalous within the 10 years of data

inspected. To control for irrigation misclassification, predictions are then filtered by an auxiliary 10m land cover map. While this filtering method reduces the number of false positive predictions, it also likely eliminates many true positive predictions, a tradeoff that must be considered in producing final predicted maps. Next, Chapter VI presents an application of the irrigation detection model to an area in Oromia with similar climatological and phenological patterns to those in Tigray and Amhara, finding that irrigated areas surrounding a reservoir near Asasa can be identified. Moreover, longitudinal studies of irrigation in Tigray show that the plot-level location of irrigated areas changes substantially from year to year, and that the 40% decline in irrigated area from 2020 to 2021 identified in Chapter V is not due to excessive positive predictions in 2020, as irrigated area in 2020 represents a 6% increase over predicted area in 2019. Another extension in Chapter VI assesses prediction performance over a set of labels collected in Tigray in 2020, demonstrating that the irrigation prediction model is robust in years and regions for which no training data was acquired. Lastly, the addendum to Chapter VI contains details about a fully open-source implementation of all methodology introduced in Chapter V using Google Earth Engine.

Chapter VII further details the theoretical underpinnings of the importance of productive electricity demand for energy system planning. While the shape and nature of all types of electricity demand are relevant during planning, the substantial potential for growth of electricity for irrigation, along with the size and flexibility of these loads, make them essential considerations for electrification efforts. This chapter then incorporates predicted irrigated areas in energy system planning via the introduced Irrigated System Electrification (ISE) model. Through its formulation as a linear program, ISE computes the least-cost method of satisfying

combined residential and productive electricity. ISE is parametrized from previous SEL experience and applied to a case study near Gebedge, Ethiopia. Model simulations reveal that grid power offers the cheapest method of meeting residential and productive demand at a cost of $0.164/kWh. In contrast, isolated solar/battery/diesel systems offer more costly alternatives when a grid network is not present: If all 0.2 hectare plots within a 300m radius are connected to the same system, the cost of electricity is $0.272/kWh; however, if all 0.2 hectare plots are independent, the electricity cost rises to $0.397/kWh.

Chapter IV: Review of previous approaches for irrigation detection and vegetation mapping

Abstract

To estimate productive electricity demand for irrigation, it is first necessary to locate areas that support dry season vegetation growth, as these sites are likely to contain additional cropping seasons with corresponding energy demands. Identifying these areas will require pulling from existing literature on vegetation characterization and land use monitoring. Of particular interest are methods developed in [183], whereby a temporal mixture model is applied to timeseries of vegetation abundance layers to create a regional phenology map. Accurate characterization of vegetation phenologies in an area of interest informs the extent to which an irrigation detection model will prove applicable. The following sections provide background and a literature review for 1) vegetation characterization via satellite imagery; and 2) irrigation detection efforts, both globally and specifically in Ethiopia.

1. Literature Review

There are a number of existing publications particularly germane to our task of irrigation detection in Ethiopia. At a high-level, detecting irrigation extent and land cover change via satellite imagery is a well-established methodological approach, consisting of many previous efforts that are detailed here in Chapter IV. Before reviewing this literature, we foreground the five novelties that separate the proposed method of irrigation detection contained in Chapter V from previously published work and qualify as original contributions to the field.

1. A method of label collection that allows 1) labeling at scaling, and 2) verification of *timeseries* of EVI signatures for new samples. This method includes a process called "cluster cleaning", which allows users to draw larger labeled polygons – i.e., label faster – and confirm that newly labeled samples demonstrate phenologies consistent with our definition of irrigation/non-irrigation. Other labeling processes typically use a single timestep of imagery for labeling. For the case of Ethiopia, it's easy to confuse non-irrigated, evergreen cropland with irrigated areas using only Sentinel-2 imagery at a single timestep. As our labeling process verifies annual vegetation phenologies in making an irrigation/non-irrigation class decision, it avoids this issue.
2. A method of applying a random phase shift to input timeseries during training to increase classifier robustness. Data augmentation techniques are well-established in the computer vision literature, but this particular method has never been deployed for remote sensing classification tasks.
3. Use of a spatiotemporal phenology map to govern the applicability of a classifier model. While methods for creating a phenology map are not new, using one to understand and visualize irrigation model transferability is. Given the tendency of other irrigation products to be generated over a large spatial extent with little, if any, justification for using the same model across different agro-ecological settings, this aspect of our work sets it apart.
4. A 10m irrigation map is produced for 2020 and 2021, covering 205,000 km^2 at >95% accuracy. There is no other product that exists for Ethiopia for these years at 10m scale with this level of accuracy.

5. A 40% decline in irrigated area from 2020-2021 is computed, a new finding regarding how recent irrigation patterns have changed in Ethiopia, one that is still relevant given the ongoing civil conflict in the country.

The following literature review is split into subsections based on the key methodological assumptions of the work in question. First, literature most closely related to this thesis work is discussed. The next subsections cover irrigation mapping via data fusion-based approaches; irrigation mapping using advanced computed vision methods; irrigation mapping across the continental US; irrigation mapping efforts at decameter resolution; global irrigation mapping at hectometer and kilometer resolution; and various relevant land cover mapping efforts. As an entry into this discussion, Ozdogan et al. present a review of opportunities and challenges in using remote sensing for irrigated agriculture monitoring [184]. Specific findings from this paper posit that single date imagery acquired during the peak of the growing season can hold sufficient information for irrigation classification, but multitemporal imagery is preferred. The authors also find that at larger scales, more work is required to identify the best spectral indices, best observed time periods, and best classification methods under different climatological and agroecological environments. Lastly, the paper concludes that further model refinement is necessary as additional labeled data is folded into the training process.

1.1 Literature most closely related to the proposed irrigation detection methodology

The two papers most relevant to our irrigation detection work were produced by Vogels et al. [185,186]. The first paper has two goals: 1) evaluate Geographic Object-Based Image Analysis (GEOBIA), an object based approach for field segmentation, for irrigation mapping; and

2) discriminate between large-scale and smallholder agriculture in the Ethiopian Rift [185]. For this task, three 1.5m resolution SPOT6 scenes from November 2013, December 2013, and February 2014 are used as imagery. An image segmentation software (eCognition*) is applied on NDVI difference maps to generate 3000 labeled polygons; these polygons are then labeled via visual inspection as irrigated, non-irrigated, or non-cropped. Here, the presence of irrigation is determined based on vegetation content in the segmented objects. As the three SPOT6 scenes are all collected during the dry season of this part of Ethiopia, visual confirmation of vegetation is assumed to be indicative of irrigation presence. Of the 3000 labeled objects, 596 were labeled as traditional smallholder and 549 were labeled as modern large-scale agriculture. Lastly, labeled objects are filtered using ancillary land-use/land-cover (LULC) maps and cloud cover masks to result in a final dataset of 2636 objects.

For each object, 57 variables are derived in total: 17 spectral variables, 8 shape variables, 22 texture variables, and 2 location variables (x and y coordinates). These variables are used as inputs to a series of random forest classifiers. The first model classifies each object as cropland or non-cropland. The second model classifies the cropland predictions from the first model as irrigated or non-irrigated for each of the 3 scenes. The third and final model classifies smallholder and large-scale agriculture.

Overall, this approach achieves 94% accuracy in predicting irrigation presence. While this performance is impressive, there are a number of limitations to the methodology. First, no ground truth is leveraged in this effort, making the visually labeled cropland objects the only data points for assessing performance. Second, this approach is limited to where meter resolution imagery is available; as most of this imagery comes from commercial sources, it is not available

to researchers over a wide extent or time period. Third, since model inputs consist of a single red-green-blue-near infrared (RGBN) image, no temporal information is incorporated. Irrigation is a land cover process that involves greening and senescence periods that extend over multiple months, and timeseries of imagery provide insight into irrigation presence that is impossible to acquire from single timestep snapshots.

The second Vogels et al. paper [186] extends the methodology of [185], using GEOBIA and Sentinel-2 imagery for differentiating irrigated and non-irrigated agriculture across the entirety of the Horn of Africa: Ethiopia, Eritrea, Djibouti, Somalia, and Kenya. This methdology produces monthly maps of irrigated agriculture between September 2016 and August 2017. First, croplands are segmented using a dry-season (October 2016 – March 2017) Sentinel-2 NDVI mosaic and the eCognition Developer. Statistical quantities for these cropland objects – e.g. mean NDVI, mean NDVI in February, texture, size – were used as features in a random forest classfier to distinguish cropland from non-cropland. Ultimately, 9632 objects were labeled as cropland and 34,695 were labeled as other LULC classes, with the resultant cropland classifier 1) reporting 96% accuracy, and 2) determining that 17% of the total land area in the Horn is cropland.

After determining cropland extent, irrigation presence was predicted using process-based rules on field object NDVI timeseries. Here, NDVI timeseries are created using monthly Sentinel-2 mosaics between September 2016 and August 2017. For each object, the NDVI differences on a monthly basis are calculated to determine whether crop growth is happening (i.e., if the NDVI different was positive). Then, the portion of vegetated objects (i.e. not roads, water bodies, or villages) within a 5km radius that also demonstrate growth is determined. If this portion is less

than a specified threshold – either 15%, 25%, or 35% – the object is determined to be irrigated. In contrast, if this portion is larger than the threshold, the object is determined to be rainfed. This method produces 3 times (for the 3 different thresholds) 11 monthly land use maps showing the spatial distribution of rainfed and irrigated agriculture. By methodological design, the authors of [186] cannot determine an optimal threshold between 15%, 25%, and 35%. These thresholds correspond to a substantial difference in overall irrigated area: 17.7 Mha, 20.75 Mha, and 23.72 MHa, respectively. These areas can be compared to existing estimates of 8.17 Mha (IAAA), 0.0004 Mha (Globcover2009), 9.93 MHa (GFSAD1000: GIAM), 0.95 Mha (GMIA), and 1.23 Mha (AQUASTAT). As this approach does not use verified labels of irrigation presence, it is entirely dependent on selecting an accurate threshold in producing a final irrigated areas map.

While this paper remains a valuable contribution to the literature, there are numerous methodological shortcomings to its approach. First, no labeled data is collected to assess the accuracy of irrigation predictions, preventing readers from knowing how well the process-based approach performs for the task of irrigation detection. Second, the methodology treats each month's Sentinel-2 imagery as independent from preceding/following images, removing all temporal information for irrigation process detection. This choice leads to the counterintuitive result that many pixels in the Horn are predicted as irrigated more than 6 times, an impossibility given the multi-month periodicity of irrigation cycles and the existence of a rainy season (or multiple rainy seasons, depending on the location) that renders supplementary crop watering events unnecessary. Third, the same approach to irrigation detection is deployed over the entire Horn of Africa regardless of substantial regional differences in climatology and agroecological conditions.

A recent working paper from IWMI maps irrigated and rainfed agriculture in Ethiopia from 2015-2016 using remote sensing methods [187]. This paper uses a single mosaic of Landsat 8 images to remove non-agricultural land cover types. Subsequently, MODIS 16-day NDVI composites and CHIRPS data are used to assess irrigation presence. Here, classification is carried out by combining temporal correlation of crop growth patterns with rainfall, the number of crop cycles in a year, and dry season moisture status. Timeseries of MODIS imagery are paired with ground truth from field surveys, secondary sources, and visual interpretation of high resolution images; unfortunately, no further information about the amount or extent of labeled data used in this paper is specified. Despite being completed with 250m MODIS imagery, the final map is produced at 30m resolution. Given the native 250m resolution of this map, it cannot capture smallholder irrigation dynamics.

Methods of mapping cropping intensity of smallholder farms are compared by Jain et al. in [188]. Four methods of intensity mapping are assessed: 1) the Landsat threshold method, which identifies whether a Landsat pixel is cropped or uncropped in each growing season based on an NDVI threshold; 2) the MODIS peak method, which determines whether there is a peak during MODIS timeseries; 3) the MODIS temporal mixture analysis method, which quantifies MODIS sub-pixel phenology heterogeneity using methods from [183]; and 4) the MODIS hierarchical training method, which computes the sub-pixel heterogeneity of cropping intensity using hierarchical training techniques. All methods are assessed using four criteria: data availability, accuracy across different spatial scales, ease of implementation, and ability to use the methods over large spatial and temporal scales. This study was performed over Madhya Pradesh and Gujarat states in India over summer and winter seasons. In Gujarat, Quickbird and

Worldview-2 imagery are used for label collection and validation; in Madhya Pradesh, no high-resolution imagery is available, so Landsat 8 is used for label collection. For each state, 25 cropped areas of interest – i.e., polygons – and 25 non-cropped areas of interest, each approximately 100m x 100m, were collected for model validation via visual interpretation. Predictions are made for "fraction of each cell cropped" at 250m, 1km, 10km spatial resolutions; performance is reported as R2 values. Predictions are also validated at the district level using Indian census data. Altogether, this paper finds that the Landsat method is most accurate, particularly at smaller scales of analysis (i.e., 250m resolution); however, this method is more difficult to implement because of a lack of image availability. When aggregating to higher spatial scales (i.e., 1km and 10km resolutions), MODIS-based approaches have similar performance to the Landsat-based approach.

A 2012 paper by Professor Small introduces a method of combining characterizations of spatial and temporal change with modeling, whereby combinations of endmembers in a temporal feature space are used to represent the relative presence of spatiotemporal processes [183]. Here, the topology of the feature space informs the hand-selection of temporal endmembers; this approach is applied to map the vegetation phenologies of the Ganges-Brahmaputra delta using MODIS vegetation index timeseries. The methodology introduced in this paper provides the foundation for the temporal endmember extraction technique and the utilization of abundance maps for understanding landscape change presented in the following sections and Chapter V.

1.2 Irrigation mapping via data fusion

As remotely-sensed data become more accessible, researchers are increasingly exploring the benefits of data fusion – the process of combining multiple satellite imagery products, often across sensing domains, to extract otherwise hidden information about irrigation processes. Ferrant et al. leverage Sentinel-1 and Sentinel-2 to 1) detect irrigated crops, and 2) estimate seasonal groundwater use in South India [189]. Here, a random forest is deployed over three growing seasons (January to March 2016, July to November 2016, January to March 2017) to classify: inundated rice paddy, irrigated crops, and dynamics of other surface water areas. NDVI and NDWI measurements are used from the Sentinel-2 imagery, while the *vv*/*vh* ratio of Sentinel-1 backscatter is extracted to produce ground observations in the cloudy monsoon season. Three seasonal surveys of land cover provide 428 plots of non-irrigated area, 286 irrigated areas, 192 flooded rice crop areas for training and validation labels. These plots encompass 9 land-cover classes – inundated rice, irrigated vegetables, irrigated maize, orchards, forested area, bare ground and natural bushes, urban areas, surface water, and rain-fed cotton – and result in 100,000 labeled pixels split across the 3 region-growing seasons. A model is trained independently for each region-growing season.

MODIS timeseries, Landsat imagery, and ancillary data are combined to detect irrigation extent, frequency and timing in northwestern China in [190]. Here, imagery timeseries are created at 30m by fusing Landsat and MODIS data: MODIS imagery is transformed to feature space, and then the first dimension of the PC transform is replaced with Landsat imagery, followed by application of the inverse transform. These images are paired with extensive ground truth: daily irrigation records of 36 villages in the experiment area for 2014, 2015; and daily

irrigation records at site scale for 2012, 2015, 2016 from Daman via HPSDC (Heihe Plan Science Data Center). Using the proposed method, the authors are able to identify water supplement stages with 90% accuracy.

A series of other papers deploy data fusion approaches for irrigation detection, combining sensed imagery from multiple different satellite missions. In [191], the authors distinguish between irrigated and rainfed crops in southwestern France using timeseries of Sentinel-1 imagery, Sentinel-2 imagery, and meteorological variables. For classification, a labeled dataset of 832 plots in 2017 (557 irrigated and 275 non-irrigated) and 942 plots in 2018 (680 irrigated and 262 rainfed) are used, collected via a field campaign. Input timeseries are processed to a monthly time resolution, and then are inputted to a random forest model; no information is provided about what features are most important for classification.

A similar effort proposes an operational framework for mapping irrigated areas at plot scale using Sentinel-1 and Sentinel-2 data, which is then applied to a study site in northwestern France [192]. The key contribution of this paper is a method for collecting labeled data without in-situ observations. Here, labeled data is collected based on either 1) the output of an irrigation event detection model applied to Sentinel-1 timeseries, or 2) maximum Sentinel-2 NDVI during the crop cycle compared to a reference threshold (>=0.8 is irrigated; <= 0.7 is non-irrigated). The irrigation event detection model introduced in this paper compares Sentinel-1 backscatter measurements at the plot level to the change in backscatter measurements at the grid scale (10km x 10km): If the backscatter at plot level increases while backscatter at grid level remains constant, then the authors assess a high probability of an irrigation event. Using these labeled

data, a random forest classifier is trained to distinguish irrigated and non-irrigated plots, achieving between 73-93% accuracy across the years and areas of deployment.

Sentinel-1 radar and Sentinel-2 optical imagery are combined to map irrigation in Catalonia and Italy in [193]. Labels are derived from administrative data in each country – SIGPAC in Catalonia [194], and iColt in Italy; 2000 labels are used in total. Here, each labeled vector plot reduces to a single collection of features via averaging of all timeseries within that plot. Feature layers also include a ratio of plot-level statistics to statistics for the surrounding 5km-by-5km area to determine whether plot-level attributes are uncharacteristic of the surrounding landscape. A Support Vector Machine (SVM) is then trained to separate irrigated and non-irrigated plots. The authors produce a high accuracy – 85% – for the binary irrigation task, which they attribute to the benefits of data fusion. Among all feature inputs, the *vv* layer was found to be most predictive of irrigation presence.

Bolognesi et al. fuse two different types of optical imagery in [195], combining Landsat 8 and Sentinel-2 timeseries to detect irrigation in southern Italy. Here, 57 Landsat 8 images and 145 Sentinel-2 images are combined to produce a single imagery timeseries with 30m resolution and a timestep of 2-3 days. The ground truth for this effort comes from a field survey, with 2992 data points from 3 main classes: bare soil and rainfed, herbaceous irrigation, and irrigated tree crop. Data points are taken one per field parcel in order to limit the spatial correlation among the training data. After organizing the imagery and labeled samples, the authors assess various pixelwise classifiers, achieving accuracies close to 90%.

Liu et al. also fuse Landsat 8 and Sentinel-2 timeseries in [196], using the two sources of imagery to map cropland intensity in northwestern China. The two types of optical imagery are

combined to create 30m, 10-day composites, with temporal gaps filled by a Savitzky-Golay filter. A field collection effort produces 700 pixelwise labeled samples for modeling; signal processing is then applied to these samples to determine whether single, double, or triple cropping was occurring, with max NDVI values above 0.5 required once during the crop cycle to qualify as cropped. This heuristic-based approach – one without any machine learning – achieves an accuracy of 94%. The authors also find that only NDVI and the Land Surface Water Index (LSWI) are necessary for classification.

1.3 Irrigation mapping using advanced computer vision techniques

Various advanced deep learning methods developed for computer science applications have also been transferred to remote sensing applications to achieve state-of-the-art performance for classification tasks. One such effort uses a self-supervised contrastive loss to learn important features in the BigEarthNet-S2 imagery dataset before tuning to an irrigation detection task [197]. In this self-supervised pretraining stage, a batch of images is first selected. For each image in this batch, two image augmentations are selected at random, with a neural network learning to minimize the contrastive difference between these images after a forward pass. Assessed over withheld samples in Europe from the BigEarthNet-S2 dataset, this pretrained model consistently outperforms the supervised baseline for the task of irrigation detection, indicating that the proposed approach produces models that generalize well on unseen data.

Ragettli et al. develop an unsupervised classification algorithm for multi-temporal irrigated area mapping in Central Asia in [198]. This process is developed in Google Earth Engine, where regions of interest are first identified through unsupervised object-based segmentation, followed by multi-temporal image analysis to distinguish productive irrigated fields from non-

productive and non-irrigated ones. Unsupervised segmentation is achieved via a region growing clustering technique; as this step produces polygons covering large irrigation schemes with relatively homogeneous land cover, subsequent supervised classification models are trained to distinguish between vegetated and non-vegetated land use classes. Classification is performed using 1) multitemporal Landsat 7, Landsat 8, Sentinel-2, and MODIS imagery; 2) 2M labels collected via single-timestep photo interpretation and hand labeling of 7 available Landsat scenes, resampled to 80m; and 3) a random forest classifier. This approach produces monthly 30m irrigation maps for Kyrgyzstan and Kazakhstan between 2000 and 2017, achieving accuracies between 77-96% based on 7 available validation scenes.

The literature on irrigation detection also contains transfer learning based approaches as well. In [199], a U-Net with a ResNet50 backbone (trained on ImageNet) yields the best detection performance. This approach is applied in Morocco on three cloud-free Landsat 8 scenes (October 2015, May 2016, and August 2016); these three scenes produce 50 224 x 224 image patches in total, with labels collected via photointerpretation of the data. Only RGB bands are used for classification. The authors of this paper also find that two methods of data augmentation – rotating and flipping input images – improve prediction accuracy.

Given the rapid introduction of many novel deep learning model architectures, recent research has focused on assessing the performance of these architectures for satellite imagery-based applications. In [200], Zhao et al. evaluates the performance of 5 different deep learning models for crop type classification in Hebei Province, China. Classification is performed using Sentinel-2 timeseries, with missing pixel values infilled using a Savitzky-Golay filter. A mixture of ground truth and visually collected samples from Google Earth Engine comprise this paper's

labeled data; in total, 2182 pixels are collected across 7 crop classes. The authors find that models that preserve temporal relationships – such as the LSTM and GRU – perform better than those that do not. Another finding is that 1D convolutional networks can learn from timeseries with missing data (i.e., those that haven't been infilled) – the highest performing models achieve accuracies around 85% when the total missing rate of samples is 43.5%.

1.4 Irrigation mapping across the continental United States

Transitioning to irrigation mapping efforts in the US – where croplands are larger than the typical smallholder scale seen across most of sub-Saharan Africa – Colligan et al. produce a deep learning-based approach in [201] for mapping annual irrigation in Montana from 2000-2019. Input imagery consists of imagery stack timeseries composed of 6 temporally averaged images (32 days apart) starting May 1, all containing blue, green, red, NIR, SWIR1, and SWIR2 spectral bands. Labeled data is provided by a novel ground truth dataset, county level stats from USDA on irrigation extent, and cadastral surveys. The novel ground truth dataset was introduced in [202] (described below), and consists of vector labels partitioned into three classes: irrigated, unirrigated, and uncultivated. These data are collected via manual interpretation of satellite images, and then are vetted with high-resolution imagery and rainfall timeseries. For the task of mapping irrigation in Montana, a standard U-Net is deployed. This network architecture learns to mask clouds and ignore Landsat-7 scan line failures without supervision, reducing the need for preprocessed data.

Ketchum et al. produce annual 30m resolution irrigation maps using Google Earth Engine for the years 1986-2018 for 11 Western states within the conterminous US. This paper classifies land into 4 classes: irrigated agriculture, dryland agriculture, uncultivated land, and wetlands

[202]. To do so, the authors compile an extensive labeled dataset containing 50,000 human-verified irrigated fields, 38,000 dryland fields, and 500,000km^2 of uncultivated lands and wetlands. These labels are the result of data aggregation from numerous state and federal agencies and research institutions, representing an unprecedented collection of verified irrigated areas. Using Google Earth Engine, Landsat imagery and auxiliary climate, meteorology, and terrain data, over 60,000 point samples across 28 years are then collected to train a random forest classifier. This method produces an overall binary irrigation vs. non-irrigated accuracy of 98%.

In [203], a similar approach is taken to map annual irrigation across the entire conterminous US from 1997-2017 using Landsat imagery and environmental variables. The authors of this paper take separate approaches for the eastern and western US. In the West, croplands with a maximum greenness index (GI) or EVI above an optimal threshold are determined to be irrigated. This threshold varies by year and by county and is determined via regression. In the eastern US, 30,000 potential training locations are manually selected, all from center pivot systems. Non-irrigated training sites for the entire US are extracted from [204]. In both settings, annual statistics (e.g., max, mean, range) of various derived Landsat bands (e.g., EVI, NDVI, water greenness index (WGI), and aridity normalized greenness index (AGI)) are used as feature layers. Classification in the eastern US also leverages MODIS imagery that is temporally smoothed and aggregated to annual and late season (May 15th to October 15th) statistical quantities. Classifiers are trained annually by county using a random forest architecture, achieving >90% pixel level accuracy. After classification, irrigated area filtering is performed using

National Land Cover Database (NLCD) products: predicted areas less than 2 Ha in size are discarded.

An early effort to track change in irrigated areas across the conterminous US merged MODIS imagery and USDA county level irrigation statistics [205]. Analysis focuses on 2002 and 2017 and uses annual peak NDVI as the primary metric in distinguishing irrigated and rainfed crops. Based on the assumption that irrigated crops demonstrate higher NDVI, the model assigns pixels to the irrigated class on a county-wide basis by ordering all pixels in descending order based on max NDVI. The corresponding number of pixels that align with the county wide irrigated area sum are deemed to be irrigated. This methodology is verified with Department of Water Resources data from California, revealing acceptable accuracies but large numbers of errors of omission.

Irrigation dynamics in the US Northern High Plains are monitored with Landsat imagery in [206]. In this paper, the authors produce 30m irrigation maps for 1999-2016 by combining all available Landsat satellite imagery with climate and soil covariates in Google Earth Engine, achieving accuracies between 92-96%. For classification, 9 Landsat variables and 11 covariables are used to distinguish between irrigated and rainfed phenologies, which typically demonstrate differences in their peak greenness indices. The training dataset contains 1401 labeled samples, created using high resolution National Agriculture Imagery Program (NAIP) scenes, EVI timeseries, the US Department of Agriculture (USDA) Cropland Data Layer (CDL), and expert assessment. The resulting classifier identifies novel vegetation indices – the water adjusted greenness index (WAGI = NDWI * GI) and the aridity-normalized greenness index (AGI =

GI/growing season aridity) – as most important for prediction, part of a comprehensive evaluation of variable importance contained within the paper.

Multiple papers have extended the methodology of [206] to map irrigation in other parts of the US. In [207], Xu et al. focus on southwestern Michigan between the years of 2001 and 2016. Here, 1536 labeled pixels are collected from visual inspection of Landsat imagery on Google Earth Engine based on the presence of visible irrigation infrastructure, high vegetation indices, and center pivot systems. Annual composites of the Landsat imagery (e.g. the mean, maximum, minimum, and range) are then used for classification, as the full Landsat timeseries have different lengths and availabilities across the area of interest based on cloud cover and satellite imaging geometry. A random forest trained on these composite statistics then achieves between 78-85% accuracy, although only after the additional bands introduced in [206] (WAGI, AGI) are included. With irrigation extent predictions between 2001 and 2016, this paper is also able to show correlation between irrigated areas and commodity prices.

Similarly, the primary author of [206] (along with a new set of coauthors) extends her original methodology for irrigation detection to the US High Plains Aquifer (an area covering portions of Nebraska and Kansas) in [208]. Here, irrigation presence is predicted from 1984 to 2017 using statistical composites of Landsat imagery. In total, 17,276 labels are collected via the same approach introduced in [206] that leverages visual inspection and Google Earth Engine. A random forest is used to predict irrigation extent, achieving 91% accuracy. Post-processing is also performed after prediction to decrease errors of commission.

1.5 International irrigation mapping at decameter resolution

Two additional studies that map irrigation in international settings at resolutions between 10m and 30m are particularly pertinent to our proposed method of irrigation detection. In southeastern Turkey, Landsat timeseries reveal simultaneous expansion and intensification of irrigated dry season cropping from 1990 to 2018 [209]. In this setting, irrigation is essential for crop growth during the near-zero precipitation time between July and September (similar to dry-season cropping dynamics in Ethiopia between December and April). The authors of this paper propose a binary classification approach for mapping – the two classes are irrigated and all other land cover types. Training data are comprised of 568 hand-drawn polygons between 9 Ha and 30 Ha in size, all collected during 2015. These training polygons yield 10,780 pixelwise samples in total. Statistical quantities of the collected Landsat timeseries (i.e., 25^{th} percentile, median, 75^{th} percentile, variance) are used as features for classification due to a lack of temporally consistent Landsat measurements across years. All steps in this process are performed on Google Earth Engine, and the final predictions achieve >90% accuracy.

Google Earth Engine facilitates an effort to identify irrigated paddy fields in the Philippines in [210]. Here, Sentinel-2 imagery is paired with a mixture of ground and visually-collected labels – 479 in total. For the visually collected labels, a custom GEE app was used to assess the quality of the referenced sample points. The Sentinel-2 imagery was then processed into 16 band composite images that include VI statistical quantities for each of the dry and wet seasons. The authors achieve 68% accuracy in the dry season and 75% accuracy in the wet season.

1.6 Global irrigation mapping at hectometer and kilometer resolutions

Using 250m MODIS imagery resampled to a global 25km grid, Zohaib et al. provide the most recent global irrigated areas map in [211]. To detect irrigation, this paper uses MODIS satellite imagery and three reanalysis-based irrigation-dependent variables: soil moisture, land surface temperature, and surface albedo; the proposed methodology assumes that irrigation is an unmodeled land surface process, and that satellite observations can detect this signal in near real-time. Compared to other previous global irrigation maps, including the Global Map of Irrigated Areas (GMIA) [212], the Global Irrigated Area Map (GIAM) [213], the International Water Management Institute Global Irrigated Area Map (IWMI-GIAM) [214], and the Global Irrigated Areas product [215], this map has reasonable agreement, overlapping with approximately 70% of the irrigated areas. In order to consider only irrigation in areas with known cropland, the authors mask their spatial extent with the ESA CCI land cover product for the year 2015 [216]. A major limitation of this study – similar to all previous global irrigated area maps – is the lack of ground truth required to verify performance; generation of reliable ground truth is difficult and expensive [217], a fact that has traditionally prevented the existence of large labeled datasets. Moreover, the hectometer and kilometer spatial resolutions of all global irrigated maps are too coarse to accurately map the distribution of smallholder irrigation, the type of which dominates across sub-Saharan Africa and is most relevant for the work at hand.

Ambika et al. looks at the Indian context in [218], using 250m MODIS NDVI imagery and 56m LULC data to develop irrigated area maps for the entirety of India for 2000-2015. Compared to agricultural statistics data for each state, the method achieves an R^2 value of 0.95. In this approach, training data is a set of NDVI curves associated with each of the crops considered,

along with whether irrigation is present. Based on the spectral similarity of NDVI curves at inference, classification decisions are made; this approach is called the Spectral Correlation Mapper. An LULC layer is also used to remove non-cropped areas from consideration, with final maps produced at 250m.

1.7 Land cover mapping

Given their frequent revisit, multiple spectral bands, and global extent, decameter and hectometer satellite imaging missions such as MODIS, Landsat, and Sentinel-2 have been used for a variety of land cover mapping applications. A set of papers – [219] and [220] – finds that land use maps are more accurate when vegetation indices are used instead of a full suite of spectral reflectances; moreover, they determine that when the number of endmembers are limited, spectral linear unmixing provides satisfactory land classification results.

A change-detection algorithm using 16-day MODIS normalized difference in vegetation index (NDVI) data is presented in [221], methodology which is used to detect and remove corrupted data values. Similarly, MODIS timeseries are decomposed into trend, seasonal, and remainder components in [222], showing that it is possible to extract features from within imagery time series at multiple time resolutions.

MODIS imagery has also been used to conduct land-use/land-cover (LULC) mapping in the Great Plains [223]. In this paper, the authors apply a decision tree classifier to multi-temporal NDVI data collected over the growing season for a series of classification problems that include determining: 1) crop presence, 2) general crop types, 3) specific summer crop types, and 4) irrigated/non-irrigated crops. All LULC predictions were compared to reported crop data, with overall accuracies ranging from 84% to 94%. These results were deemed satisfactory, although

there were some classification errors when cropped extent was smaller than the 250m resolution of the imagery.

Four sets of Landsat imagery are leveraged to determine LULC change over 25 years in the western Nile delta of Egypt in [224]; classification results were further augmented with visual interpretation. The authors also note that satellite imagery is the most appropriate data source for large-scaled LULC change detection, due to its repetitive acquisition, suitability for digital processing, and accurate georeferencing. In fact, multiple studies have shown that the 30m resolution of Landsat to be adequate for mapping vegetation [225,226] .

The authors of [227] use 10m-resolution Sentinel-2 imagery to monitor crop change around Barrax, Spain. The high spatial and temporal resolutions offered by Sentinel-2 allow for single crop fields to be distinguished, even at a relatively small size (a few hectares). Extrapolating this finding to settings in sub-Saharan Africa where crop fields are much smaller on average than those found in the US, 10m Sentinel-2 imagery is likely to provide a valuable tool for landscape analysis.

Crop change in Spain is also tracked in [228], although with Sentinel-1 Synthetic Aperture Radar (SAR) imagery. As SAR detects radio waves which are not impeded by the presence of clouds, SAR imagery is always cloud-free. Both Sentinel-1 and -2 imagery are incorporated for crop monitoring in [229]; by pairing longer-availability Sentinel-1 data with 10m resolution, multispectral Sentinel-2 data, researchers can gain multi-year insights into cropping behaviors.

In [230], Xiong et al. produce a 30m cropland extent map for continental Africa by integrating pixel-based and object-based algorithms on Sentinel-2 and Landsat-8 imagery with Google Earth Engine. The map is created from two biannual composites (July to December 2015,

and January to June 2016), with 9791 points of training data that was collected via visual inspection of 5m National Geospatial Agency (NGA) imagery by multiple analysts who are all country-specific experts. During training, labels are only used if they reside within a homogenous 90m square to avoid pixel mixing. The imagery composites and labels are then used to train a random forest; a SVM was also trained to compensate for defects in certain areas. The resultant model achieves an accuracy of 94% over 1754 independent validation samples. After training and inference, predictions are segmented into fields using the Recursive Hierarchical Segmentation (RHSeg) software. Following segmentation, polygons are assigned a cropland class: If >85% of polygon is predicted as cropland, the polygon is determined to be cropland. If <15% are assessed as cropland, then entire polygon is determined to be non-cropland. Otherwise, the polygon is kept as mixed cropland with its pixel based classifications unchanged. The final map is called Global Food Security-support Analysis Data @ 30m for African Continent Cropland Extent (GFSAD30AFCE).

Another paper of interest in this category is [231], whereby bimonthly Landsat composites, Google Earth Engine, and a random forest classifier are used for mapping croplands of Europe, the Middle East, Russia, and Central Europe. The resulting map contains binary classification of cultivated/non-cultivated land for 2015 at 30m resolution and is deemed the GFSAD30. For this study, training data takes the form of 64,000 3x3 sampling squares (i.e., 90m x 90m) that are selected across all areas of interest. Each sampling square is labeled as cropland/non-cropland by expert enumerators using very high resolution imagery and a double blind assessment. The final predictions achieve 94% accuracy, with NDVI and SRTM features layers being the most important for classification.

In a part of the world near the Ethiopian Highlands, [232] presents a method of differentiating fallow fields from cropland across the Sahel using Sentinel-2 imagery. For this work, 1M labels are collected over cropped areas only via the CGLS-L100 land cover map [233], and then split into two classes via unsupervised clustering. The pixels with the lowest 25% of NDVI values are kept as fallow samples; the pixels with the highest 25% of NDVI values are kept as cropped samples. From this set, 1000 cropped pixels and 1000 fallow pixels are randomly sampled for training, with another 100 of each class selected for validation. An accuracy of 84% is achieved with a random forest classifier

In Kenya, Sentinel-1 and Sentinel-2 imagery are combined for crop mapping over heterogeneous landscapes in 2018. Here, all available Sentinel-2 images are stacked, mosaiced, and then reduced to a median value for each of the four seasons. Various vegetation indices are also derived from the Sentinel-2 imagery, including NDVI, EVI, EVI2, SAVI, and GNDVI. All Sentinel-1 images are calibrated, despeckled, terrain-corrected, stacked, and then reduced to median seasonal values. Labeled data comes in the form of 872 polygons collected from a field survey that covers 12 classes and 10,982 pixels. Using a random forest, the authors achieve between 75% and 90% accuracy based on the crop type, and also present a comprehensive overview of feature importances for classification.

While classification methodologies abound for distinguishing land cover types, standardized datasets for this task are few and far between. Tsendbazar et al. address this gap in the literature by developing and applying a multi-purpose land cover validation dataset for Africa in [234]. The resultant dataset has a 100m resolution and is compiled from 3617 sample sites across the continent. Each sample site is annotated by one of 6 regional experts using the

GeoWiki platform. This dataset is then used to validate the CGLS-LC100 land cover map, which is assessed to have an accuracy of 75%.

A data fusion based approach is taken to map crop types and crop sequences in Germany from 2017-2019 in [235]. For this work, Sentinel-2 and Landsat 8 imagery are resampled to 10m and collected at 5 day timesteps; Sentinel-1 imagery is produced as monthly mean composites. Using the Land Parcel Identification Service (LPIS), 1000 pixels per crop type are extracted from administrative polygons. In total, 483 features are fed into a random forest for classification; the final model achieves between 78-80% accuracy. A subsequent evaluation of feature importance reveals that the three most important features are SAVI, then NDVI, then NIR, indicating that radar imagery may not provide additional value for classification when optical imagery is available.

A very similar effort maps winter crops in Germany from 2016-2019 [236]. Here, Sentinel-2 imagery is collected between July and April, yielding between 10 and 40 different timesteps per year. This imagery is combined with 1000 polygons for each class – winter catch crop, or no winter catch crop – with each polygon yielding a single averaged pixel timeseries. Various vegetation indices are then derived from the Sentinel-2 data and combined with the labels and a random forest classifier to achieve 84% accuracy.

In [237], Gumma et al. use Sentinel-2 imagery for crop type identification and spatial mapping for three districts in India. The Sentinel-2 imagery consists of NDVI measurements in 15-day timeseries at 10m resolution. A spectral matching approach is taken for classification, whereby NDVI signatures are compared to ideal temporal profiles of various crops for

classification. The labeled data comes from a ground survey collected in 2019 and is comprised of 732 sample points in total.

1.8 Endmember-based land cover mapping

Spectral endmembers are proxy spectra for materials on the ground [238]; temporal endmembers are timeseries that represent phenological change in landscapes. Endmember modeling can quantify continuous spatial and temporal change, in contrast to classification approaches that focus on discrete categorization.

The different methods of endmember techniques and their relative advantages are exhaustively reviewed in [239]. While endmembers retrieved from a library are known to accurately represent a specific spectrum in a certain environment, these spectra are rarely acquired under the same conditions as airborne data. In contrast, image endmembers can be easily associated with features on the scene, as they occupy the same distribution as the data; however, image endmembers are often noisy, difficult to scale, and require domain knowledge to extract. The paper also notes that when unmixing images with endmembers to create abundance maps, two constraints are usually imposed: the abundance nonnegativity constraint, whereby no endmember contribution to a pixel timeseries can be negative; and the abundance sum-to-one constraint, whereby the contribution of individual endmembers to a pixel timeseries must sum to one. This finding dovetails with others that note how nonnegativity and sum-to-one constraints are usually implemented to ensure physically meaningful abundance estimates [240].

Convex cone analysis (CCA) is a frequently used method for selecting endmembers from an image. Here, a user extracts endmembers from the exterior of a convex region formed by the

pixels in pixel space, leveraging the assumption that the timeseries of pixels residing within the convex region can be recreated by linear combinations of pixels that reside on the exterior of the space [241,242].

In determining the appropriateness of selected endmembers, [243] proposed root mean square error (RMSE) as a method of comparing abundance maps to the basis imagery, coining the term “Multiple Endmember Spectral Mixture Analysis” (MESMA) in the process. While much of the original work on MESMA focused on how best to select and combine spectral endmembers for a single multispectral image, researchers have also determined that endmember decomposition strategies can be applied to temporal image stacks. In fact, [244] found that when time series are used to define endmembers, subpixel classification accuracy can drastically increase as plant type specific phenological development cycles will be captured in different temporal reflectance profiles.

A weighted loss strategy for spectral unmixing is demonstrated in [245]. Here the authors apply different weights to endmember recreation losses in order to influence the resulting abundance map, a useful technique for recreating spectra of different scales.

An iterative method of extracting endmembers best able to recreate an original image, determined by overall residual error, is proposed in [246]. Others papers have extended this approach to select endmembers from among multi-temporal, seasonally-mixed spectral libraries [247]. While an iterative process ensures endmembers that can closely recreate the original data, it is a computationally intensive approach, especially when applied to hyperspectral data cubes [248]. A tradeoff therefore exists between accuracy and computational intensity in the endmember selection process [249].

Other more complex methods for endmember extraction exist in the literature, including one that takes a Bayesian approach to understand the impact of endmember variability on subpixel vegetation fractions in an urban environment [250]. Here, an iterative approach selects the best endmember signatures for an image; however, the endmembers are no longer treated as constants but represented by probability density functions, thus incorporating spectral signature probability.

The issue of endmember organization in high-dimensional space is addressed in [251] and [252] by applying a principal component analysis (PCA) to the pixel spectra. Spectra for adjacent pixels across a continuous land cover setting will be highly correlated, and as result, will contain redundant information. In lowering the dimensionality of the pixel space, the principal component transform organizes spectra along dimensions which capture the largest amount of variance in the image, making the hand-selection of endmembers an easier task.

Urban irrigation is predicted in [253] using a combination of Landsat imagery, aerial imagery, climatic records, and land use maps. Methodologically, the paper uses 3 endmembers – representing irrigated vegetation, non-irrigated-vegetation, and impervious substrate (i.e., airport buildings, runway, pavement) – to estimate the fraction of each pixel covered by irrigated landscaping.

Land use change over 6 years is monitored in [254] using a collection of MODIS NDVI 16-day composites. Three predominant land use classes are detected: orchard, non-cultivated area, and annual crops. Endmembers representing the typical signatures of these land classes are retrieved annually using an automatic extraction mechanism that selects the pixel profiles that most accurately capture the variance in the study area. The authors validate their results using a

land use map and 30m resolution Landsat imagery, achieving a minimum classification error of 0.1 at 1km spatial resolution.

Professor Christopher Small has published multiple papers demonstrating the robustness of using spectral endmembers to characterize and model landscapes. [255] demonstrates that it is possible to represent 98% of image variance in 30 spectrally diverse Landsat ETM+ subscenes with three endmembers corresponding to substrate, vegetation, and dark surface (SVD) spectra. Professor Small builds upon this work in a later paper, finding that dark, vegetated, and high albedo endmembers can faithfully recreate spectral signatures of 28 urban settings: For just about all urban images, the RMS error in the abundance map is less than 0.1 [256]. Moreover, [257] demonstrates that endmembers retrieved from Landsat ETM+ subscenes may be used to model mixed reflectance spectra from a Worldview-2 sensor, indicating that linear spectral mixture models can be standardized across spatial scales using common endmembers. Taken together, these results demonstrate that it is possible to represent accurately scene settings of multiple spatial resolutions with a limited number of endmembers, provided they enclose the majority of the global mixing space.

2. Initial vegetation characterization applications in Ethiopia

2.1 Background on irrigation in Ethiopia

With 110 million people, Ethiopia is the second most populous country in Africa; it also has the fastest growing economy on the continent, experiencing growth averaging 9.9% a year between 2007 and 2018 [258]. In Ethiopia, agriculture contributes about 44% of GDP and 70% of export earnings, and is dominated by small land-holdings between 0.5 ha and 2 ha in size [259].

Irrigated agriculture only accounts for 5% of food crops in the country; 37% of all vegetable production is irrigated; and industrial crops such as sugarcane, cotton, and fruit are mostly irrigated [259]. However, despite its relative underdevelopment of agriculture with modern irrigation, Ethiopia has a large potential for growth: The Food and Agriculture Organization of the United Nations (FAO) estimates that the country's potential irrigable land amounts to 2.7 million ha [260]. Additionally, Ethiopia has 12 river basins providing an estimated average annual runoff of 125 billion cubic meters, 45% of which is contributed by the Abbay basin. Estimates of groundwater potential vary from 2.6 to 13.5 billion m^3, but as a lack of reliable data exists to verify these assessments, some experts suggest that the amount could be much higher [261].

Detailed studies of existing irrigation schemes in Ethiopia are scarce, but they do exist: One report assesses 12 irrigation schemes, ranging in size from 1200 to 14,600 ha for medium and large size schemes and from 850 to 2224 ha for small-scale irrigation schemes [262]. The study finds that large-scale irrigation initiatives cultivate sugarcane, cotton, tobacco, and fruits such as oranges, mandarin, tomato, guava, grape vine; small-scale irrigation schemes commonly crop onion, potato, pepper, tomato, and banana. Furthermore, most of the studied irrigation schemes use diverted water from perennial rivers by means of a weir or direct pumping. Surface irrigation, usually via trapezoidal earthen canals, is the most common irrigation method; however, the authors have noted a gradual change from field ditches and siphons to more labor-intensive, water-saving technology [262].

Another published study on irrigation extents evaluates larval habitats in irrigated areas in the Rift Valley [263]. This study aligns with recent conversations with local partners regarding irrigation development in Ethiopia: many new projects are underway in the Rift. Other authors

find that a lack of governance capacity and accountability are critical challenges for the sustainability of the irrigation projects in Ethiopia, issues which are often compounded by a lack of equitable access to the irrigation schemes [264].

2.2 Characterization via temporal mixture modeling

The 2012 paper from Professor Small discussed in the last paragraph of Chapter IV, Section 1.8 provides a method of producing regional vegetation phenology maps by applying a temporal mixture model to timeseries of vegetation abundance layers [183]. The first step in this process involves transforming a timeseries stack of two-dimensional enhanced vegetation index (EVI) layers – also known as an image cube – into principal component (PC) space. For reference, EVI is calculated from the near infrared (NIR), red, and blue bands, per Eq. (IV-1):

$$EVI = \frac{2.5 * (NIR - RED)}{(NIR + 6 * RED - 7.5 * BLUE + 1)} \tag{IV-1}$$

An example EVI image cube is shown in Figure IV-1. This figure is taken directly from [183]; it is included here because of its clear and effective representation of the three-dimensional nature of a timeseries stack of single band layers.

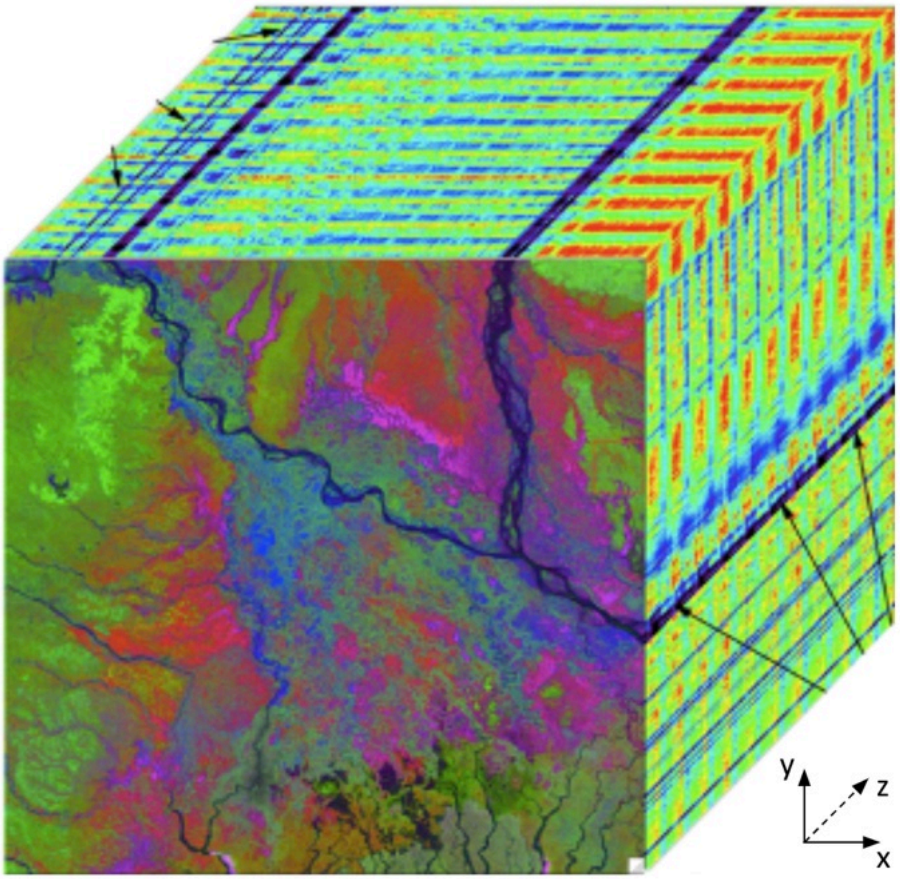

Figure IV-1: Example imagery cube published in [183]. This image cube contains a 10-year stack of enhanced vegetation index (EVI) layers over the Ganges-Brahmaputra Delta in southern Bangladesh. The x- and y-dimensions of the image cube characterize spatial extent, while the z-dimension represents time.

Dimensionality reduction via the PC transform reduces the amount of highly correlated, redundant information in the EVI image cube. Next, temporal endmembers (tEMs) are extracted from the exterior of the PC space representation of the EVI image cube; these tEMs, representing unique vegetation phenologies, are selected to bound the temporal feature space of all vegetation phenology cycles observed in the area of interest. Figure IV-2(a) displays an example point cloud distribution of EVI timeseries in PC space, and (b) shows the result of fitting a convex hull to the feature space. In this figure, the convex hull is shown to illustrate the exterior of the

point cloud, i.e. from where tEMs will be extracted. Extracting tEMs from the exterior of the image cube point cloud leverages the assumption underlying convex cone analysis: Linear combinations of endmembers on the exterior of the pixel space can recreate the signatures of interior pixels.

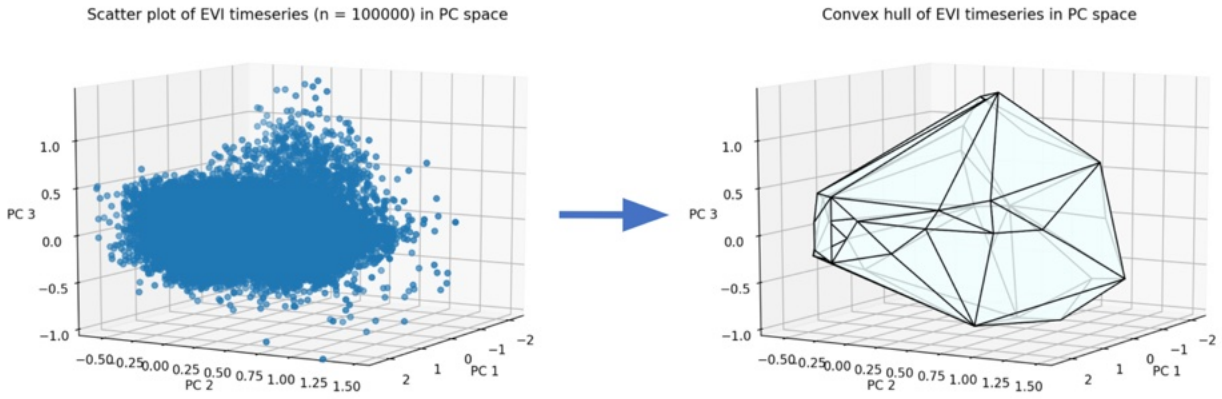

Figure IV-2: (a) Example enhanced vegetation index (EVI) point cloud for an image cube transformed into principal component (PC) space; and (b) a convex hull fit to the point cloud, indicating from where temporal endmembers will be extracted, per the temporal mixture modeling approach introduced in [183].

Extracted tEMs form the basis of a linear temporal mixture model that can be inverted to provide tEM fraction estimates for each pixel's vegetation phenology; these resultant estimates constitute an endmember abundance map. Through this inversion process, users can determine the degree to which select vegetation phenologies contribute to the overall pixel timeseries.

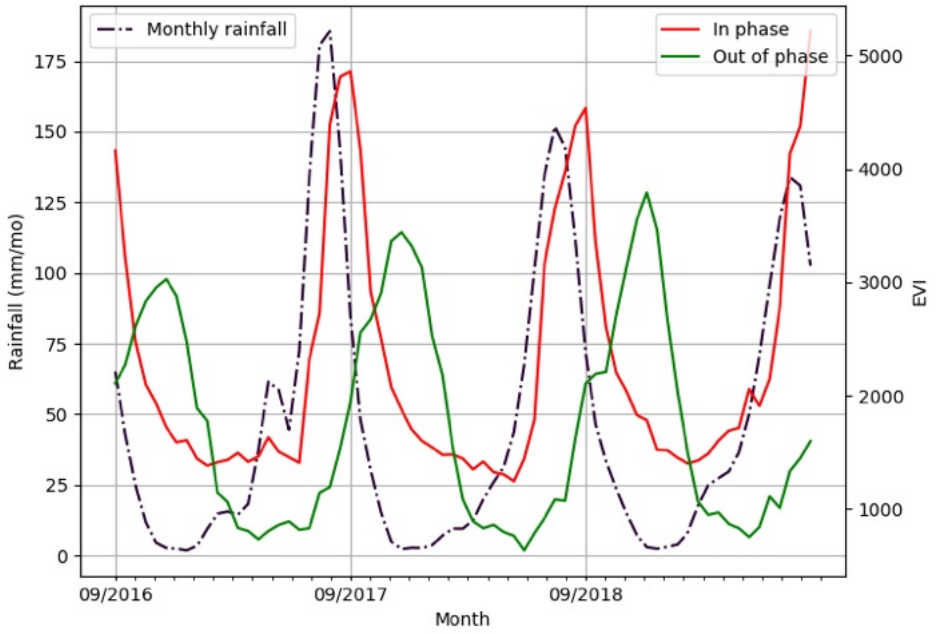

Figure IV-3: Extracted temporal endmembers (tEMs) in-phase and out-of-phase with precipitation estimates provided by Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) in Amhara, Ethiopia. tEMs are extracted from an image cube of 16-day MODIS enhanced vegetation index (EVI) layers between September 2016 and September 2019. Monthly precipitation values are measured on the left-hand y-axis; EVI values are measured on the right-hand y-axis.

Figure IV-3 shows two extracted tEMs from an EVI image cube created with 16-day MODIS EVI layers between September 2016 and September 2019 over the Amhara state of Ethiopia; and average monthly rainfall estimates for the entirety of Amhara from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) [265]. The tEM that closely follows the rainfall timeseries is shown in red and deemed “in-phase”; the tEM that represents vegetation growth that cannot be attributed to rainfall is shown in green and is deemed “out-of-phase”. Together with a zero-valued tEM representing non-vegetated surfaces, these tEMs are used to create an endmember abundance map that displays the existence of vegetation in-phase and out-of-phase

with rainfall in Amhara; the ensuing map is presented in Figure IV-4. For this temporal mixture model, sum-to-one least square unmixing is selected as the inversion strategy.

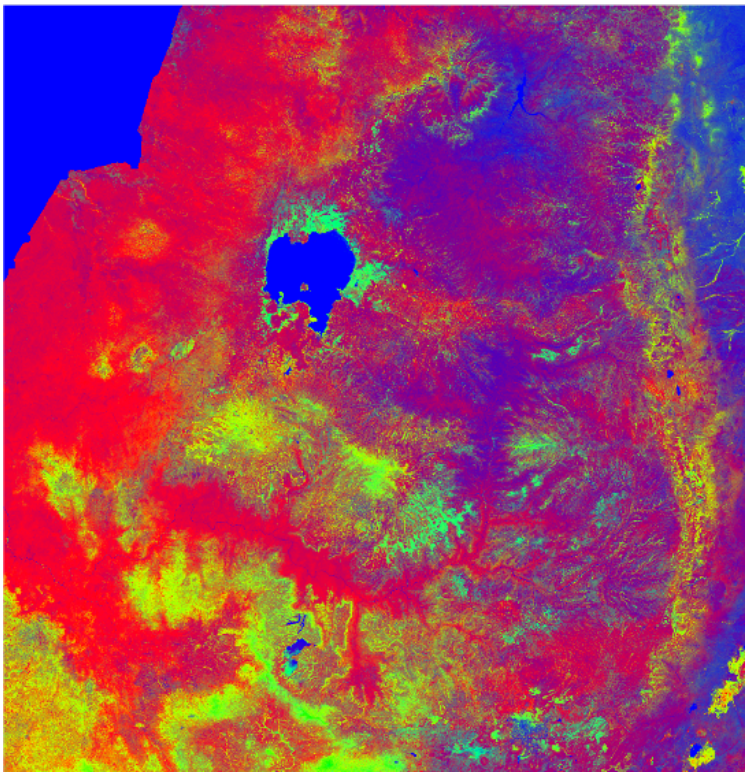

Figure IV-4: Temporal endmember abundance map over Amhara, Ethiopia, created from 16-day MODIS enhanced vegetation index (EVI) layers between September 2016 and September 2019. Single phase endmember abundance is presented in red; evergreen endmember abundance is presented in green; non-vegetated endmember abundance is presented in blue.

In Figure IV-4, in-phase vegetation abundances are shown in the red channel, out-of-phase vegetation abundances are shown in the green channel, and non-vegetated abundances are shown in the blue channel. Given its red-dominant hues, Figure IV-4 demonstrates that vegetation in Amhara primarily cycles in-phase with the state's main summer rains. However, green pixels around Lake Tana, seen in blue in the center of the figure, indicate the presence of out-of-phase vegetation growth. Out-of-phase vegetation growth can also be seen further south

around Choke Mountain, and along the Main Ethiopian Rift (MER) running north-south in the far right of Figure IV-4. While future work will need to assess how accurately an abundance map can characterize vegetation phenologies in Amhara, and whether other tEMs should be included in the modeling process, Figure IV-4 provides a preliminary example of how continuous, spatiotemporal characterization via a temporal mixture model can identify vegetation in Ethiopia not attributable to rainfall.

2.3 Characterization via statistical analysis

Another method of characterizing vegetation growth in Ethiopia assesses various statistical properties of vegetation and rainfall timeseries to determine misalignment between the two indicative of dry season crop growth. Here, vegetation phenologies are determined from 16-day 250m MODIS EVI imagery between June 1, 2011 and June 1, 2021. Precipitation estimates are compiled using the CHIRPS quasi-global rainfall dataset [265]: Monthly 0.05-degree precipitation layers are resampled temporally and spatially to the EVI image cube resolution via bilinear interpolation. The CHIRPS image cube is shifted backed by 2 timesteps (32 days) to align with expected vegetation growth and is then normalized to be between 0 and 1.

Figure IV-5 displays the misalignment between vegetation phenology and rainfall in Ethiopia. The misalignment between EVI and shifted rainfall, β, is presented in the red channel of Figure IV-5: the redder the image appears, the more misalignment. β is calculated with the following two equations:

$$\alpha_x = \left| EVI_x - \lambda_x * CHIRPS_{norm,x} \right| \tag{IV-2}$$

$$\beta_x = \frac{f_{90}(\alpha_x)}{mean(EVI_x)} \tag{IV-3}$$

where *x* indicates the pixel location in two-dimensional space; *EVI* and *CHIRPS* represent the 10-year phenology and rainfall image cubes, respectively; λ_x is specified to minimize the RMS of timeseries α_x; *norm* indicates an image cube normalized between 0 and 1; and f_y finds the y^{th} percentile of a timeseries. In Eq. (IV-2), α_x is the absolute difference between the vegetation and rainfall timeseries at pixel location *x*, after the λ_x multiplier scales the normalized CHIRPS timeseries to account for the vectors' differing units. In Eq. (IV-3), we use the 90th percentile function to extract the high-end of the phenology-rainfall misalignment, as dry season crop phenologies will contain timesteps with high vegetation-rainfall misalignment. We then divide by the mean timeseries EVI to reduce the visibility of high misalignment over evergreen areas.

The green and blue channels in Figure IV-5 both present the maximum difference in vegetation strength over the 10-year time period, calculated as the pixelwise difference between the 90th percentile and the 10th percentile of EVI image cube, per Eq. (IV-4). This quantity, termed γ, highlights parts of Ethiopia where there is substantial change in vegetation strength across the timeseries, as is the case with cropland (and to a lesser extent in evergreen areas). In contrast, sparsely vegetated and barren parts of the country will have little change in vegetation strength across the timeseries and will contain low values in both green and blue channels.

$$\gamma_x = f_{90}(\alpha_x) - f_{10}(\alpha_x) \tag{IV-4}$$

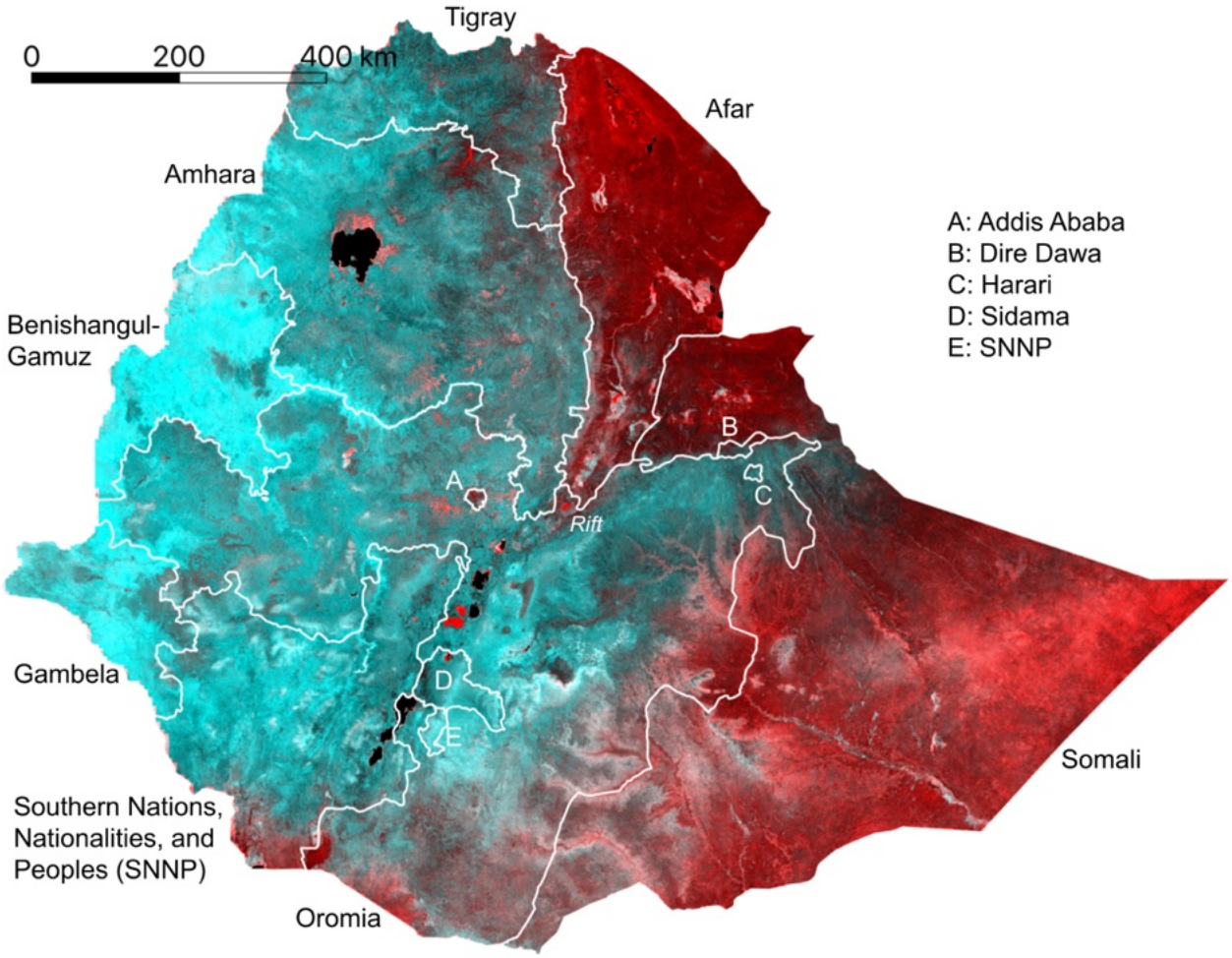

Figure IV-5: Vegetation phenology-rainfall misalignment map. The red channel in the figure contains the misalignment between normalized rainfall and EVI, β (see Eqs. (IV-2) and (IV-3)); maximum EVI difference, γ, is presented in the green and blue channels (See Eq. (IV-4)). A 2% linear stretch is applied, with administrative boundaries outlined in white.

From Figure IV-5, one can see that the bulk of the Western Ethiopian Highlands presents as cyan, indicating 1) low misalignment between vegetation phenology and rainfall; and 2) substantial difference between the maximum and minimum EVI values. However, pockets of Amhara and Tigray do contain red and pink pixels: As higher values in the red channel designate greater misalignment between vegetation phenology and rainfall, such coloration signals that

dry-season irrigation may exist in these areas. We also note that the high misalignment values in Somali and Afar are due to the sparse vegetation that populates much of these areas – vegetation which yields small denominators in Eq. (3) – and as such are of not of our concern. Accordingly, Figure IV-5 provides a comparison point for future dry season crop classifiers in Ethiopia: If a more complex classifier reveals dry season cropping in the areas of Figure IV-5 that appear red/pink surrounded by cyan, this statistical approach of vegetation characterization will be shown to identify areas of interest in a quicker, more transparent manner.

3. Conclusion

Locating areas that currently sustain dry season crop growth informs energy system planners as to where productive electricity use can be expanded: These are areas without any barriers to irrigation, such as excessively deep water tables, no market access, or a lack of local know-how. To locate areas of dry season crop growth, it is necessary to pull from the literature on vegetation characterization and irrigation detection, including those authored by Professors Small and DeFries. Chapter IV, Section 1 contains a review of relevant papers for the task of identifying dry season crop growth.

While there are many endmember-based approaches to understanding vegetation conditions via satellite imagery, a 2012 paper by Professor Small that introduces the concept spatiotemporal characterization using a temporal mixture model is particularly relevant for the work at hand [183]. Adapting this approach to Ethiopia can characterize the main types of vegetation growth throughout the entire country, facilitating the detection of vegetation cycles in-phase and out-of-phase with rainfall. A separate strategy for detecting dry season vegetation curates a set of statistical properties that indicate the misalignment of rainfall and vegetation

growth. This second methodology provides a continuous characterization of vegetation-rainfall dynamics, one that informs users about areas where vegetation is likely not attributable to rainfall. Given the relatively straightforward manner of computing these statistical quantities, results from this method of characterization can also be compared to predictions from more complex dry season irrigation detectors to determine how closely quicker analysis can approximate more robust predictions.

Altogether, a summary of relevant vegetation and irrigation mapping methods provides a basis for the development of a formal irrigation detection classifier for the Ethiopia highlands, a process contained in the following chapter.

Chapter V: A multiscale spatiotemporal approach for smallholder irrigation detection

Abstract

In presenting an irrigation detection methodology that leverages multiscale satellite imagery of vegetation abundance, the following chapter introduces a process to supplement limited ground-collected labels and ensure classifier applicability in an area of interest. Spatiotemporal analysis of MODIS 250m Enhanced Vegetation Index (EVI) timeseries characterizes native vegetation phenologies at regional scale to provide the basis for a continuous phenology map that guides supplementary label collection over irrigated and non-irrigated agriculture. Subsequently, validated dry season greening and senescence cycles observed in 10m Sentinel-2 imagery are used to train a suite of classifiers for automated detection of potential smallholder irrigation. Strategies to improve model robustness are demonstrated, including a method of data augmentation that randomly shifts training samples; and an assessment of classifier types that produce the best performance in withheld target regions. The methodology is applied to detect smallholder irrigation in two states in the Ethiopian Highlands, Tigray and Amhara, where detection of irrigated smallholder farm plots is crucial for energy infrastructure planning. Results show that a transformer-based neural network architecture allows for the most robust prediction performance in withheld regions, followed closely by a CatBoost random forest model. Over withheld ground-collection survey labels, the transformer-based model achieves 96.7% accuracy over non-irrigated samples and 95.9% accuracy over irrigated samples. Over a larger set of samples independently collected via the introduced method of label supplementation, non-irrigated and irrigated labels are predicted

with 98.3% and 95.5% accuracy, respectively. The detection model is then deployed over Tigray and Amhara, revealing crop rotation patterns and year-over-year irrigated area change. Predictions suggest that irrigated area in these two states has decreased by approximately 40% from 2020 to 2021.

1. Introduction

Between 1970 and 2008, global irrigated area increased from 170 million to 304 million hectares [185]. In sub-Saharan Africa however, as little as 4-6% of cultivated area is irrigated, given the lack of electric grid infrastructure and the high cost of diesel [266]. Locating isolated irrigation identifies areas that can support higher quality energy provision services -- e.g. a grid connection or minigrid installation -- as these sites can sustain higher energy demands and the attendant electricity costs [267]. Facilitated through informed planning, irrigation expansion has a direct impact on poverty reduction: In Ethiopia, one study found that the average income of irrigating households was double that of non-irrigating households [181].

In data poor locations, satellite imagery provides a source of detailed synoptic observations of irrigated agriculture [268]. A previous irrigation mapping effort in Ethiopia used three 1.5m resolution SPOT6 images to distinguish between large-scale and smallholder irrigation in the Ethiopian rift [186]. This approach was then adapted to intake a timeseries of 10m Sentinel-2 imagery to predict irrigation presence across the horn of Africa [185]. While both studies demonstrated high accuracies over collected observations, limited labels precluded a more rigorous performance assessment over the entire area of interest. Other studies have used multiscale imagery to detect irrigation, including one that fuses MODIS and Landsat imagery to identify irrigated extent, frequency, and timing in northwestern China [190]. Here, unique

advantages of satellite imagery products at different resolutions are exploited: 250m MODIS imagery is valuable for characterizing vegetation over large areas [269], while decameter resolution imagery from Landsat or Sentinel-2 missions can better discern plot extent [270].

Deep learning techniques have become widely used for land process classification, as they uncover intricate structures in large, complex datasets [271]; and provide a robust method of handling phenological variability [272]. Deep learning approaches using 1D and 2D convolutional neural networks (CNNs) have demonstrated strong performance for remotely-sensed image classification tasks, e.g. [273,274]. Researchers have also deployed more complex deep learning model architectures for improved performance on LC monitoring tasks: In lieu of a full review of relevant papers, [275,276] use recurrent neural networks to learn a joint spectral-spatial-temporal feature representation; while [277,278] provide two examples of cropland mapping with Long Short-Term Memory (LSTM)-based networks. A few studies have also explored using transformer blocks – encoders that have achieved state-of-the-art performance in many natural language processing tasks – for LC classification: Among others, [279,280] show that transformer models can effectively extract temporal dependencies that contribute to high-level feature learning, and [281,282] demonstrate that transformer-based networks can outperform CNNs for classification of imagery timeseries.

Despite increasing availability of remotely sensed imagery, computing resources, and advanced algorithms for information extraction, high-quality labels remain scarce and expensive to acquire. Methods of overcoming label scarcity generally fall into one of four categories: 1) using pretrained networks; 2) unsupervised and self-supervised learning; 3) data augmentation; or 4) additional label collection [283]. Even as pretrained networks like ImageNet [284] are highly

effective for true-color image classification, these networks' weights do not translate to tasks that intake multispectral or hyperspectral imagery [285]. Unsupervised learning techniques, including those that ensemble different clustering methods -- e.g. [286] -- have been shown to effectively organize unlabeled imagery. Existing work has also demonstrated that training a Generative Adversarial Network (GAN) -- itself a type of unsupervised learning -- has allowed for improved change detection performance on multispectral imagery, e.g. [287]. For data augmentation, three techniques are often implemented: image translation, rotation, and flipping [288,289]; however, these techniques do not have obvious analogues for pixel-based classification. Lastly, methods of collecting labels have included thresholding pixels based on normalized difference in vegetation index (NDVI) [192] and inspecting a single layer of high-resolution imagery [186].

Another lingering issue in land process mapping is determining the conditions under which a model can be utilized in locations beyond where it was trained. Site-specific methods may not be easily transferable to other places or climes [184,290], and the performance of transferred models can often only be assessed *after* full implementation in a novel setting [291]. Therefore, processes that yield insights about model transferability *before* training and inference offer benefits to researchers seeking to understand the maximum spatial applicability of their approaches.

As current methods primarily focus on already well-understood areas of interest with existing datasets, new techniques and products need to be developed for parts of the world lacking labeled data. In the realm of irrigation detection, new methodologies and mapping products can help identify locations for further energy system planning and investment, as these

areas contain latent energy demands that can make higher quality energy services cost-effective and increase incomes. To this end, the following chapter presents a multiscale methodology that leverages 250m MODIS imagery for regional phenological characterization and 10m Sentinel-2 imagery for irrigation detection on smallholder plots. This approach is then applied to the 205,000 km^2 Ethiopian Highlands, whereby it introduces a novel method of label collection; an evaluation of different classifier architectures and training strategies that ensure model applicability within the area of interest; and an assessment of irrigated area in the Tigray and Amhara states of Ethiopia for 2020 and 2021.

3. Background

Identification of dry season greening as potentially irrigated agriculture must account for spatiotemporal variations in native vegetation phenological cycles. The complex topography of the Ethiopian Highlands and East African rift system, combined with the latitudinal movement of the InterTropical Convergence Zone (ITCZ) and seasonal upwelling of the Somali current in the Arabian Sea produces a diversity of rainfall patterns that control annual vegetation phenological cycles in the study area[xix]. In order to provide phenological context with which to identify anomalous dry season greening, a regional vegetation phenology map is derived from spatiotemporal analysis of timeseries of vegetation abundance maps. Using the spatiotemporal characterization and temporal mixture modeling approach given by [183] applied to timeseries of MODIS enhanced vegetation index (EVI) maps, four temporal endmember (tEM) phenologies are identified that bound the temporal feature space of all vegetation phenology cycles observed

[xix] See [336] for a fuller discussion of rainfall patterns in Ethiopia.

on the East African Sahel. These four tEM phenologies form the basis of a linear temporal mixture model that can be inverted to provide tEM fraction estimates for each pixel's vegetation phenology. Figure V-1 presents a spatiotemporal phenological characterization for the country, created from 16-day 250m MODIS EVI imagery between June 1st, 2011 and June 1st, 2021.

The four tEMs extracted for Ethiopia are as follows: a *single cycle* tEM, representing a single annual vegetation cycle per year that peaks in September/October; an *evergreen* tEM, representing perennial vegetation; a *double cycle* tEM, representing semiannual vegetation cycles observed on the Somali peninsula; and a *non-vegetated* tEM, representing barren or non-existent vegetation. The ensuing phenology map in Figure V-1 contains unmixing root mean square (RMS) error less than 10% for 90% of the pixels; additional unmixing error statistics and the locations of the extracted tEMs in principal component (PC) feature space are shown in Figures V-2 and V-3.

Figure V-1 roughly divides into 4 quadrants. In the northeast quadrant, Afar appears as dark green, indicating that none of the 4 tEMs contribute significantly to phenologies in this part of the country: The vegetation that does exist in this mostly barren area is represented by low levels of evergreen tEM abundances. In the southeast quadrant, dominated by Somali and a portion of Oromia, vegetation patterns cycle twice annually. This is an area with bimodal rainfall but low total annual precipitation that results in the *double cycle* tEM containing peak vegetation abundances lower than those of the *single cycle* and *evergreen* tEMs. It follows that southeast Ethiopia is more pastoral with sparser vegetation than other parts of the country.

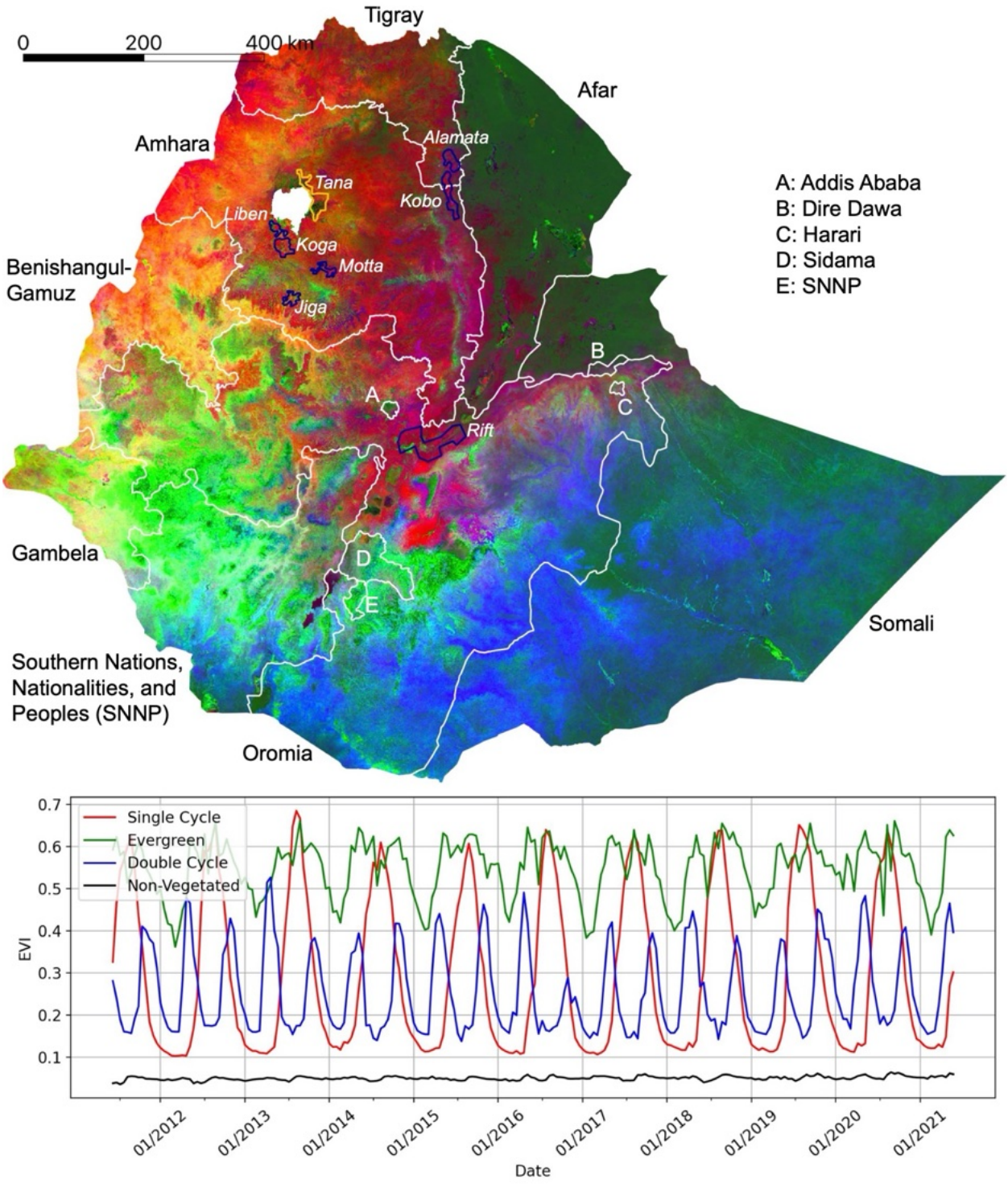

Figure V-1: Continuous endmember fraction map derived from a temporal mixture model of 250m MODIS enhanced vegetation indices (EVI). Smooth gradients and abrupt transitions in phenology are primarily related to topography and variations in precipitation. Region names showing locations of labeled polygons are italicized: The region containing ground collection (GC) labels is delineated in gold; the regions containing visual collection (VC) labels are delineated in blue.

The southwest quadrant – covering Southern Nations, Nationalities, and Peoples' (SNNP) Region, Sidama, and the western portion of Oromia – contains significant amounts of evergreen vegetation, as is demonstrated by its bright green hue. Here, evergreen vegetation is supported by bimodal rainfall with higher levels of annual precipitation than in eastern Ethiopia. In contrast, the northwest quadrant of the phenology map contains red-dominant color gradients, indicating phenologies similar to the *single cycle* tEM. This portion of the country, known as the Ethiopian Highlands and comprising of Amhara and Tigray, is highly agricultural; the main cropping season lasts from June to October and coincides with the primary *kiremt* rains, with some secondary cropping following the lighter *belg* rains from March to May. Accordingly, cropping that occurs during the dry season between November and March is likely to be irrigated.

In presenting a map of dominant vegetation phenologies in Ethiopia, Figure V-1 provides a guide for land cover classification applicability within the country. For instance, a dry season irrigation detector trained in Amhara will perform poorly in SNNP, as phenological patterns differ significantly across these states, and dry season crop cycles exhibit different vegetation signatures. In contrast, a dry season irrigation detector developed across Amhara can be transferred to Tigray or Benishangul-Gamuz, due to regional phenological similarities.

The named, italicized outlines in Figure V-1 represent the 8 areas containing labels used in this manuscript, referred to as *regions*: The yellow outline indicates a region where labels were collected via a ground survey, and the purple outlines indicate regions where labels were collected by means of visual interpretation and timeseries inspection. Full information on the labeled data collection process is presented in Section V-3.

Figure V-2 presents the locations of tEM extraction from the image cube transformed into principal component (PC) space. The four extracted tEMs are then used to create the phenology map via unconstrained least-squares linear unmixing per the methodology introduced in [183]. Relatedly, Figure V-3 presents the temporal mixture model inversion error and the cumulative error statistics for Figure V-1. Interpreting Figure V-1 and V-3 together reveals that the locations of highest error occur over evergreen vegetation, primarily in the southeast of Ethiopia. As the unmixing error remains low over Tigray and Amhara, the authors stipulate that Figure V-1 contains an accurate assessment of vegetation cycles in the area of interest.

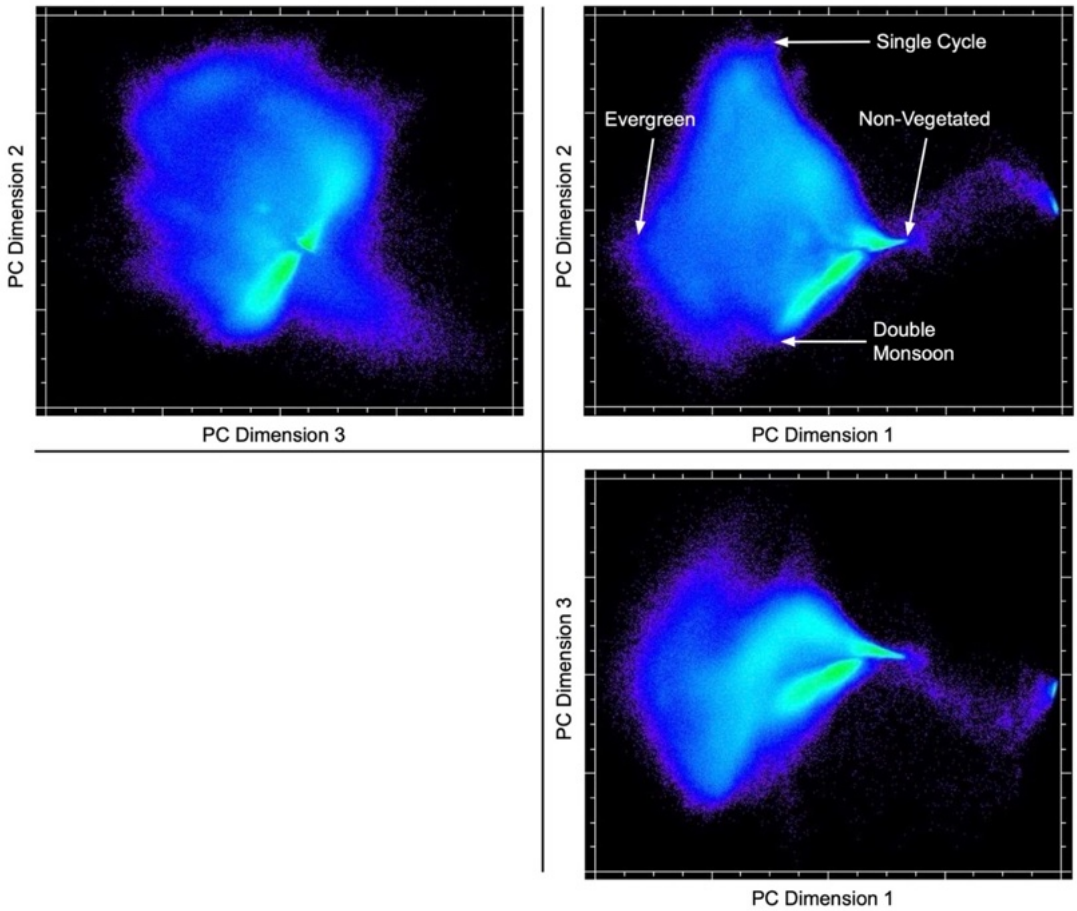

Figure V-2: Views of the first three principal component (PC) dimensions of the transformed 10-year Ethiopia MODIS enhanced vegetation index (EVI) imagery cube. Locations of temporal endmembers used to construct the phenology map in Figure V-1 via unconstrained linear unmixing are presented in the PC Dimension 1 vs. PC Dimension 2 plot.

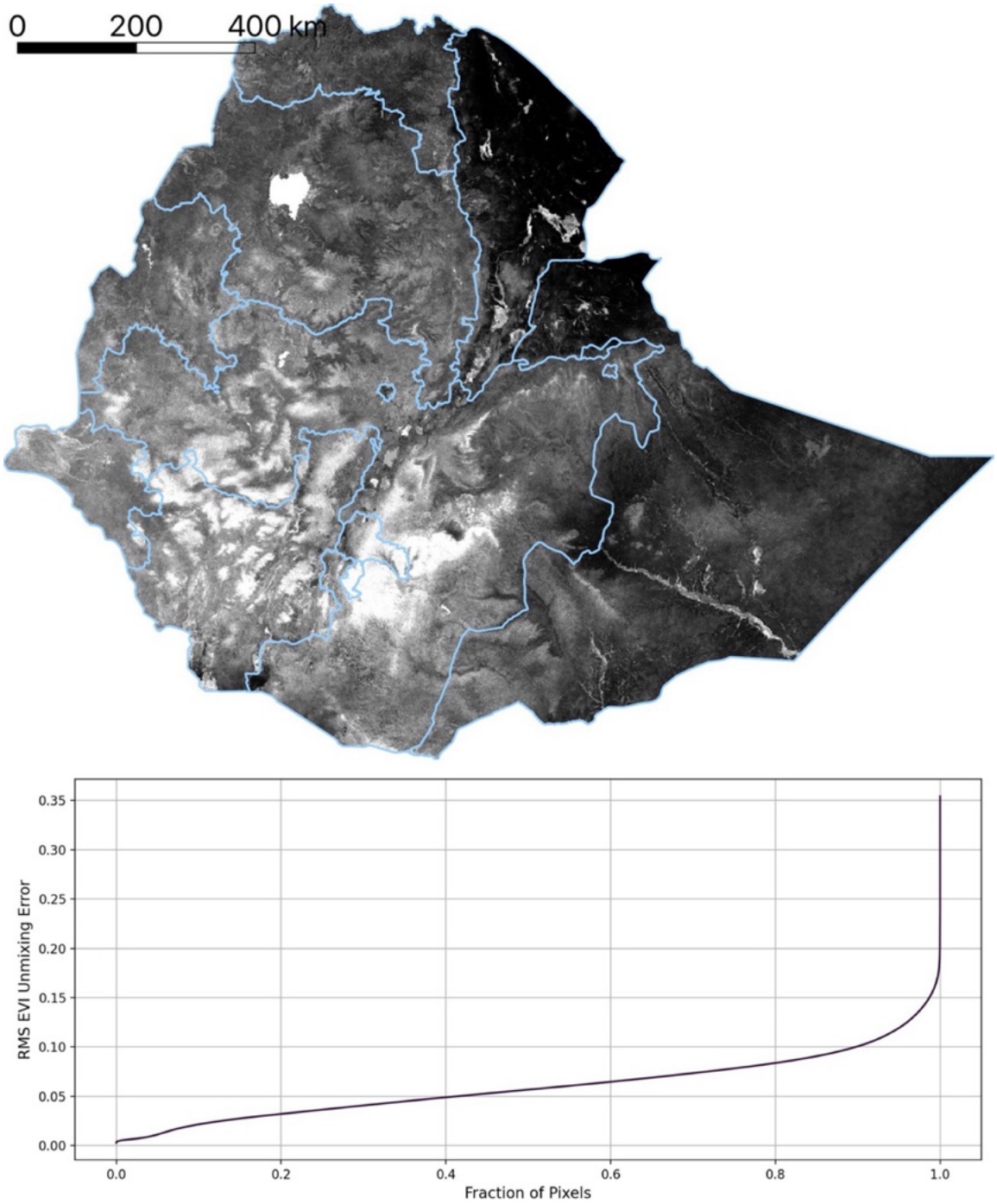

Figure V-3: Root mean square (RMS) error of the temporal mixture model inversion for Figure V-1. A 2% linear stretch is applied to the error map, with administrative boundaries outlined in light blue. The displayed cumulative density function shows the distribution of RMS errors for an increasing fraction of pixels in Figure V-1.

3. Materials and Methods

The data collection portion of this manuscript's methodology consists of pairing Sentinel-2 imagery with labeled polygons to train an irrigation detector. Here, a pixel timeseries paired with a binary irrigation/non-irrigation label constitutes a sample. Irrigation is defined as such: A pixel is irrigated if its phenology includes at least one non-perennial vegetation cycle during the dry season, December 1st to April 1st for the Ethiopian Highlands. Conversely, a pixel is non-irrigated if its phenology demonstrates only vegetation growth that can be attributed to the area's known rainy seasons. Irrigated areas are only of interest if they contain dry season vegetation cycles; this strict definition of irrigation excludes supplemental irrigation practices and perennial crops that may be consistently irrigated throughout the year.

3.1 Sentinel-2 imagery collection

The following analysis uses bottom-of-atmosphere corrected (processing level L2A) Sentinel-2 temporal stacks – four dimensional arrays created by stacking a set spatial extent of imagery bands over multiple timesteps – using the Descartes Labs (DL) platform, a commercial environment for planet-scale geospatial analysis. Images are collected at a 10-day time resolution. To focus on the 2020 and 2021 dry seasons, the time period of interest is defined as between June 1st, 2019, and June 1st, 2021. Given the 10-day timestep, 72 image mosaics are collected – 36 per year.

Using the Descartes Labs platform, imagery mosaics are generating by collecting all Sentinel-2 imagery available within a 10-day timestep that come from 100km-by-100km granules with less than 10% aggregate cloud cover. These images are then sorted by cloud cover and masked using the cloud masks provided by the Sen2Cor algorithm [292]. Given the 5-day revisit

period of Sentinel-2 near the equator, a 10-day timestep ensures that there are two separate satellite passes per image mosaic.

Once imagery is collected for each timestep, values are assigned to individual pixels, pulling first from the image with the lowest amount of cloud cover. If there are masked pixels in this image, pixel values are determined for these locations using the image with the next lowest amount of cloud cover; this process repeats until either all the images available at a timestep are cycled through or each pixel in the 10-day mosaic is filled with valid, non-clouded values. For mosaics that retain invalid pixels due to persistent cloud cover across the timestep (often during the rainy season in Ethiopia, which stretches from June to September), pixel values are assigned via temporal interpolation: Each invalid pixel is given a linearly interpolated value based on the nearest preceding and subsequent image mosaic with a non-clouded value for that pixel.

Image mosaicking is performed bandwise. All 10m and 20m Sentinel-2 bands are extracted (10 bands in total); the 60m coastal aerosol and water vapor bands are ignored, as these bands contain atmospheric information not relevant for the land process monitoring task at hand. The final image processing step involves temporal smoothing of all timeseries using a 3rd order polynomial Savitzky-Golay filter with a window length of 5.

To assist with temporal interpolation of clouded pixels at the start and end of the specified time period, 82 image mosaics are collected in total – the 72 image mosaics that make up the 2 years of imagery (June 1, 2019 – June 1, 2020, and June 1, 2020 – June 1, 2021), plus 5 additional timesteps before and after the beginning and end of full time period. After interpolation and smoothing, these additional image mosaics are discarded to leave cloud-free, smoothed, 10 band

Sentinel-2 imagery for only the desired 72 timesteps. The imagery is then split into annual temporal stacks, with all training and inference done on a single year's 36 timesteps of imagery.

3.2 Label collection

Two types of labeled data are leveraged for irrigation mapping: *ground collection* (GC) labels, acquired via an in-person survey; and *visual collection* (VC) labels, acquired via visual identification of dry season vegetation from Sentinel-2 imagery using the DL platform and subsequent cleaning via timeseries clustering. The locations of these GC and VC regions are shown in italics in Figure V-1, with all labels collected for the 2021 dry season. As the GC labels constitute our highest quality irrigation observations, verified by in-situ visits to individual plots, we do not use them for training during the model sensitivity analysis, instead reserving them for validation of classifier performance.

3.2.1 Ground collection

The ground collection survey was conducted during the months of March and April 2021. Enumerators collected labels across an area north and east of Lake Tana (referred to as "Tana"; see Figure V-1) in a process that involved traveling to individual plots of lands, collecting four coordinate points corresponding to the corners of the plot, and specifying whether irrigation was present on the plot during the visit. The ground collection survey team collected 2002 polygons in Tana: 1500 were labeled non-irrigated and 502 were labeled irrigated. In total, these polygons cover 1867 Ha, 78% of which was designated as non-irrigated.

3.2.2 Visual label collection

To supplement the GC labels located in Tana, visually collected labels are acquired for seven separate regions via a three-step process of 1) visual inspection, 2) EVI timeseries confirmation, and 3) cluster cleaning. Each of these steps is described in its eponymous subsection below.

Visual inspection

The first step in the VC labeling process involves drawing polygons around locations that either: a) present as cropland with visible vegetation growth (for the collection of irrigated samples), or b) present as cropland with no visible vegetation growth (for the collection of non-irrigated samples), based on dry-season, false-color Sentinel-2 imagery presented on the DL platform. Sub-meter resolution commercial satellite imagery from Google Earth Pro is also used to confirm the existence of cropland in the viewing window. For the collection of non-irrigated labels, polygons are restricted to areas that contain non-perennial cropland; however, because only phenologies that contain dry season vegetation cycles are considered irrigated, non-irrigated polygons occasionally overlap other types of land cover – e.g., perennial crops, fallow cropland, or areas with human settlement – with any overlap likely to improve training robustness.

EVI timeseries confirmation

After drawing a polygon around a suspected irrigated or non-irrigated area, the second step in the VC label acquisition process entails inspection of the median Sentinel-2 EVI timeseries of all pixels contained within the polygon; this step is shown in the plot windows of Figure V-4.

Here, all available Sentinel-2 imagery with less than 20% cloud cover between June 1, 2020, and June 1, 2021 is retrieved; a cubic spline is then fit to all available data to generate continuous EVI timeseries. For potential irrigated polygons, if the EVI timeseries shows a clear peak above 0.2 during the dry season, it is confirmed as irrigated. Similarly, for potential non-irrigated polygons, an EVI timeseries that demonstrates a single vegetation cycle attributable to Ethiopia's June to September rains is taken as confirmation of a non-irrigated VC polygon. However, if the EVI timeseries does not confirm the expected irrigated/non-irrigated class, or if the plotted EVI error bars (representing $\pm$one standard deviation of the EVI values at that timestep) indicate a level of signal noise within the polygon that prevents the identification of a clear vegetation phenology, the polygon is discarded.

Figure V-4(a) demonstrates an example of irrigated VC label collection in the Koga region – here, the double vegetation peak present in the EVI timeseries confirms the purple polygon in the center of the window as irrigated (blue polygons indicate areas already saved as irrigated VC labels). Figure V-4(b) demonstrates the same process for non-irrigated VC labels, also in Koga: The single EVI peak in October 2020 confirms the pink polygon in the top left of the window as non-irrigated (red polygons indicate areas already saved as non-irrigated VC labels).

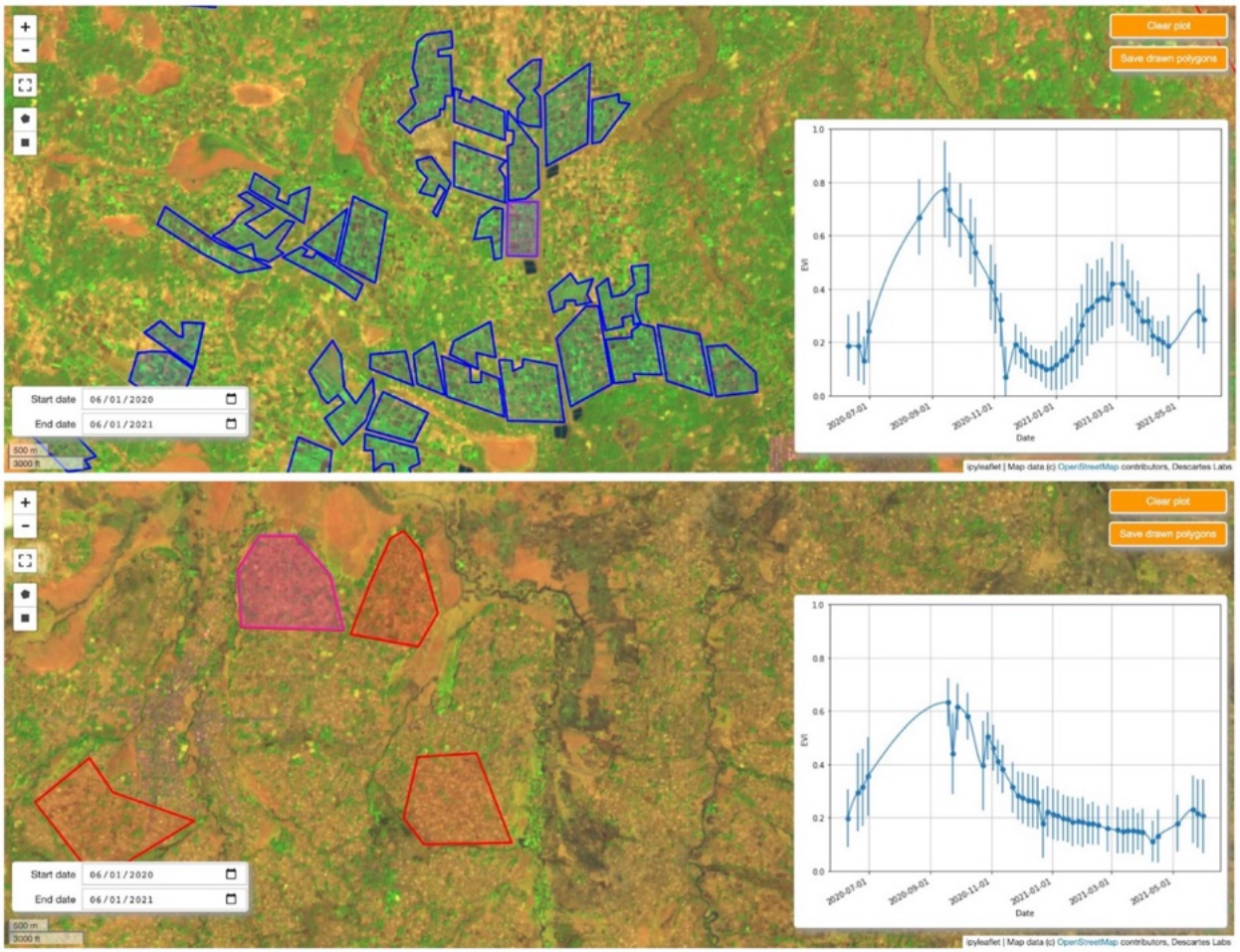

Figure V-4: Example of the visual collection (VC) labeling process in Koga using the Descartes Labs platform. Blue polygons denote areas determined to be irrigated; red polygons are determined to be non-irrigated. Background imagery is a false-color Sentinel-2 image taken in March 2021: red, near-infrared, and blue bands are presented in the RGB channels, respectively. In (a), the Sentinel-2 enhanced vegetation index (EVI) timeseries is shown for the drawn purple rectangle in the middle of the window; in (b), the Sentinel-2 EVI timeseries is shown for the drawn pink, semi-octagonal polygon in the top left of the window. Both timeseries present the median EVI values for all pixels contained within the drawn polygon; the error bars show one standard deviation of these values above and below the median. In both figures, the drawn polygons are confirmed as VC labels, since they meet the definitions of irrigation/non-irrigation, respectively.

Cluster cleaning

The third step in the VC label acquisition process involves bulk verification of the collected timeseries by means of cluster cleaning. For each VC region, all pixels that reside within labeled

polygons are collected and split based on the irrigated/non-irrigated class labels of the polygons. Fifteen-component Gaussian mixture models are fit to each class's data to extract the dominant phenologies contained within the region's samples; the EVI timeseries representing the cluster centroids are then plotted, with the plot legend displaying the number of samples per cluster. Figure V-5(a) presents the results of this initial clustering for the Koga region.

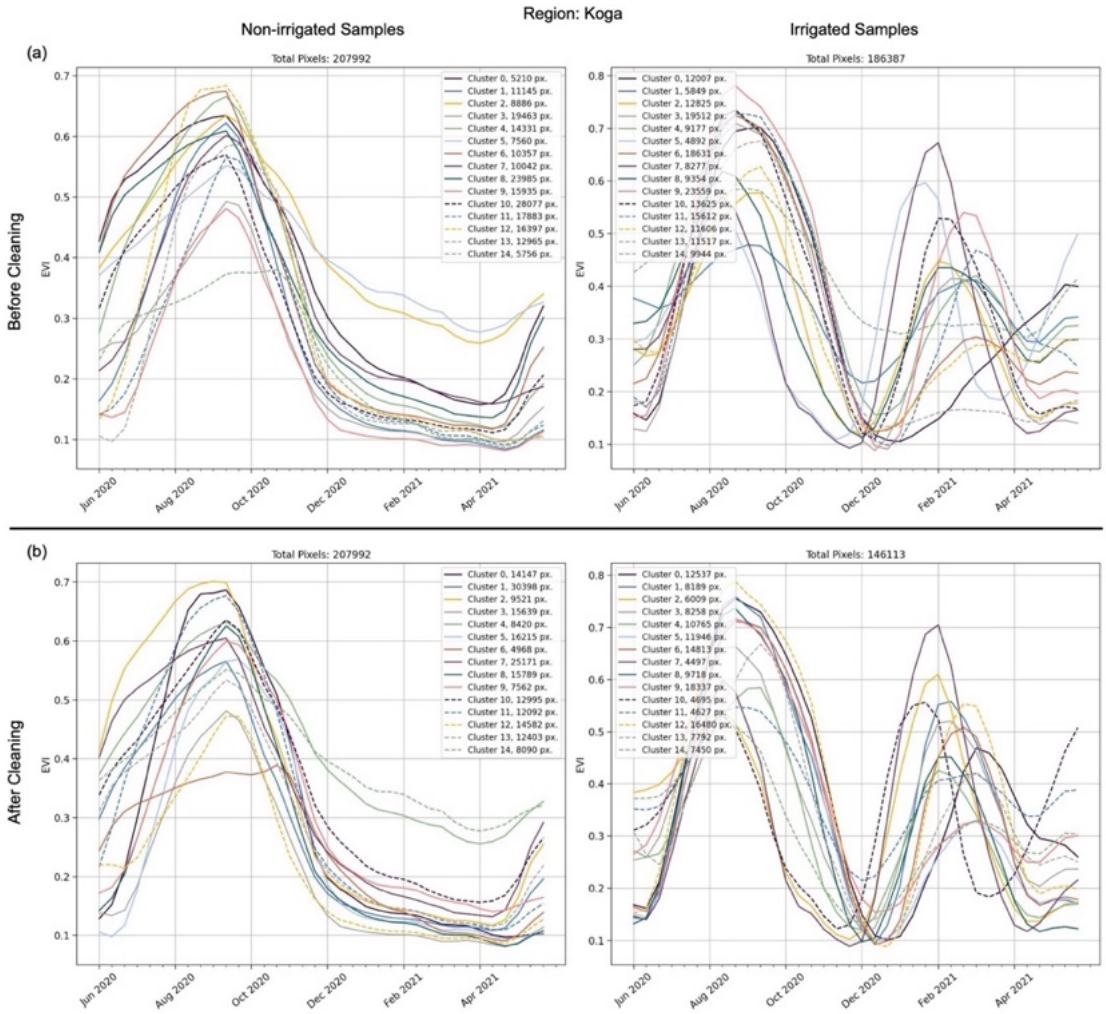

Figure V-5: Clustered enhanced vegetation index (EVI) timeseries before and after cluster cleaning for the Koga visual collection (VC) region. Before and after cleaning, pixels are grouped into one of 15 randomly indexed clusters. In (a), Clusters 12, 13, and 14 of the irrigated samples do not achieve an EVI peak of 0.2 during the dry season (December 1st to April 1st) or do not contain multiple successive values below 0.2 and are discarded. All non-irrigated clusters display a single vegetation peak aligned with the main rainy season, and the irrigated clusters after cleaning (b) all display a vegetation cycle during the dry season.

From the initial cluster timeseries, an iterative process begins to ensure that all cluster timeseries align with the specified class label. For an irrigated cluster timeseries to be kept, it must contain multiple successive EVI values above and below 0.2, and it must contain a clear EVI peak above 0.2 during the dry season. Analogously, non-irrigated cluster timeseries are discarded if they display a clear dry-season EVI peak above 0.2. If these conditions are not met – as is the case for Clusters 12, 13, and 14 of the Koga irrigated samples, which do not contain a clear EVI peak above 0.2 between December 1, 2020 and April 1, 2021 and/or do not senesce below an EVI threshold of 0.2 for successive timesteps – all pixels associated with that cluster are discarded from the labeled data. This process is repeated until all 15 clusters for both classes demonstrate EVI signals that meet the non-irrigated/irrigated class definitions. The final, cleaned cluster timeseries for the Koga region are shown in Figure V-5(b).

Cluster-cleaning is performed for all regions' labeled data, including labeled data collected from the GC region, Tana. For increased visibility into the labeled data collected and used for training, these regions' clusters before and after cleaning are included in Appendix A in Section V-6.

Table V-1: Summary of labeled polygons, split by region and model training configuration. GC stands for ground collection labels; VC stands for visual collection labels.

		Number of Labeled Polygons							
Region	Type of Labels	Training		Validation		Testing		Total	
		Non-Irrigated	Irrigated	Non-Irrigated	Irrigated	Non-Irrigated	Irrigated	Non-Irrigated	Irrigated
Tana	GC	1050	351	225	76	225	75	1500	502
Rift	VC	12	25	3	6	3	6	18	37
Koga	VC	27	46	6	10	6	10	39	66
Kobo	VC	26	28	6	6	6	7	38	41
Alamata	VC	17	16	4	4	4	4	25	24
Liben	VC	24	25	5	5	6	6	35	36
Jiga	VC	15	13	4	3	3	3	22	19
Motta	VC	17	17	4	4	4	4	25	25
Total	GC + VC	1188	521	257	114	257	115	1702	750

A summary of the number of collected polygons and cleaned pixel timeseries samples in each region is shown in Tables V-1 and V-2: In total, 1,207,233 non-irrigated samples and 907,887 irrigated samples are used, taken from 1702 and 750 labeled polygons, respectively. For model training and evaluation, data are divided among training, validation, and test splits[xx]. Here, polygons in each labeled region are split according to a 70/15/15 training/validation/test ratio; this method ensures that highly similar pixels from within the same polygon do not exist across training configurations, a division of data that would artificially inflate model performance for

[xx] In splitting the labeled data, the training/validation/testing terminology standard in machine and deep learning literature is adopted.

the task of predicting irrigation over pixel timeseries unseen by the model. All training, validation, and testing is performed pixelwise (i.e., having removed the spatial relationships of samples).

Table V-2: Summary of labeled polygons, split by region and model training configuration. GC stands for ground collection labels; VC stands for visual collection labels.

		Number of Labeled Pixels							
Region	Type of Labels	Training		Validation		Testing		Total	
		Non-Irrigated	Irrigated	Non-Irrigated	Irrigated	Non-Irrigated	Irrigated	Non-Irrigated	Irrigated
Tana	GC	63,729	24,675	14,283	5089	13,910	5361	91,922	35,125
Rift	VC	92,157	104,682	19,149	19,269	20,378	20,286	131,684	144,237
Koga	VC	150,378	98,697	29,661	23,015	27,953	24,401	207,992	146,113
Kobo	VC	93,838	123,946	30,549	36,494	31,473	48,077	155,860	208,517
Alamata	VC	58,310	21,176	14,356	4601	11,083	6447	83,749	32,224
Liben	VC	132,999	113,733	26,027	31,212	35,394	21,895	194,420	166,840
Jiga	VC	113,640	79,143	33,244	15,368	38,734	12,204	185,618	106,715
Motta	VC	94,153	47,915	34,267	11,127	27,568	9074	155,988	68,116
Total	GC + VC	799,204	613,967	201,536	146,175	206,493	147,745	1,207,233	907,887

3.2.3 Labeled data exploration

To better understand the vegetation phenologies contained within this study's labeled data, the similarities of EVI timeseries of the same class are explored across regions. This process first involves applying a PC transform to all labeled training data. The samples' dimensionality is then reduced by using only the first 10 dimensions of the transformed data; these first 10 dimensions explain 91% of the variance contained within the samples' EVI timeseries.

After dimensionality reduction via the PC transform, the two sample Kolmogorov-Smirnov (KS) test statistic is calculated between sample distributions of the same class across regions. The two-sample KS statistic determines the largest absolute distance between two 1D empirical distributions, and is presented in Eq. (V-1):

$$D_{KS} = sup_x|F_1(x) - F_2(x)| \tag{V-1}$$

where $F_1(x)$ and $F_2(x)$ are the two empirical distribution functions of 1D variable *x*, and *sup* is the supremum function. The KS statistic is assessed for two reasons: 1) the statistic depends on no assumptions about the underlying data distributions; and 2) the statistic has been adapted for multivariate distributions via the pseudo-1D KS metric [293]. In this adaptation, the pseudo-1D KS metric, $D_{KS,P1D}$, is the Euclidean KS statistic calculated between successive orthogonal dimensions of two multivariate distributions:

$$D_{KS,P1D} = \sqrt{(D_{KS,1})^2 + (D_{KS,2})^2 + \cdots + (D_{KS,n})^2} \tag{V-2}$$

where

$$D_{KS,n} = sup_{y_n}|F_1(y_n) - F_2(y_n)| \tag{V-3}$$

Here, Eq. (V-3) represents the KS statistic between the empirical distribution functions F_1 and F_2 of the *n*th dimension of multivariate variable *y*. As only the first 10 PC dimensions of the transformed data are used, *n* ranges between 1 and 10.

Table V-3 presents pairwise pseuso-1D KS statistics between regions' non-irrigated samples; Table V-4 presents pairwise pseuso-1D KS distances between regions' irrigated samples.

In these tables, the relative statistics between distributions are compared, as the absolute statistics cannot be interpreted in a physically meaningful way. The cells with typographical marks in the two tables indicate statistics to be interpreted with the results in Table V-7, discussed alongside that table. Tables V-3 and V-4 show that the relative pairwise statistic between regional distributions is larger among the irrigated sample sets, indicating that irrigated samples are more dissimilar across regions compared to the non-irrigated samples. This takeaway reflects the varying nature of irrigation practices across Ethiopia – irrigation can occur at different parts of the dry season for a variety of different crops. In contrast, the phenologies of non-irrigated cropland must mirror Ethiopia's primary rains, which are consistent in time for the regions included in this analysis.

Table V-3: Pairwise pseudo-1D KS metric between regions' non-irrigated training samples. Values with typographical symbols are to be interpreted alongside Table V-7.

	Tana	Rift	Koga	Kobo	Alamata	Liben	Jiga	Motta	Mean
Tana	0.00	1.13	1.85	1.00	1.25	1.69	1.50	1.33	1.39
Rift	1.13	0.00	1.48	0.62	0.71	1.43	0.95	0.57	0.98
Koga	1.85*	1.48	0.00	1.48	1.33	0.41	1.02	1.26	1.26
Kobo	1.00	0.62	1.48	0.00	0.76	1.45	1.16	0.91	1.05
Alamata	1.25	0.71	1.33	0.76	0.00	1.33	0.97	0.72	1.01†
Liben	1.69	1.43	0.41	1.45	1.33	0.00	0.91	1.23	1.21
Jiga	1.50**	0.95	1.02	1.16	0.97	0.91	0.00	0.51	1.00
Motta	1.33	0.57	1.26	0.91	0.72	1.23	0.51	0.00	0.93‡
									1.11

Table V-4: Pairwise pseudo-1D KS metric between regions' irrigated training samples. Values with typographical symbols are to be interpreted alongside Table V-7.

	Tana	Rift	Koga	Kobo	Alamata	Liben	Jiga	Motta	Mean
Tana	0.00	1.88	2.25	1.73	1.65	1.13	1.60	1.27	1.64
Rift	1.88	0.00	1.48	0.61	0.37	2.01	1.30	1.51	1.31
Koga	2.25*	1.48	0.00	1.37	1.45	2.46	1.98	2.20	1.88
Kobo	1.73	0.61	1.37	0.00	0.65	1.98	1.51	1.57	1.35
Alamata	1.65	0.37	1.45	0.65	0.00	1.78	1.11	1.31	1.19†
Liben	1.13	2.01	2.46	1.98	1.78	0.00	1.41	0.98	1.68
Jiga	1.60**	1.30	1.98	1.51	1.11	1.41	0.00	0.89	1.40
Motta	1.27	1.51	2.20	1.57	1.31	0.98	0.89	0.00	1.39‡
									1.48

3.3 Prediction admissibility criteria

Given that irrigated phenologies exist over a small fraction of the total land area of the Ethiopian Highlands, and that there are many types of land cover that do not fall within this manuscript's non-irrigated/irrigated cropland dichotomy, a set of criteria are imposed to exclude pixel phenologies that are not cropland or are highly unlikely to be irrigated. Table V-5 presents five criteria that must all be met for a pixel timeseries to be potentially irrigated and the motivation behind each.

The criteria in Table V-5 are also used to create a reference irrigation classifier that does not rely on machine learning. For this reference classifier, if all 5 conditions are met, the sample is deemed irrigated; if any of the conditions is not satisfied, the sample is deemed non-irrigated.

Table V-5: Prediction admissibility criteria. All criteria need to be satisfied for a prediction to be admitted as irrigated.

Admissibility Criteria	Motivation
10th percentile of EVI timeseries < 0.2	Filter out evergreen pixels
90th percentile of EVI timeseries > 0.2	Filter out barren/non-vegetated pixels
Maximum of the EVI timeseries during the dry season (Dec 1 – Apr 1) > 0.2	Filter out pixels with no vegetation growth in the dry season
Ratio of the 90th:10th percentile of the EVI timeseries > 2	Filter out evergreen pixels
Shuttle Radar Topography Mission slope measurement < 8%	Filter out highly sloped settings where cropping is impractical

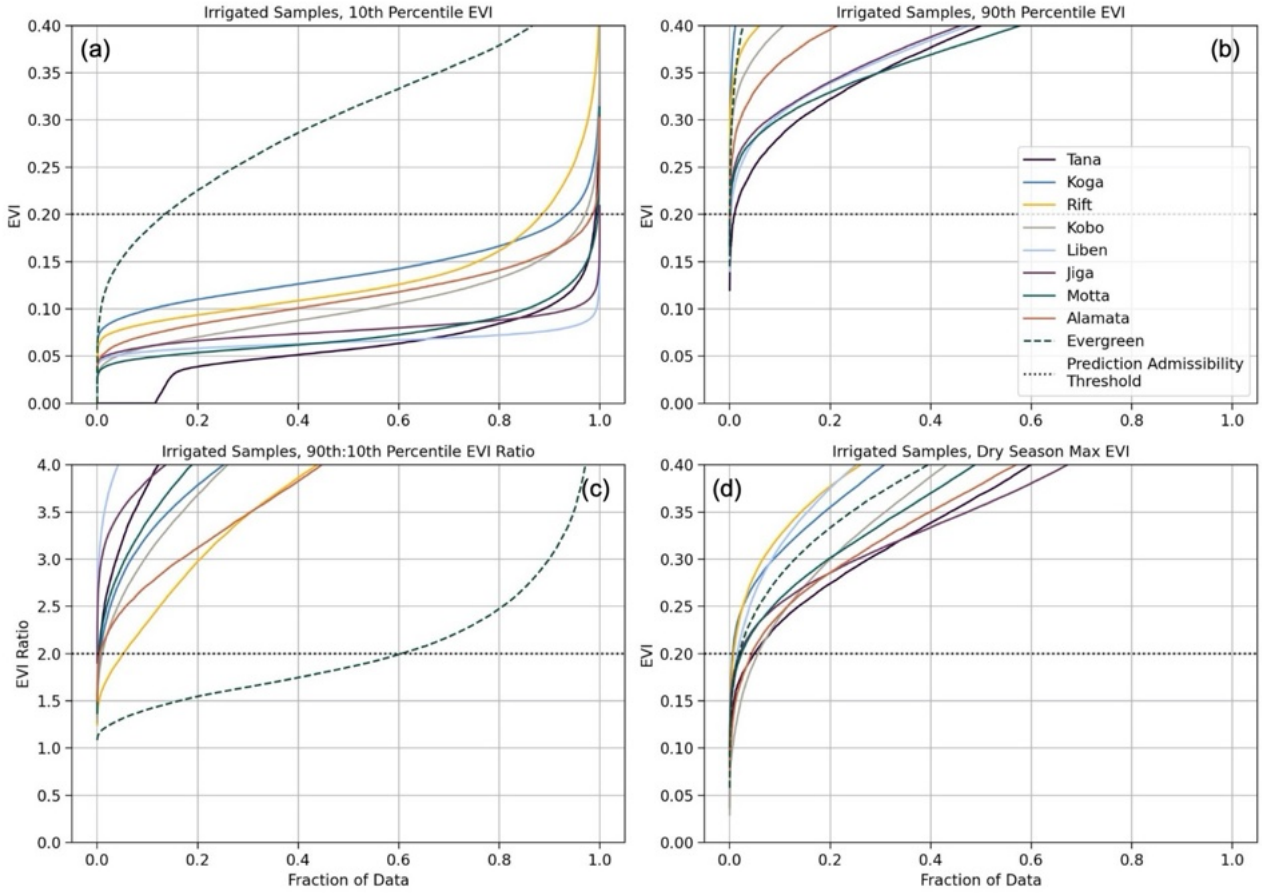

Figure V-6: Cumulative distribution functions (CDFs) for the (a) 10th and (b) 90th EVI timeseries percentiles; (c) the 90th:10th EVI timeseries percentile ratio; and (d) the maximum EVI value during the dry season (December 1st, 2020, to April 1st , 2020).

The vegetation-specific criteria in Table V-5 are informed by the EVI distributions of labeled irrigated samples for all label collection regions: Figure V-6 contains cumulative distribution functions (CDFs) for the 10^{th} and 90^{th} EVI timeseries percentiles, the 90^{th}:10^{th} EVI timeseries percentile ratio, and the maximum EVI value during the dry season. CDFs are presented for all regions' irrigated samples, including for a set of polygons collected over evergreen land cover areas.

Figure V-6(a) shows that a maximum of 0.2 for the 10^{th} percentile of the EVI timeseries is achieved by nearly all irrigated samples, and how this ceiling filters out 85% of all evergreen samples. Similarly, a minimum 90^{th}:10^{th} percentile EVI ratio of 2 is satisfied by nearly all irrigated samples and excludes 60% of evergreen samples (Figure V-6(c)). While no EVI timeseries for barren or non-vegetated areas are shown in this figure, the criteria specifying a 90^{th} percentile EVI value above 0.2 and a dry season max EVI value above 0.2 are met by the vast majority of irrigated samples (Figure V-6(b,d)), and would filter out many of these non-cropped pixels.

3.4 Model training

3.4.1 Model architectures

Five separate classifier types are compared to determine the model architecture with the most robust irrigation detection performance across regions. The first two classifiers are decision tree-based: A random forest with 1000 trees [294]; and a CatBoost model that uses gradient boosting on up to 1000 trees [295]. The other three classifiers are neural networks (NN): A baseline network, an LSTM-based network, and a transformer-based network. For comparability, these three classifier architectures are designed to have similar structures, based on the strong

baseline model structure proposed in [296]; as seen in Figure V-7, they differ only in the type of encoding blocks used.

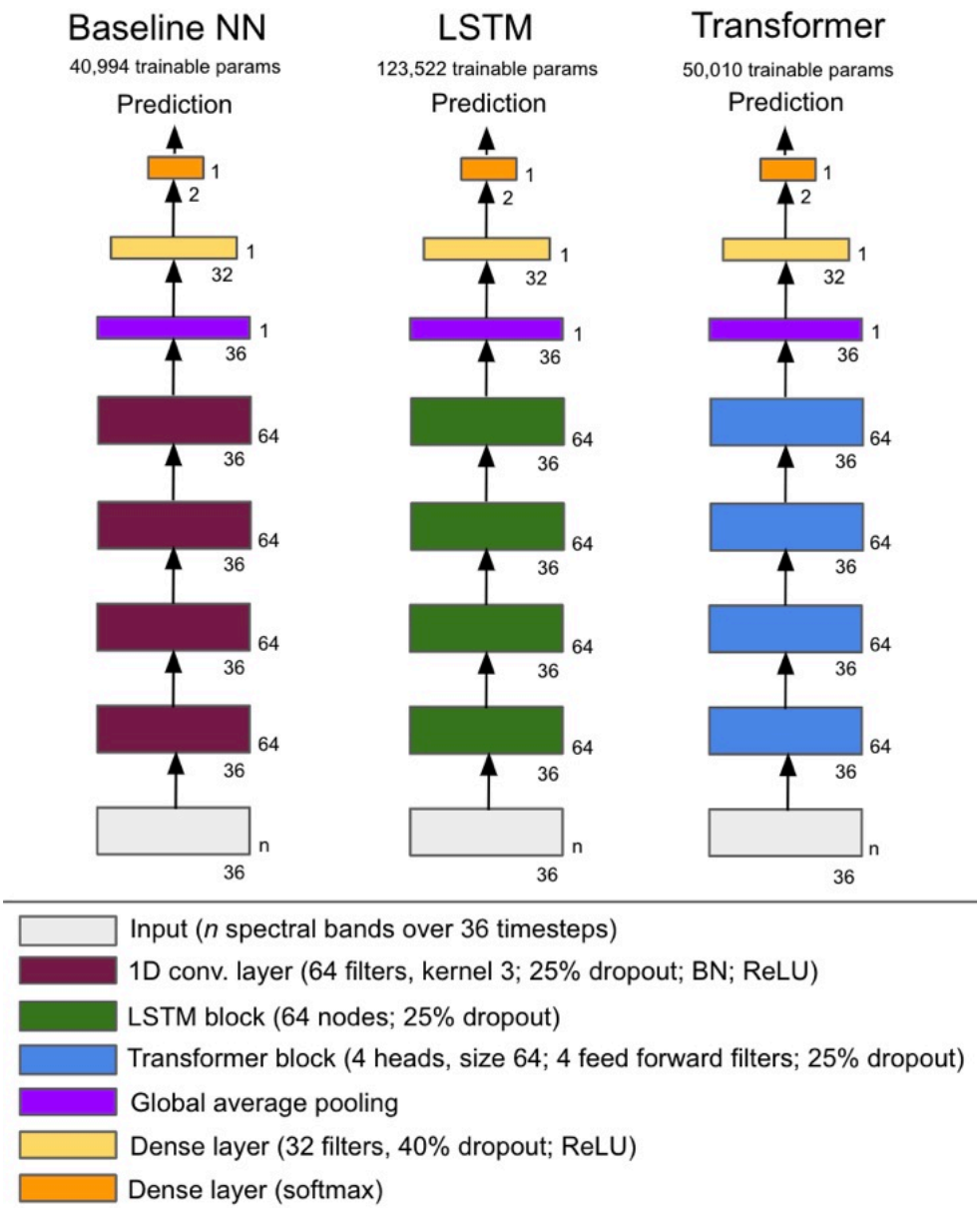

Figure V-7: Neural network (NN) model architectures tested as irrigation detection classifiers. Model architectures are consistent by design; only encoding blocks differ across networks.

3.4.2 Model training strategies

The implemented model training strategy addresses two potential pitfalls among training processes: 1) imbalanced samples across region and class; and 2) high similarity among samples within a region that may not reflect the sample distributions across all regions. Consistent with best practices in dealing with imbalanced data, this first issue is addressed with a) class balancing weights specific to each region, based on the "balanced" heuristic inspired by [297]; and b) a region-specific weight equal to the ratio of the maximum number of samples in any region to the number of samples for the region in question. Both class-balancing and region-balancing weights are used in all training configurations.

To address potential redundancy and time-specificity among samples within a region, random shifts are applied to all input timeseries. The sizes of these random shifts vary between -3 and +3 timesteps (corresponding to between -30 and +30 days), with an equal probability of each shift occurring (including a shift by 0 timesteps). Random shifts are applied to all samples in the training and validation sets and differ for each sample every time it's seen by the model. No shifts are applied to the samples in the testing sets.

The primary metric for performance evaluation is the F_1 score on the test datasets of regions withheld from training. Accordingly, performance is assessed in a manner that prioritizes classifier robustness – i.e., performance in regions unseen during training – and not in a manner that could be inflated by close similarity of samples within a region. For reference, the F_1 score balances prediction precision and recall, and is calculated per Eq. (V-4).

$$F_1 = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

(V-4)

where *TP* indicated true positive predictions, *FP* indicates false positive predictions, and *FN* indicates false negative predictions.

The training strategy differs for the tree-based classifiers and for the neural network-based classifiers. As training the tree-based classifiers occurs across a single batch with no iteration across epochs, there is no need for separate validation and testing datasets: The training and validation datasets of all included regions are therefore combined to create a single training dataset. After training on this combined dataset, performance is evaluated across the test datasets.

In contrast, training neural network-based models takes place by batch across epochs, and a validation set is required to guide the training process. For a given training step, one batch from each region is concatenated, with the combined output shuffled before model intake. After the epoch is finished, performance is assessed on the validation set of each region included in training. If the minimum F_1 score among all regions' validation sets has increased from its previous maximum, the model weights are saved; however, if the minimum F_1 score has not increased from its previous high point, the model weights are discarded. Minimum F_1 score across all validation regions is selected as the weight update criteria to ensure model robustness: Consistent performance across the entire area of interest is desired, not high performance in one set of regions and poor performance in another. Training concludes once the minimum validation region F_1 score has not improved for 10 training epochs, or after 30 epochs have been completed. After training, model weights are loaded from the epoch with the highest minimum validation region F_1 score; performance of this model on the test datasets of all regions is then reported.

For all training runs, a binary cross-entropy loss, a learning rate of 1e-4, and an Adam optimizer [298] are specified. Inputs are standardized to a mean of 0 and standard deviation of 1 using statistics from the entire set of labeled samples.

4. Results

4.1 Model sensitivity

Figure V-8 presents withheld VC region test dataset F_1 scores for three different types of model input – one that includes all spectral bands for all timesteps; one that includes only the EVI layer for all timesteps; and one that includes only the EVI layer for all timesteps with the random sample shift applied. Here, the performance of models trained on all combinations of VC regions is evaluated; these results are organized along the *x*-axis by the number of VC regions included during training. Each *x*-axis tick label also includes in parentheses the number of withheld VC region test dataset evaluations for all models trained on *x* included VC regions. All results are presented for the transformer model architecture; however, these findings are agnostic to the classifier architecture selected.

Figure V-8 demonstrates that models trained on samples containing only EVI timeseries outperform those that include all spectral bands at all timesteps, both on average (a) and in low performing regions (b). The 10^{th} percentile of withheld regions' F_1 scores is shown in order to understand the low-end of model performance without accounting for outliers. For reference, classifier performance based on the prediction admissibility criteria is also included. Figure V-8 shows that explicitly feeding classification models information about samples' vegetation content – i.e., feature engineering – allows for better performance compared to models that

intake 10 Sentinel-2 spectral bands. Introducing a random temporal shift to the EVI timeseries further increases performance; by increasing the sample variance seen by the model, randomly shifting the input timeseries improves model transferability.

Taken together, randomly shifted EVI timeseries increase withheld region F_1 scores by an average of 0.22 when only 2 VC regions are included in the training data, compared to models that use all spectral bands. As performance begins to plateau with 4 or more VC regions included in the training data, this gap shrinks to an improvement of 0.10. Similar results can be seen in Figure V-8(b) for the low-end of performance: Extracting and randomly shifting EVI timeseries increase the 10^{th} percentile of withheld region F_1 scores by 0.40 when 2 VC regions are included in the training data, a difference that shrinks to approximately 0.14 with 5 or more VC regions in the training data. Two additional findings are gleaned from the results for the models trained on randomly shifted EVI timeseries (i.e., the grey curve). First, a classifier trained on data from 2 VC regions or more outperforms the pixel filtering baseline. Second, increasing the number of VC regions included in the training set improves withheld region prediction performance up until 4 VC regions before tapering off.

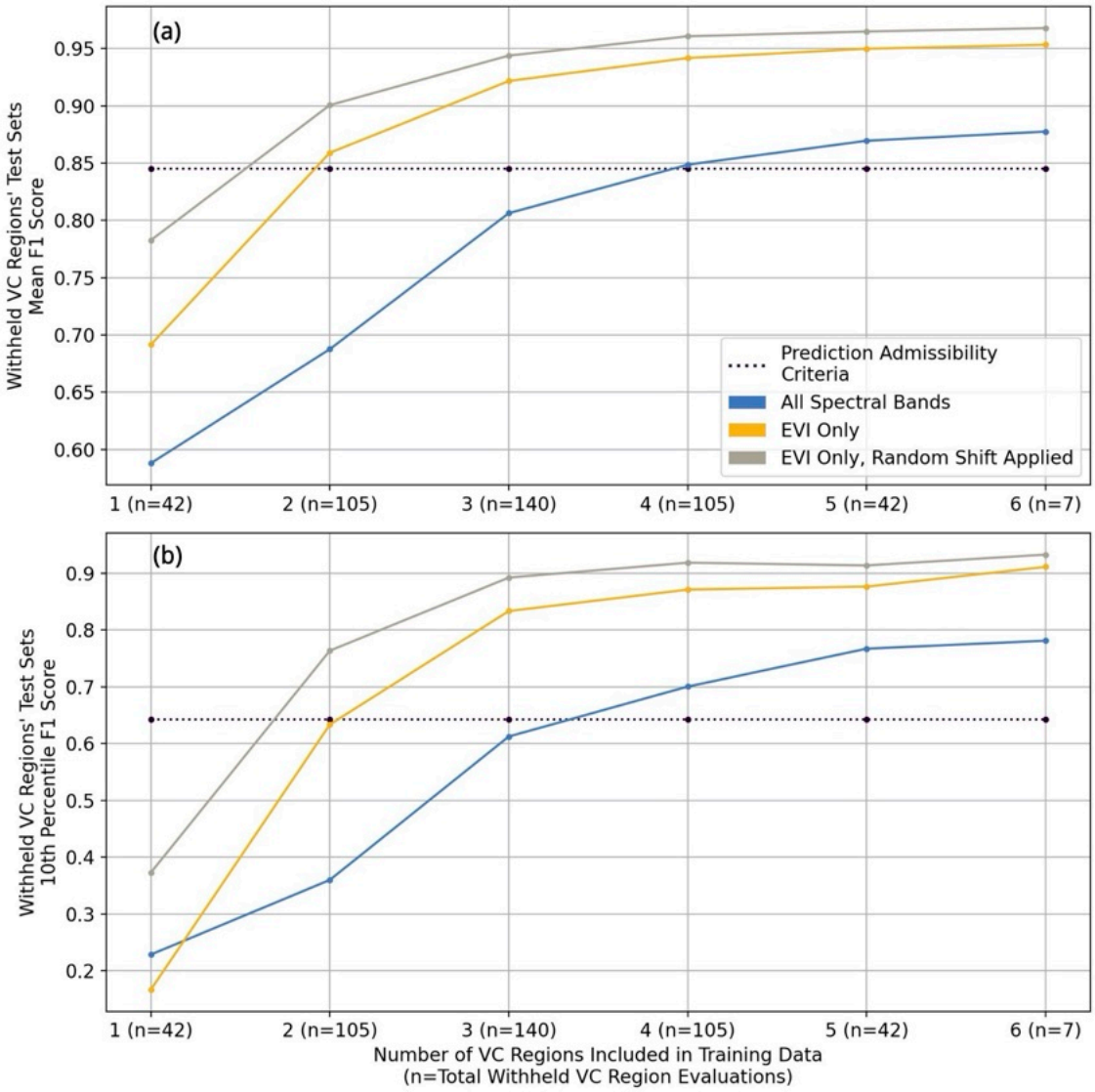

Figure V-8: Withheld region test dataset performance for different types of model input, organized along the x-axis by the number of regions included during training. (a) presents mean F_1 score over the withheld regions; (b) presents the 10^{th} percentile F_1 score over the withheld regions. Results indicate that model inputs of randomly shifted enhanced vegetation index (EVI) timeseries yield the best classifier performance. F_1 scores from classification based on the prediction admissibility criteria are presented for reference.

Implementing a modified Gradient-Class Activation Map (Grad-CAM) for visual prediction explanation provides further evidence for improved prediction robustness from randomly shifting input EVI timeseries. A Grad-CAM uses the gradients flowing into the final layer of a

neural network to produce a localization map highlighting important portions of the input for predicting a concept. Originally developed for images in [299], this technique can be applied analogously to timeseries. To do so, a transformer-based classifier model with its 32-node penultimate dense layer removed is trained on all VC regions' training datasets; by removing this fully connected layer, the importance of each timestep input for prediction can be visualized, as there is no longer a layer obscuring the gradient flow into the final prediction nodes. Figure V-9 displays the normalized timestep prediction importances for 16 randomly selected non-irrigated and irrigated EVI timeseries from the Koga region. Results are presented for two models: the first trained without randomly shifting input timeseries, and the second trained with the random shift applied.

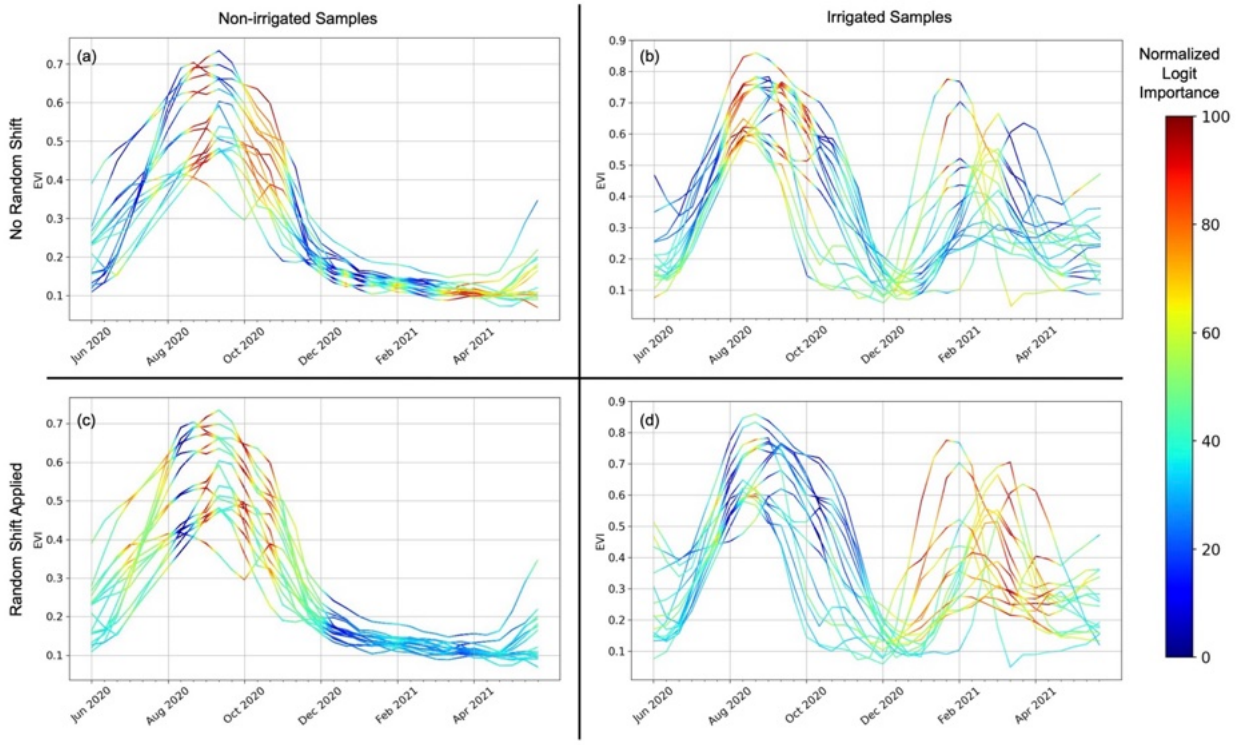

Figure V-9: Modified Grad-CAM timestep importances for 16 randomly selected non-irrigated and irrigated enhanced vegetation index (EVI) timeseries from Koga, before and after the timeseries shift is applied.

Figure V-9 demonstrates that the input timesteps most important for prediction – displayed in red per the Normalized Logit Importance colorbar – are more continuous and better reflect a common understanding of what portions of a phenology should be predictive once the model is trained on randomly shifted input timeseries. For the non-irrigated samples, the model trained on the randomly shifted timeseries identifies a larger portion of the timesteps during the rainy season as highly predictive (Figure V-9(c)); this model also correctly identifies vegetation growth during dry season timesteps as important for identifying irrigated samples (Figure V-9(d)). In comparison, the model trained on the non-shifted timeseries identifies scattered timesteps as predictive for both non-irrigated and irrigated samples (Figure V-9(a,b)); it does not emphasize dry season vegetation growth as predictive of irrigation presence (Figure V-9(b)). Instead, this model learns to identify isolated, non-intuitive timesteps, and as a consequence is more likely to misclassify input timeseries that differ slightly from those in the training data.

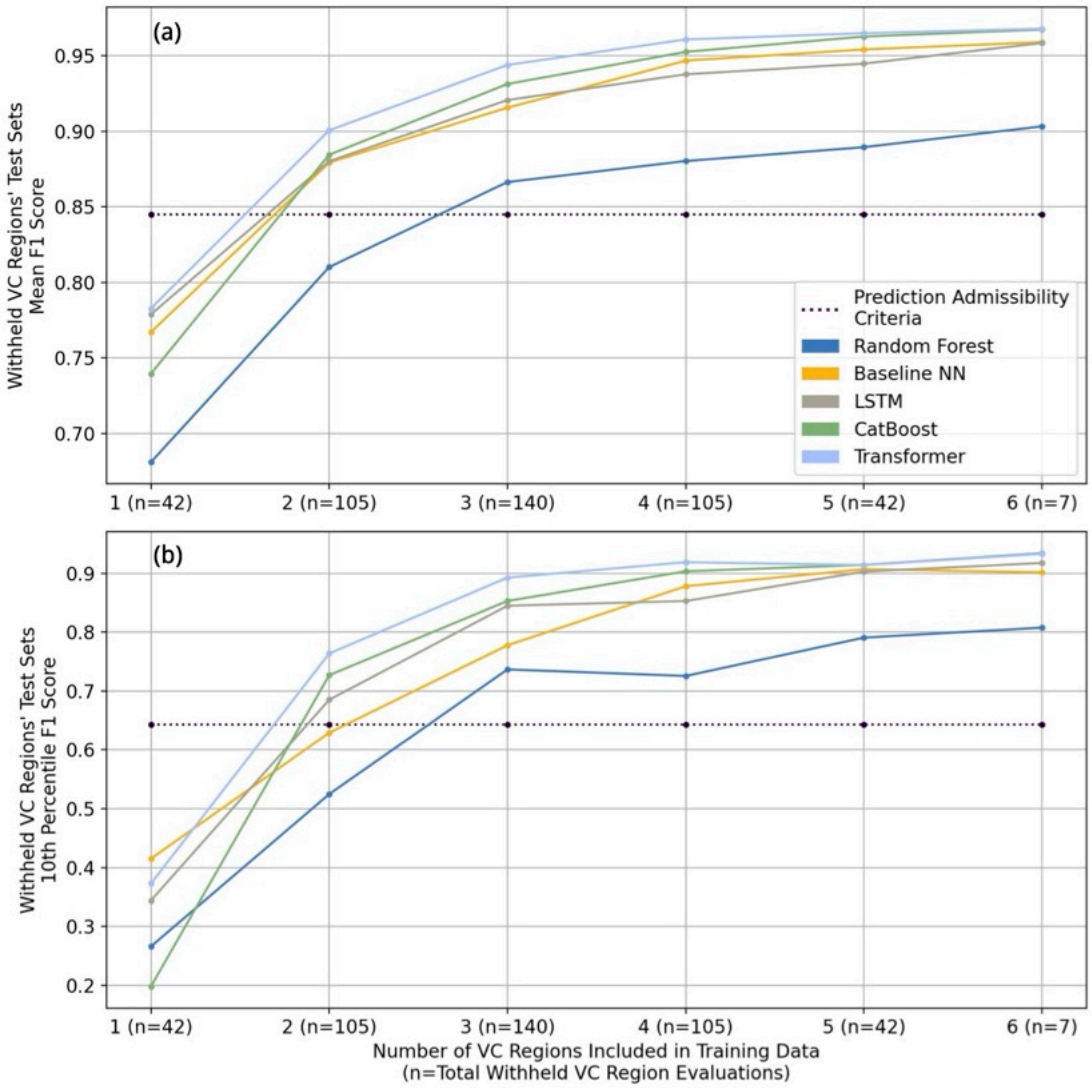

Figure V-10: Withheld region test dataset performance for different classifier models, organized along the x-axis by the number of regions included during training. (a) presents mean F_1 score over the withheld regions; (b) presents the 10^{th} percentile F_1 score over the withheld regions. Results indicate that the transformer-based classifier yields the best performance, followed closely by the CatBoost model. F_1 scores from classification based on the prediction admissibility criteria are presented for reference.

Figure V-10 displays (a) mean and (b) 10^{th} percentile F_1 score for all combinations of VC regions included in training for the 5 classification models tested, along with the reference classifier based on the prediction admissibility criteria. Figure V-10 demonstrates that the

transformer architecture is most robust for all combinations of VC training regions, followed closely by the CatBoost architecture for all training configurations with 2 or more VC regions. Moreover, for models with 5 or 6 VC regions included in training, mean and low-end F_1 scores for these two architectures are practically indistinguishable at 0.97 and 0.92, respectively. Lastly, Figure V-10 shows that the LSTM architecture does not noticeably improve performance compared to the baseline neural network, and that the trained Random Forest models yield the worst performance in withheld regions.

Table V-6: Comparison between Transformer and CatBoost model predictions for models trained on all 7 visual collection (VC) regions' training datasets.

Region	Type of Labels	Comparison between Transformer and CatBoost model predictions				
		Number of aligned sample predictions		Number of misaligned sample predictions		Fraction of predictions aligned across models
		Non-irrigated	Irrigated	Transformer: non-irrigated, Catboost: irrigated	Transformer irrigated, CatBoost non-irrigated	
Tana	GC	88,587	35,035	728	2690	0.973
Rift	VC	130,399	142,114	2367	1024	0.988
Koga	VC	207,536	144,284	1401	827	0.994
Kobo	VC	156,946	203,510	2313	583	0.992
Alamata	VC	84024	31,978	615	251	0.993
Liben	VC	193,428	165,933	884	971	0.995
Jiga	VC	184,116	106,384	955	833	0.994
Motta	VC	155056	65,956	1689	1363	0.986

To ensure that different irrigation detection architectures converge on similar decision boundaries, the alignment of predictions across the transformer based and CatBoost models is

investigated. Here, both architectures are training on the randomly shifted EVI timeseries of all 7 VC regions' training datasets; predictions are then made over all labeled samples. Table V-6 presents the alignment of these predictions, showing a high degree of prediction similarity: An average regional prediction alignment of 98.9% is calculated. The close alignment of predictions made by both these models expands the basis for the solution set.

Moreover, to understand the impact of the amount of training data on model performance, an ablation study is conducted where the fraction of labeled polygons included in each region's training dataset is varied between 0.15 and 0.85; the complementary fraction of each region's polygons comprises the test dataset. For each fraction of training polygons, the CatBoost model architecture is trained on all combinations of all 7 visual collection (VC) regions' training datasets; performance is assessed on the withheld VC regions in a process identical to the one described for Figures V-8 and V-10. All models are trained on randomly shifted EVI timeseries.

Figure V-11 presents the results of this ablation study, in which minimal impact is observed when varying the fraction of polygons included in each region's training dataset between 0.15 and 0.85. On average, withheld region F1 score decreases by approximately 0.05 as the fraction of training polygons drops from 0.85 to 0.15 when 1 VC region is included in the training data; this gap shrinks as additional VC regions are incorporated during training, becoming negligible for all models trained on 3 or more VC regions' data. A larger performance delta among the 10th percentile of withheld regions' F_1 scores exists when fewer than 3 VC regions are included during training; similar to the average performance metrics, this gap collapses when the classification model is trained on labeled data from 3 or more VC regions.

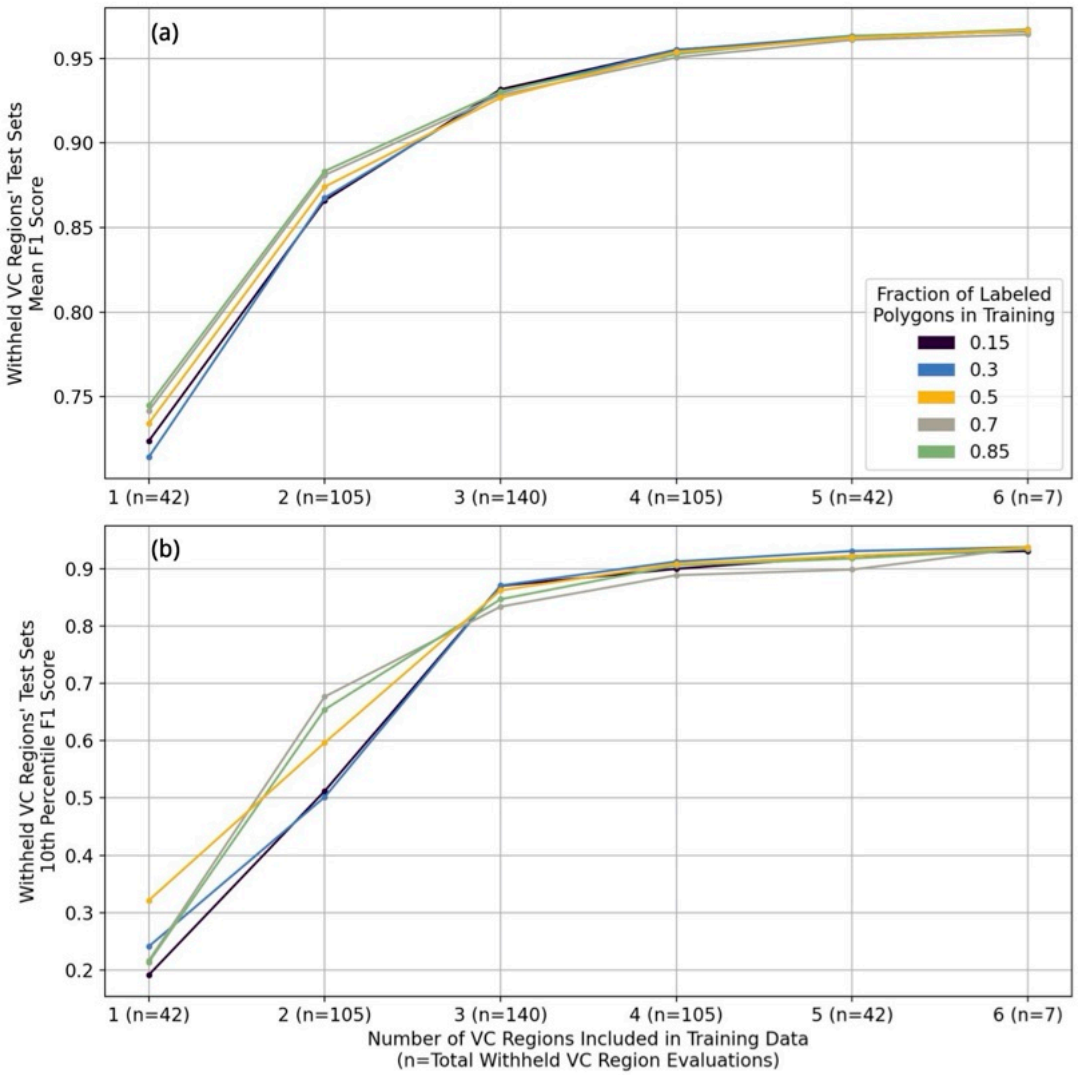

Figure V-11: Withheld region test dataset performance for different fractions of labeled polygons included in the training datasets; the complementary fractions of labeled polygons constitute the test datasets. Predictions are made using a CatBoost model architecture. (a) presents mean F1 score over the withheld regions; (b) presents the 10th percentile F1 score over the withheld regions.

Figure V-11 demonstrates that the irrigation prediction models are robust even when limiting the fraction of polygons included in training datasets to 15% of the total. Here, the inclusion of labeled data from multiple regions and randomly shifting the EVI timeseries inputs

introduces enough variance to the classification model during training so that performance over regions unseen by the classifier remains high.

Next, prediction performance over the unseen ground-collected labels in Tana is assessed. As the transformer model demonstrates the most robust performance over withheld regions' samples, it is selected for prediction, achieving 96.7% accuracy over irrigated samples (88,128/91,898) and 95.9% accuracy over non-irrigated samples (33,954/35,121) for an F_1 score of 0.932.

Lastly, through a pair of ordinary least-squares (OLS) regressions, the contribution of each VC region to target region performance can be assessed. Table V-7 presents OLS regression coefficients and P-values on target region F_1 scores for the 7 VC regions used during training, where the F_1 scores are collected over all withheld regions for all transformer classifier models presented in Figure V-10. In interpreting the regression results, variables with P-values above 0.05 are considered not statistically significant.

Table V-7 shows that training data from the regions of Alamata or Kobo have the largest impact on Tana test dataset performance, increasing F1 score on average by 0.032 or 0.024, respectively. The non-statistically significant contributions of Koga and Jiga's training data to Tana test dataset performance are highlighted, shown by the values marked with * and **. Comparing these non-statistically significant results with the relevant cells in Tables V-3 and V-4 – also marked with * and ** – reveals that non-irrigated and irrigated samples from both Koga and Jiga are more dissimilar from Tana labeled samples compared to the regional average, determined by the KS distance between the regions' data.

Table V-7: Ordinary least squares regression on withheld target region F_1 scores. F_1 scores are collected over all withheld regions for all transformer classifier models presented in Figure V-10. Values with typographical symbols are to be interpreted alongside Tables V-3 and V-4.

	Ordinary Least Squares Regressions on Withheld Region F1 Scores			
	Tana (R2=0.317, n=127)		Excluded VC Regions (R2=0.18, n=441)	
Source VC Region	Coefficient	P-value	Coefficient	P-value
Rift	0.018	0.008	-0.012	0.327
Koga	0.013*	0.056*	0.016	0.172
Kobo	0.024	0.001	0.029	0.016
Alamata	0.032	0.000	0.071†	0.000†
Liben	-0.011	0.122	0.028	0.019
Jiga	0.013**	0.063**	0.047	0.000
Motta	0.016	0.024	0.080‡	0.000‡

Table V-7 also shows that the inclusion of labeled samples from Motta and Alamata during training causes the largest increase in F1 score over the withheld VC regions; these increases, shown by values marked with † and ‡, amount to 0.08 and 0.071, respectively. Again, comparing these data points to the KS distances marked with † and ‡ in Tables V-3 and V-4 demonstrates that non-irrigated and irrigated samples from Motta and Alamata are more similar to withheld VC regions' labeled data on average, as compared to samples from other VC regions.

Taken together, the results from Tables V-3, V-4, and V-7 yield the intuitive finding that labeled samples more similar to those in a target region have a greater positive impact on performance, while more dissimilar labeled samples have a weaker effect on performance.

4.2 Model inference

For model inference, the transformer architecture is trained on the randomly shifted EVI timeseries of the labeled data from the 7 VC and one GC regions. The trained model is then

deployed over Tigray and Amhara for the 2020 and 2021 dry seasons (using imagery collected between June 1, 2019 and June 1, 2020; and between June 1, 2020 and June 1, 2021, respectively). Two post-processing steps are then taken: 1) the prediction admissibility criteria are applied, and 2) contiguous groups of predicted irrigated pixels smaller than 0.1 Ha are removed in order to ignore isolated, outlier predictions.

During inference, another step is taken to verify the accuracy of irrigation predictions. Here, five additional enumerators collect 1601 labeled polygons for the 2020 and 2021 dry seasons – 1082 non-irrigated polygons covering 3807 Ha and 519 irrigated polygons covering 582 Ha – across the extent of Amhara via the same labeling methodology used to collect the training, validation, and testing data. The locations of these independently labeled polygons are shown in Figure V-12 After cluster cleaning and applying the prediction admissibility criteria, these polygons yield 361,451 non-irrigated samples and 48,465 irrigated samples. An F_1 score of 0.917 is achieved over these samples – 98.3% accuracy over non-irrigated samples and 95.5% accuracy over irrigated samples, performance that remains in line with the reported test dataset metrics from Figure V-10 and accuracies over the withheld Tana ground-collected labels.

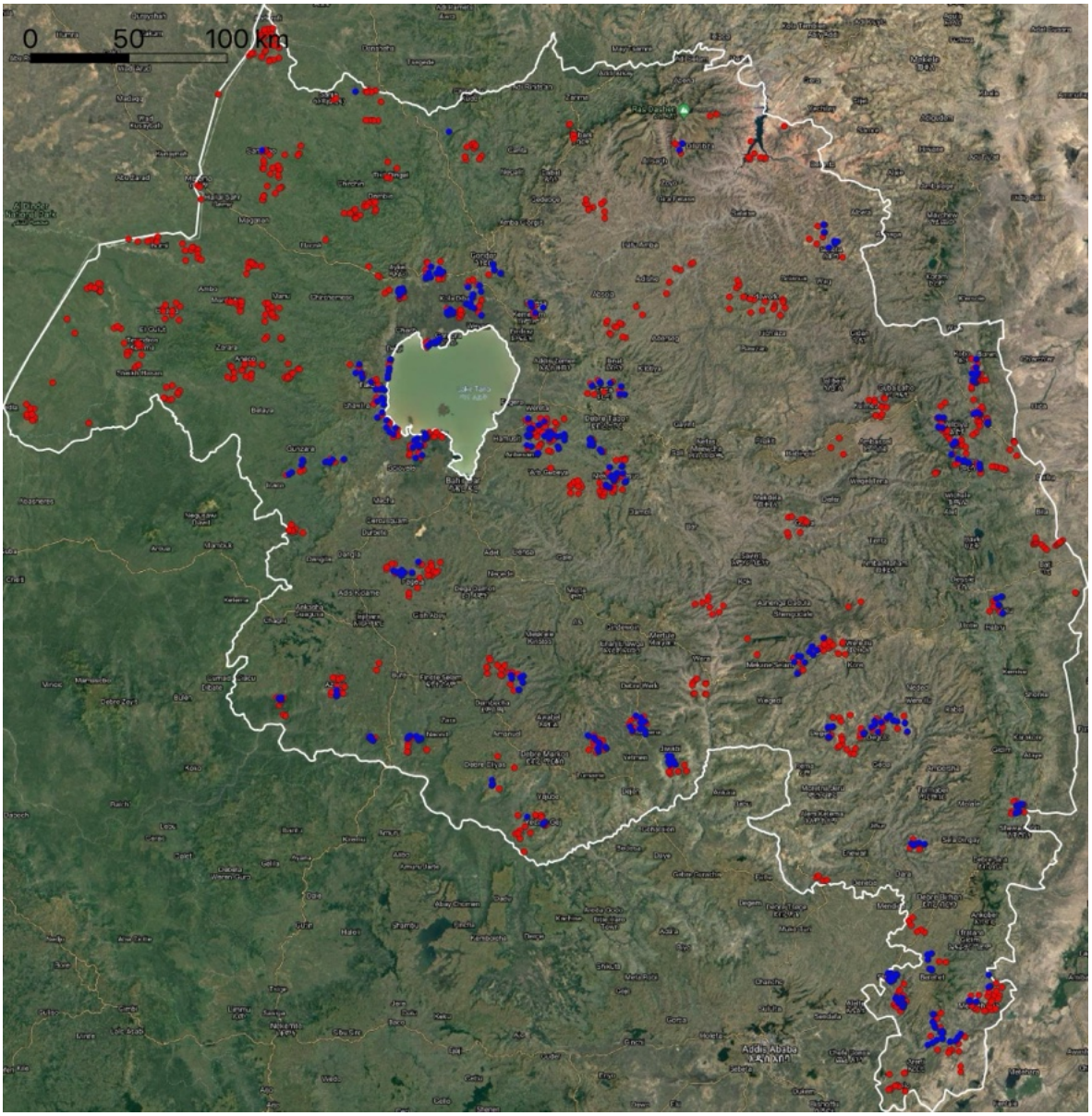

Figure V-12: Locations of independently labeled polygons for additional model performance assessment. The centroids of non-irrigated polygons are shown in red, 1082 in total; the centroids of irrigated polygons are shown in blue, 519 in total. These polygons produce 361,451 non-irrigated samples and 48,465 irrigated timeseries samples.

4.2.1 Tigray

Figure V-13 presents predicted irrigated areas in Tigray for 2020 and 2021, with 2020 irrigation predictions in red and 2021 irrigation predictions in cyan. To better understand the

nature of changing vegetation phenologies across this time period, the inset of Figure V-13 contains example timeseries that produced an irrigation prediction in one of 2020 or 2021. These example timeseries show that a second crop cycle with vegetation growth peaking in January is associated with a positive irrigation prediction; in contrast, the non-existence of this cycle is associated with non-irrigated prediction. Table V-8 displays the total predicted irrigated area for Tigray for 2020 and 2021, along with the total land area, organized by zone. Between 2020 and 2021, Table V-8 quantifies a 39.8% decline in irrigated area in Tigray.

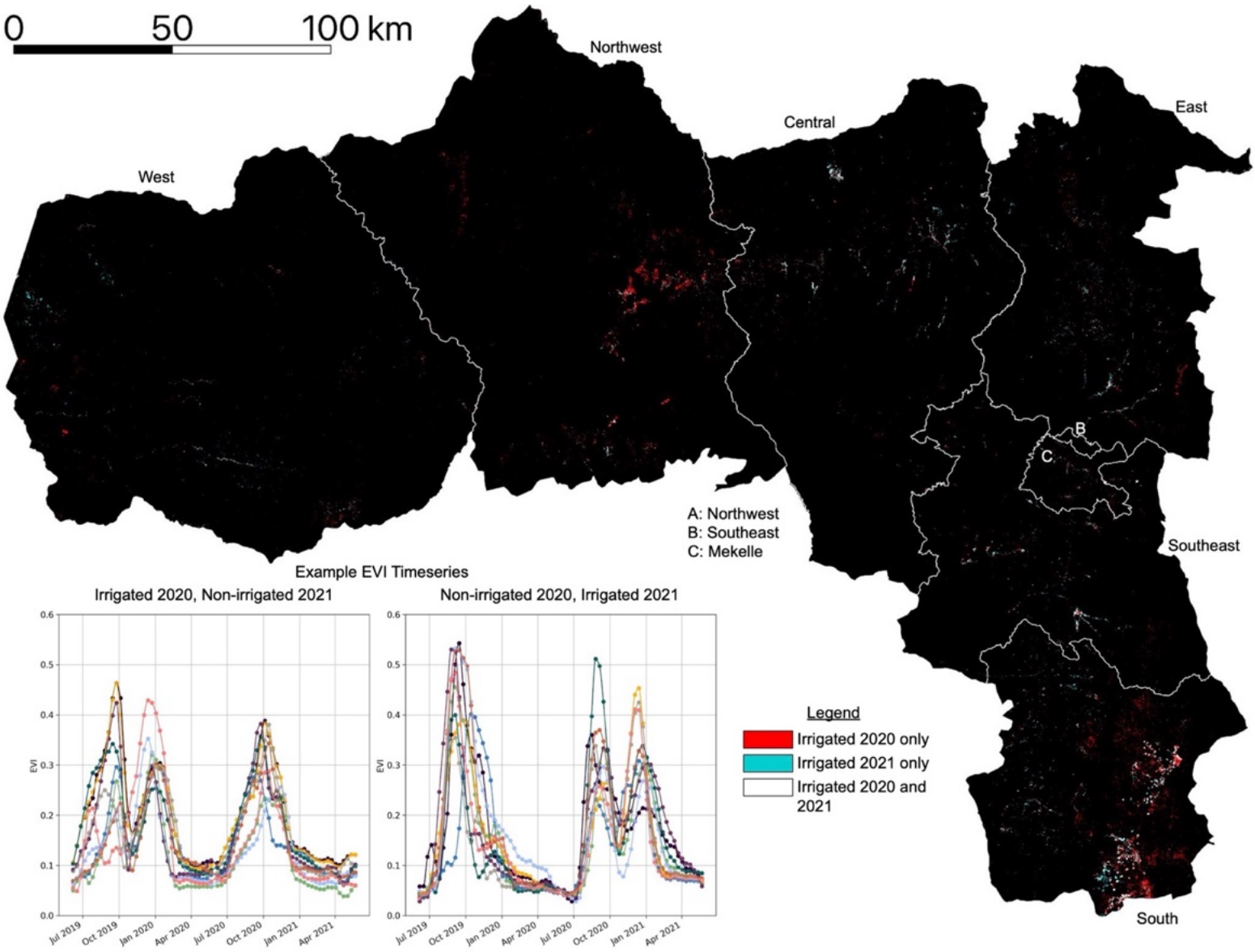

Figure V-13: Bitemporal irrigation map for Tigray. Figure inset contains example EVI timeseries predicted as irrigated in either 2020 or 2021. A predominance of red indicates that many parts of Tigray contain irrigation detected in 2020 but not in 2021.

Table V-8: Predicted irrigated area statistics in Tigray for 2020 and 2021, organized by zone.

Zone	Irrigated Ha., 2020	Irrigated Ha., 2021	Total Ha.	Percent Change, 2020 to 2021	Percent Change as Fraction of Total Area, 2020 to 2021
Central	3710	3554	954,616	-4.2%	0.0%
Eastern	3068	2863	635,670	-6.7%	0.0%
Mekelle	556	397	52,313	-28.5%	-0.3%
Northwestern	7439	2062	1,246,715	-72.3%	-0.4%
Southeastern	2658	2301	533,334	-13.4%	-0.1%
Southern	16,474	8064	506,151	-51.1%	-1.7%
Western	2278	2557	1,331,652	12.3%	0.0%
Total	36,181	21,799	5,260,451	-39.8%	-0.3%

4.2.2 Amhara

Figure V-14 presents a bitemporal irrigation map for Amhara, also with 2020 irrigation predictions in red and 2021 irrigation predictions in cyan. This map contains large clusters of irrigated predictions around Lake Tana in the zones of Central Gondar, South Gondar, and West Gojjam, an intuitive finding given the availability of water from Lake Tana and the rivers that extend off it. Irrigation is also detected in the portions of Amhara's easternmost zones that fall within the Main Ethiopian Rift (MER); as the valley formed by the MER extends north into Tigray, irrigation predictions in the North Wello, Oromia, and North Shewa zones align with irrigation predictions in the Southern zone of Tigray shown in Figure V-13. Table V-9 displays the total predicted irrigated area for Amhara for 2020 and 2021, along with the total land area, organized by zone. From 2020 to 2021, Table V-9 quantifies a 41.6% decline in irrigated area in Amhara.

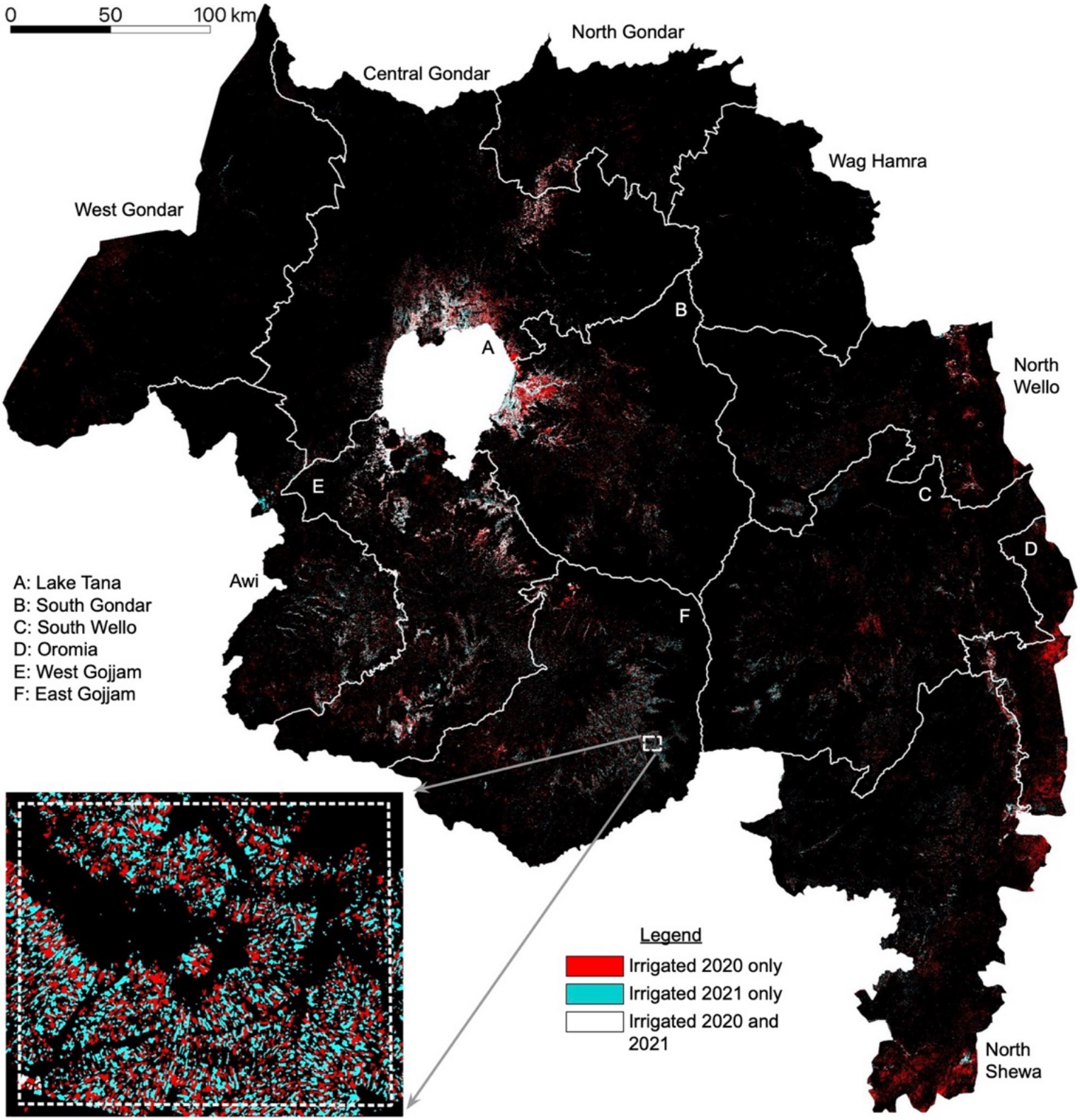

Figure V-14: Bitemporal irrigation map for Amhara. Figure inset contains example predictions around Choke Mountain displaying interannual irrigation patterns. A predominance of red indicates that many parts of Amhara contain irrigation detected in 2020 but not in 2021.

The inset of Figure V-14 presents interannual irrigated cropping patterns for an area southwest of Choke Mountain. Interlocking red and cyan plots indicate the spatial rotation of irrigated crops from 2020 to 2021; no white plots are observed, which would signify dry season crop growth in both years.

Table V-9: Predicted irrigated area for Amhara in 2020 and 2021, organized by zone.

Zone	Irrigated Ha., 2020	Irrigated Ha., 2021	Total Ha.	Percent Change, 2020 to 2021	Percent Change as Fraction of Total Area, 2020 to 2021
Awi	27,443	20,547	906,682	-25.1%	-0.8%
Central Gondar	73,450	50,954	2,095,018	-30.6%	-1.1%
East Gojam	44,975	33,888	1,405,689	-24.7%	-0.8%
North Gondar	7381	3367	684,247	-54.4%	-0.6%
North Shewa (AM)	62,933	21,362	1,622,197	-66.1%	-2.6%
North Wello	21,367	8250	1,110,856	-61.4%	-1.2%
Oromia	30,875	5285	380,773	-82.9%	-6.7%
South Gondar	72,682	43,046	1,406,698	-40.8%	-2.1%
South Wello	28,215	16,302	1,849,812	-42.2%	-0.6%
Wag Hamra	447	698	890,004	56.4%	0.0%
West Gojam	97,206	71,052	1,348,158	-26.9%	-1.9%
West Gondar	6180	1342	1,529,197	-78.3	-0.3%
Total	473,155	276,093	15,229,329	-41.6%	-1.3%

5. Discussion

This manuscript makes a set of contributions to the literature for learning from limited labels. First, it demonstrates a process of collecting training data to supplement ground-collected labels that improves on previous methods of sample collection -- such as using imagery from a single timestep or simple vegetation content heuristics -- as it verifies the existence or non-existence of full vegetation cycles during the dry season. Second, an evaluation of inputs, classifier architectures, and training strategies is presented for achieving irrigation classifier applicability to a larger area. Results indicate that enhanced vegetation (EVI) timeseries outperform a full set of spectral bands as inputs; that randomly shifting input timeseries prevents classifier models from overfitting to region-specific input features; and that a transformer-based neural network produces the highest prediction accuracies in unseen target regions. Due to the close similarity of performance metrics and alignment of predictions, the faster training, more easily interpretable CatBoost architecture is also shown as a suitable alternative for irrigation mapping efforts.

Prediction results indicate strong classifier performance over sample timeseries from regions not seen during training. On data from withheld target regions, transformer-based classifiers achieve mean F_1 scores above 0.95 when four or more regions' data are included during training; using labels from all 7 visual collection (VC) regions, the transformer-based classifier achieves an F_1 score of 0.932 on the ground collection (GC) labels around Lake Tana. Over an independently collected set of more than 400,000 samples collected for performance assessment, the same classifier achieves 98.3% accuracy over non-irrigated samples and 95.5%

accuracy over irrigated samples, demonstrating strong performance throughout the entire Ethiopian Highlands.

Deploying a transformer-based classifier trained on samples from all 8 label collection regions yields insight into changing irrigation patterns. Results suggest that from 2020 to 2021, irrigation in Tigray and Amhara decreased by 40%. In Tigray, this decline was most precipitous in the Northwest and Southern zones, which saw percent changes in irrigated area of -72.3% and -51.1%, respectively. The Western zone of Tigray was the only zone to see an increase in irrigated area from 2020 to 2021; even so, this increase amounted to 279 Ha in a zone with a total area of 1,331,652 Ha. Amhara is predicted to have had similar decreases in irrigated area: Apart from the Wag Hamra zone, which was predicted to have less than 0.08% of its area irrigated in 2020 or 2021, all zones in Amhara experienced a change in irrigated area between -25.0% and -82.3%. The largest declines by area occurred in North Shewa (-41,572 Ha), South Gondar (-29,636 Ha), and West Gojam (-26,154 Ha). Combined, results for Tigray and Amhara predict severe reductions in dry season crop growth from 2020 to 2021, findings that that align with recent reports of food insecurity following the eruption of civil conflict in Ethiopia in late 2020.

Despite presented performance metrics indicating high levels of prediction accuracy, there are a few limitations to the proposed methodology that are important to mention. First, the study area is limited to the Ethiopian Highlands, a highly agricultural, climatologically consistent area that is dominated by rainfed cropped phenologies. As the irrigation classifiers are only trained to separate dry season crop cycles from rainfed vegetation cycles -- associating identified dry-season cropping with irrigation presence -- they will perform poorly in settings with different rainfall and phenological patterns. Relatedly, the trained irrigation classifiers do not

identify irrigation used to supplement rainy season precipitation, irrigation of perennial tree crops, evergreen vegetation in riparian areas, or irrigation that supports continuous cropping, as the phenological signatures of these types of vegetation are difficult to distinguish from evergreen, non-cropped signatures. This discrimination task is left for future work. Lastly, classifiers are trained only on cropped phenologies, which constitute a portion of the vegetation signatures that exist in the area of interest. To manage the other phenologies present at model inference, prediction admissibility criteria are implemented. Nevertheless, these criteria are imperfect: There are surely irrigated pixels which have been mistakenly assigned a non-irrigated class label, along with non-cropped pixels which have evaded the admissibility criteria.

While the presented methodology is applied only for the task of irrigation identification in the Ethiopian Highlands, the strategy of regional phenological characterization to provide context for geographically informed selection of training samples and model applicability can be transferred more broadly to a range of land process mapping objectives. The suitability of this approach in the field of machine learning with limited labels is supported by results comparing classifier architectures and hyperparameter choice to assess the question of result uniqueness that overshadows all land cover classifications. As discussed by [300], what is presented as *the* map is often just *a* map -- one of many different products that can be obtained from the same set of inputs with different classifiers and hyperparameter settings. By assessing multiple classifier architectures and quantifying prediction sensitivity, this approach demonstrates consistency in results and indicates the uncertainty that can be expected of the resulting irrigation maps; as such, it provides a process for building robust classifiers in settings with scarce labeled data.

6. Appendix A

Appendix A presents labeled samples before and after cluster cleaning for all label collection regions except Koga, which is presented in Figure V-5.

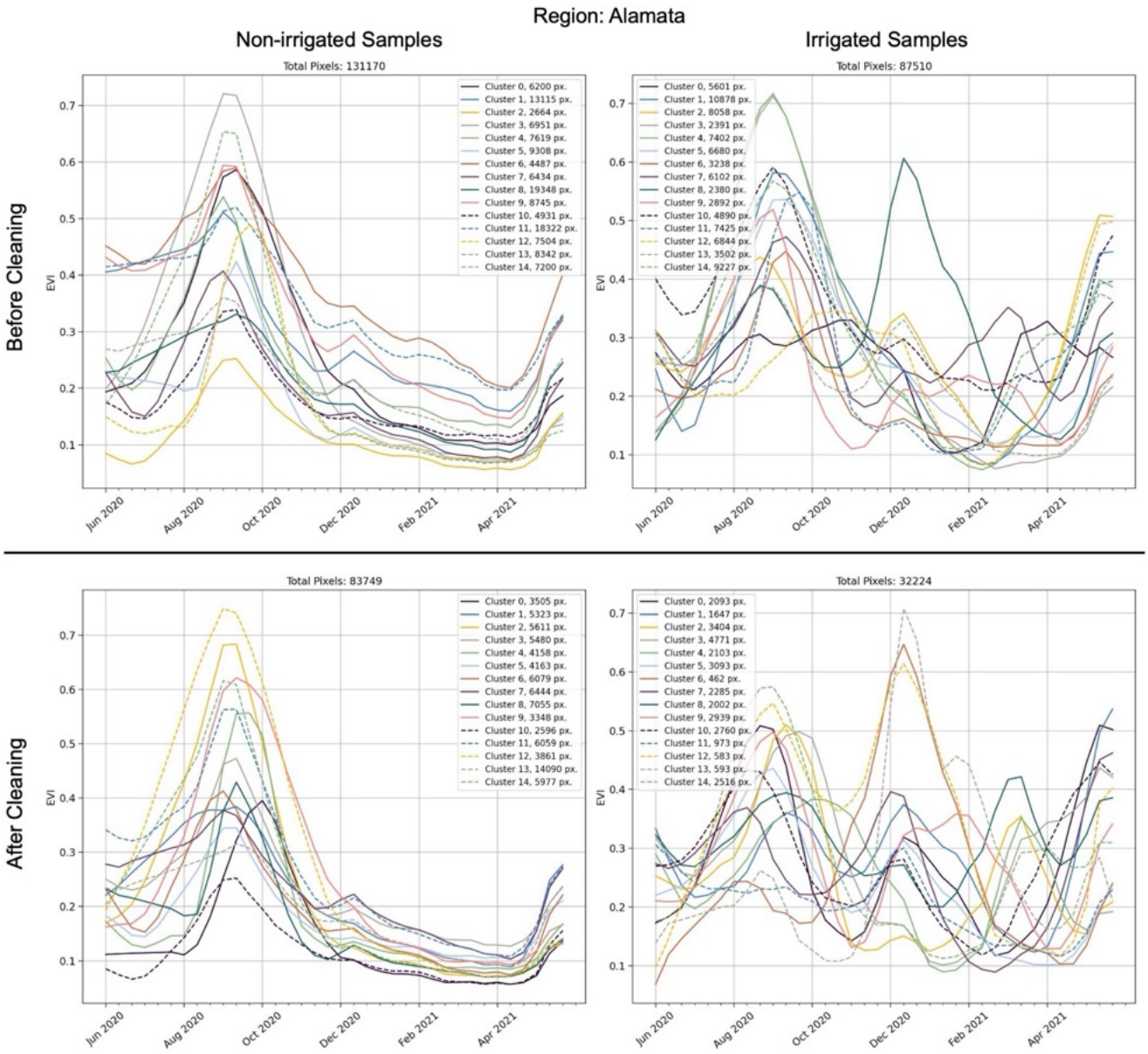

Figure V-15: Clustered enhanced vegetation index (EVI) timeseries before and after cluster cleaning for the Alamata region. After cleaning, all non-irrigated clusters display a single vegetation peak aligned with the main rainy season, and the irrigated clusters all display a vegetation cycle during the dry season.

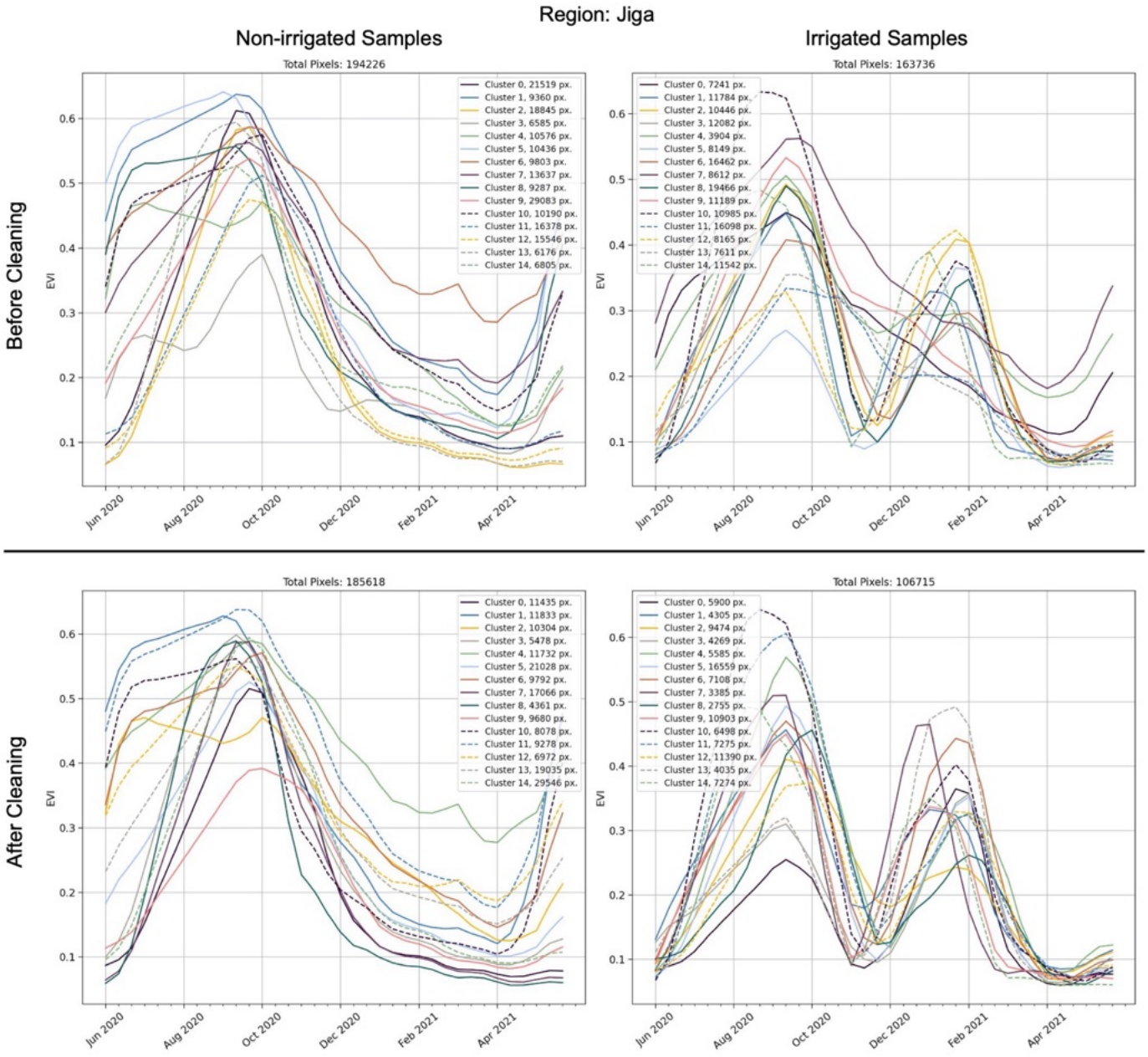

Figure V-16: Clustered enhanced vegetation index (EVI) timeseries before and after cluster cleaning for the Jiga region. After cleaning, all non-irrigated clusters display a single vegetation peak aligned with the main rainy season, and the irrigated clusters all display a vegetation cycle during the dry season.

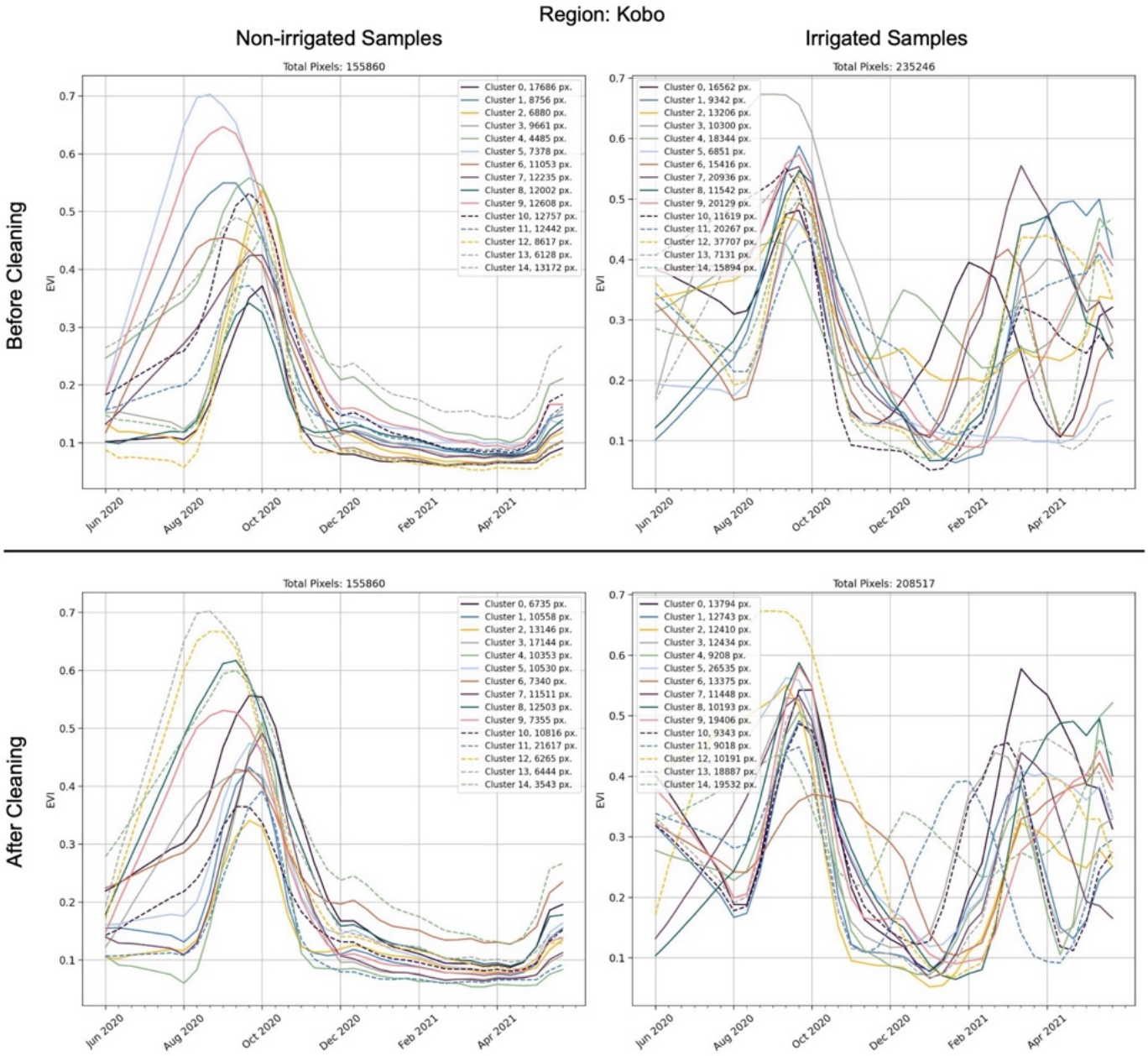

Figure V-17: Clustered enhanced vegetation index (EVI) timeseries before and after cluster cleaning for the Kobo region. After cleaning, all non-irrigated clusters display a single vegetation peak aligned with the main rainy season, and the irrigated clusters all display a vegetation cycle during the dry season.

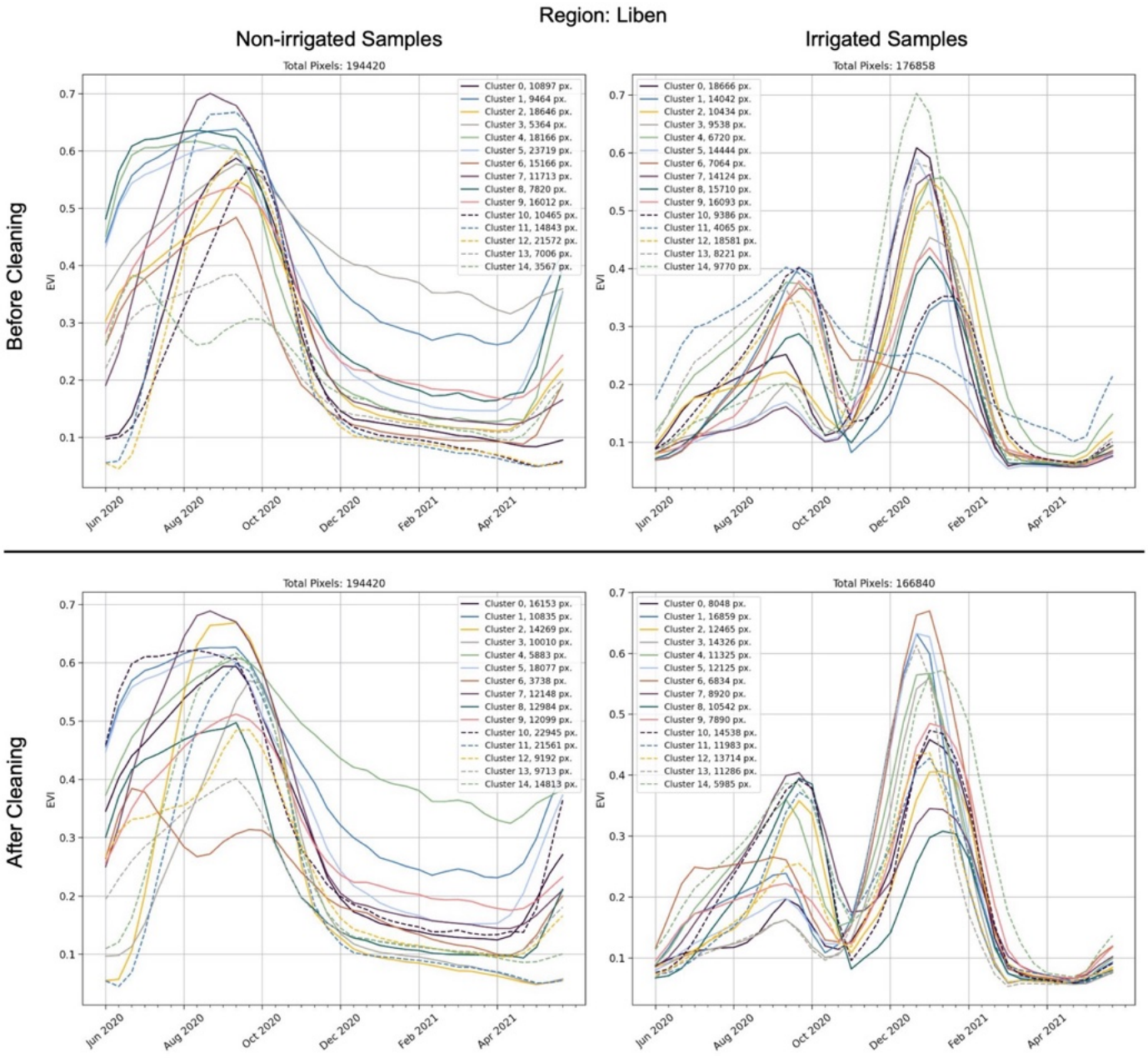

Figure V-18: Clustered enhanced vegetation index (EVI) timeseries before and after cluster cleaning for the Liben region. After cleaning, all non-irrigated clusters display a single vegetation peak aligned with the main rainy season, and the irrigated clusters all display a vegetation cycle during the dry season.

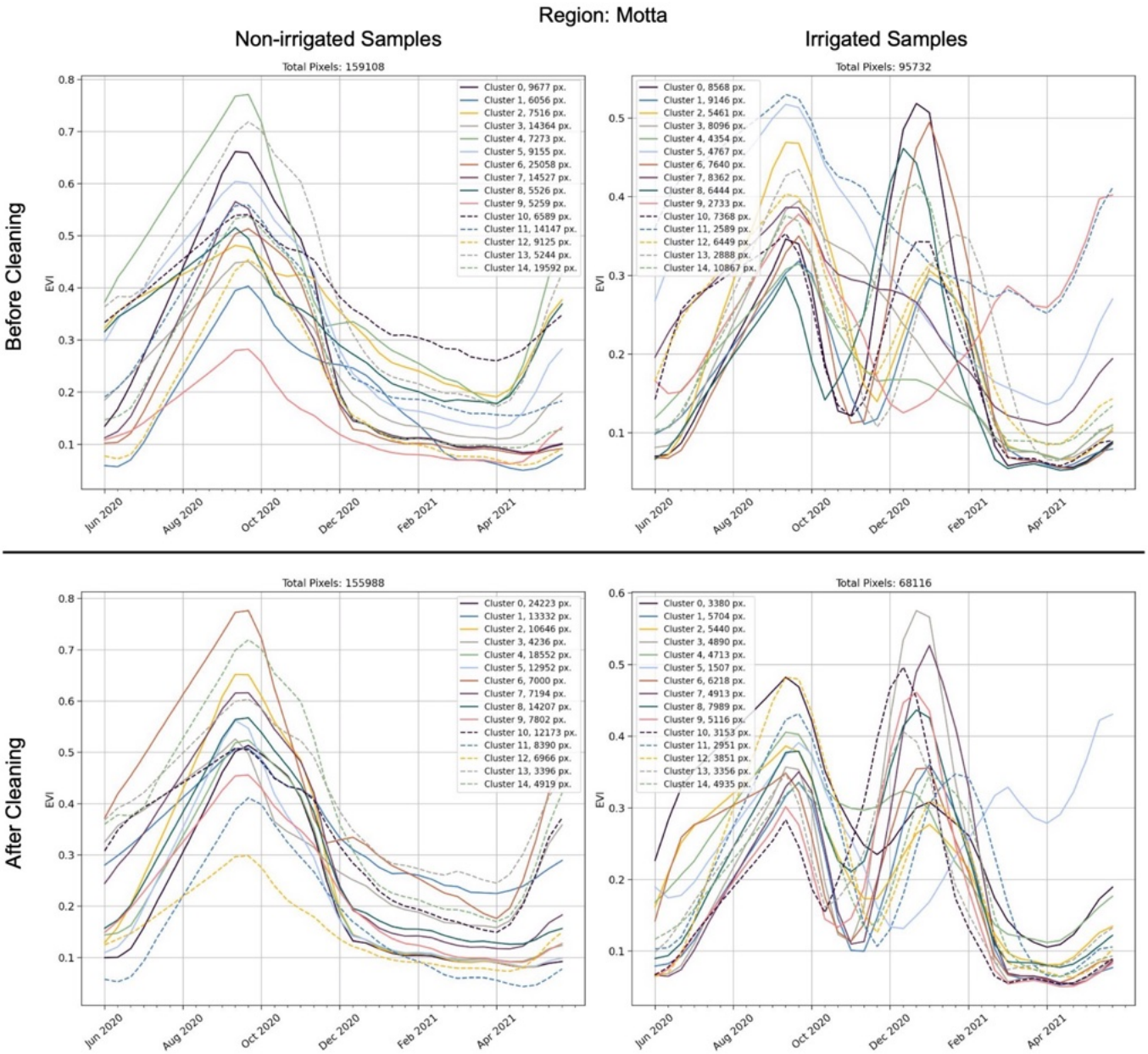

Figure V-19: Clustered enhanced vegetation index (EVI) timeseries before and after cluster cleaning for the Motta region. After cleaning, all non-irrigated clusters display a single vegetation peak aligned with the main rainy season, and the irrigated clusters all display a vegetation cycle during the dry season.

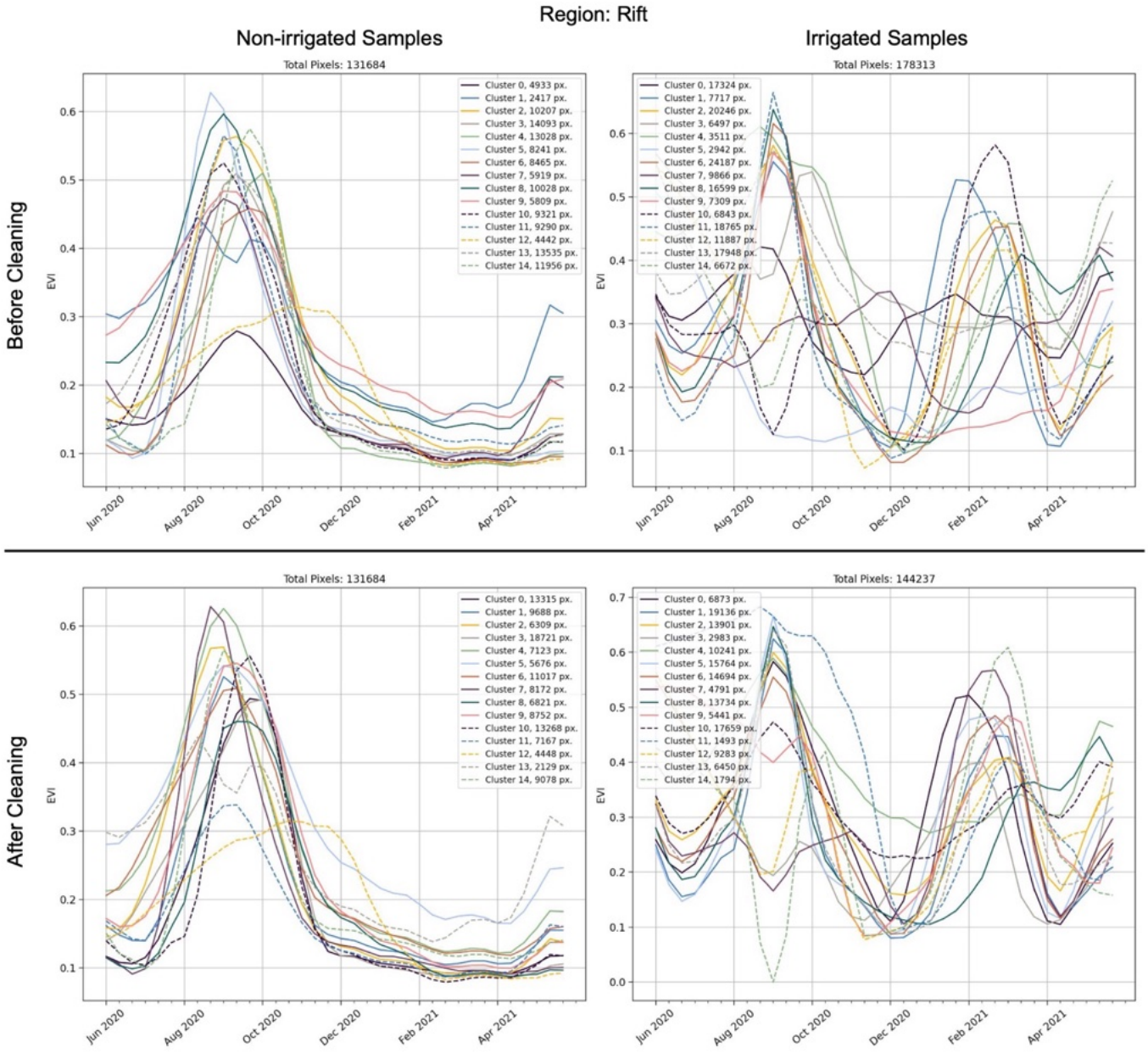

Figure V-20: Clustered enhanced vegetation index (EVI) timeseries before and after cluster cleaning for the Rift region. After cleaning, all non-irrigated clusters display a single vegetation peak aligned with the main rainy season, and the irrigated clusters all display a vegetation cycle during the dry season.

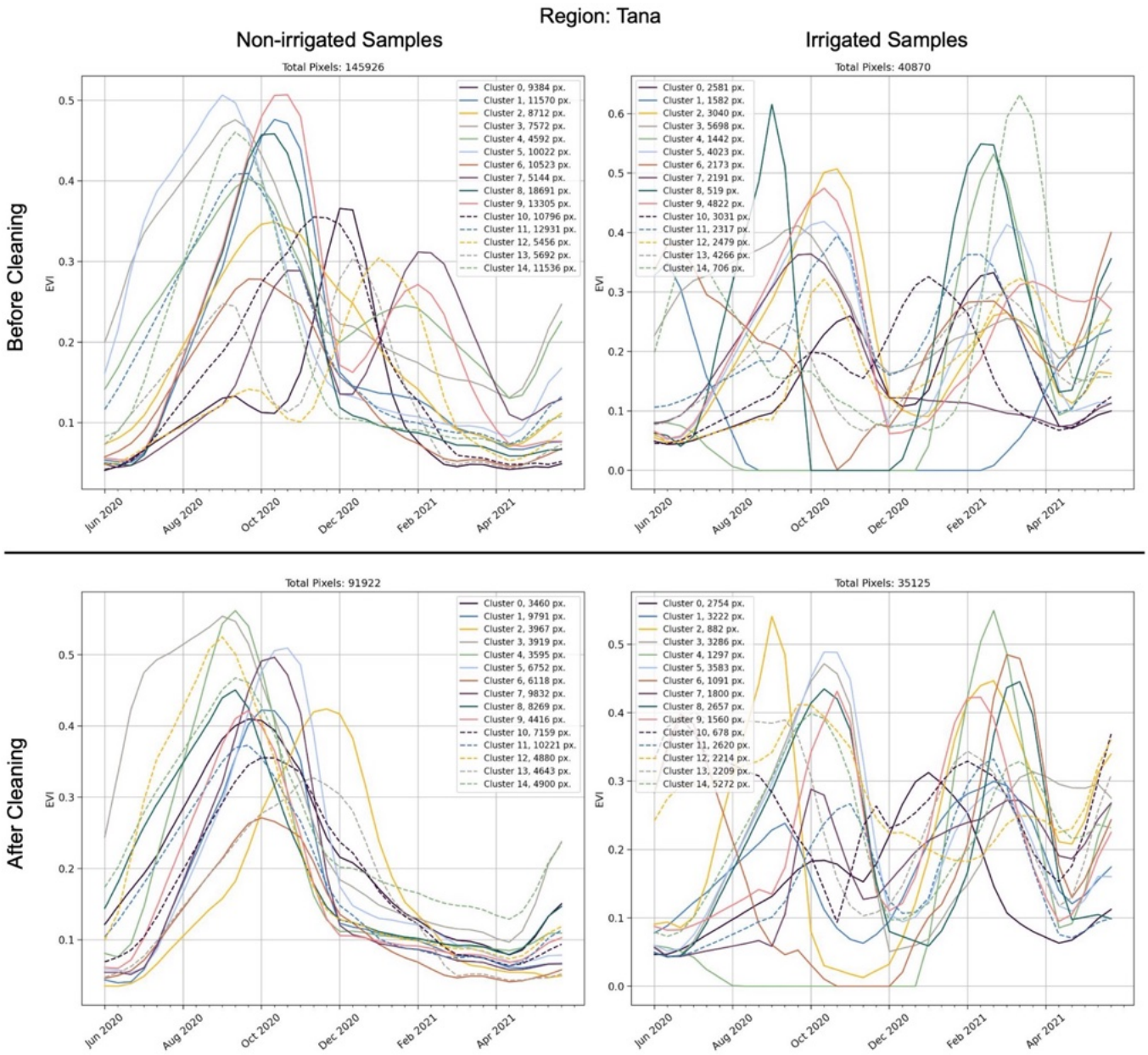

Figure V-21: Clustered enhanced vegetation index (EVI) timeseries before and after cluster cleaning for the Tana region. After cleaning, all non-irrigated clusters display a single vegetation peak aligned with the main rainy season, and the irrigated clusters all display a vegetation cycle during the dry season.

Chapter VI: Extensions of irrigation detection efforts in Ethiopia

Abstract

Following the development of an irrigation detection methodology for the Ethiopian Highlands, further irrigation prediction extensions are explored. First, segment area of irrigated predictions is quantified, revealing that while the area distribution of segments is highly similar across Tigray and Amhara – 99% of predicted segments in both states are smaller than 10 Ha – the largest 1% of predictions in Amhara are substantially larger than the largest 1% in Tigray. Next, anomalous predictions of irrigation over non-cropped land from Chapter V are investigated. Inspection of these predictions uncovers atypical rains and vegetation growth during the 2020 dry season, phenomena that occur once during the 10-year span examined and cause the erroneous predictions in question. Additional analysis leverages the 10m European Space Agency (ESA) WorldCover v100 land cover map to filter out all predictions that don't fall on estimated cropland. Restricting irrigation predictions per the ESA land cover map eliminates many of the previously identified false positive predictions, in total removing approximately 8% of predicted irrigated area across Tigray and Amhara in 2020 and 2021.

Another research extension involves irrigation mapping throughout Tigray for the 2019 dry season, predictions that reveal a 6% increase in irrigated area from 2019 to 2020 and substantial year-to-year change in the plot-level locations of irrigation. As irrigated area in Tigray sees moderate annual growth from 2019 to 2020, we can further conclude that the sharp decline in irrigated area from 2020 to 2021 posited in Chapter V is not due to excessive false positive predictions for the region in the 2020 dry season. Irrigation predictions are also made for a part

of southern Oromia with vegetation patterns highly similar to those in the Ethiopian Highlands. These predictions identify dry season crop growth along a reservoir near the city of Asasa, demonstrating that the developed irrigation prediction model is not bound by any administrative borders and can be applied where background phenological cycles are similar to those in regions where the model is trained.

Two additional irrigation detection extensions contained in this chapter involve first predicting longitudinal irrigation growth from 2017 to 2021 near Hawzen, Tigray, to compute irrigation expansion over a five-year period. Next, prediction accuracy for the 2020 irrigation season in Tigray is assessed using the method of visual label collection introduced in Chapter V. Here, we determine that irrigation predictions for a year and spatial extent where the irrigation prediction model was not trained maintain levels of accuracy above 95%, proving the robustness of our detection methodology.

Lastly, the addendum to this chapter contains a fully open-source implementation of the methodology introduced in Chapter V. In Chapter V, the Descartes Labs platform was used to generate Sentinel-2 imagery timeseries for irrigation detection. Here, I replicate that functionality using Google Earth Engine, a platform freely available to researchers. To prove its utility, this new open-source code is deployed for end-to-end model training and prediction for a region near Lake Tana in Amhara for the 2020 irrigation season. These new predictions demonstrate up to 99% agreement with the 2020 predictions introduced in Chapter V, revealing the robustness of the introduced irrigation detection methodology.

1. Irrigation detection extensions

1.1 Irrigated segment evaluation

Chapter V contains maps for predicted irrigation presence throughout Tigray and Amhara for 2020 and 2021. Evaluating the geometric distributions of these predictions yields insights about the sizes of predicted irrigated segments. Figure VI-1 contains cumulative density functions (CDFs) for total irrigated area and total irrigated segments, both plotted against the logarithm of the total irrigated segment area. Here, segments are defined as contiguous pixels predicted as irrigated surrounded by non-irrigated predictions. Positive predictions are linked using queen's case[xxi] adjacency: If two positive (i.e., irrigated) predicted pixels are connected along corners or edges, they are considered part of the same segment.

Figure VI-1(a) reveals that total predicted irrigated area in Amhara is constituted by larger segments compared to Tigray. In Amhara in 2020, segments less that 10 Ha and 100 Ha make up 54% and 71% of irrigated area, respectively; segments smaller than the same thresholds in Tigray in 2020 make up 63% and 88% of irrigated area, respectively. For both Tigray and Amhara, predicted irrigated area in 2021 is constituted by smaller segments compared to 2020. This gap appears for all segments larger than 1 Ha, as an equal proportion of the states' predicted irrigated areas is made up of segments less than 1 Ha for both regions in 2020 and 2021. However above 1 Ha, both states contain larger predicted irrigated segments in 2020 compared to 2021. This finding mirrors the annual quantities of predicted irrigated areas in Tigray and Amhara, presented first in Chapter V and displayed in the legend of Figure VI-1(a), which determine an overall reduction in irrigated area from 2020 to 2021.

[xxi] Queens case adjacency is named in reference to the movement pattern of a queen in chess.

In contrast, Figure VI-1(b) shows minimal differences in segment area across years and regions for all predicted irrigated segments. For both years and states, approximately 99% of segments are less than 10 Ha in size. Where the distributions diverge is in the largest 1% of segments: For Amhara, these largest segments reach a size of 22,155 Ha in 2020 and 8844 Ha in 2021; for Tigray, these largest segment areas are 883 Ha in 2020 and 235 in 2021. Comparing to Figure VI-1(b), one observes that the largest 1% of segments – i.e., those over 10 ha – makes up either a plurality or a majority of predicted irrigated areas, as is quantified in the preceding paragraph.

The results contained in Figure VI-1 present characterizations of irrigated area segments larger than or equal to 0.1 ha. Per an assumption introduced in Chapter V, all predicted irrigated segments smaller than 0.1 ha are ignored, as these are likely isolated, anomalous predictions. Moreover, since the irrigation detection model utilized operates in a pixelwise fashion, no spatial relationships are considered during model inference. By limiting final predictions to only those residing in segments larger than or equal to 0.1 Ha, spatial relationships between adjacent pixels are incorporated and many false positive predictions are removed.

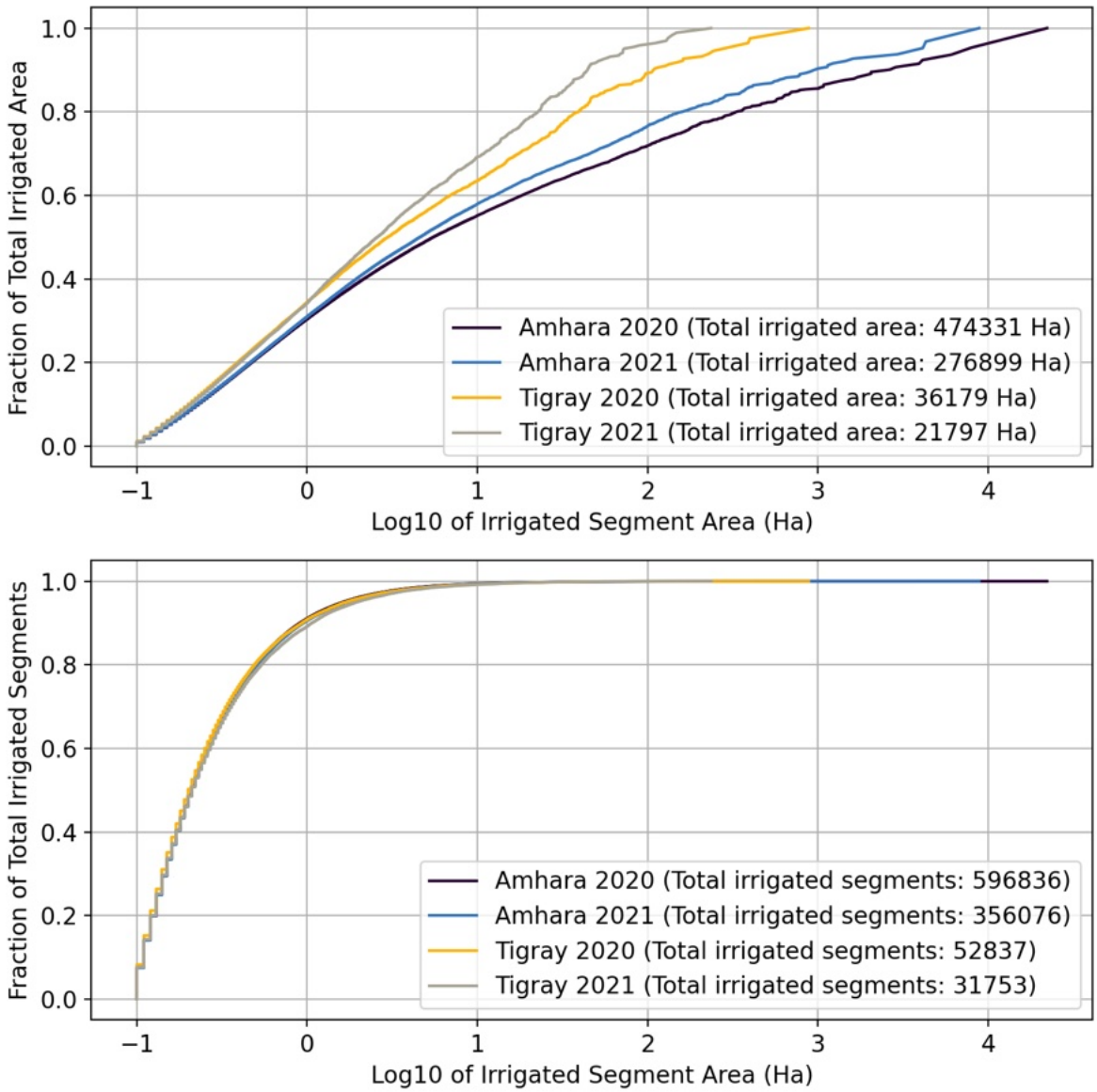

Figure VI-1: Cumulative distribution functions (CDFs) of (a) total irrigated area, and (b) total irrigated segments, plotted against logarithm base 10 of the irrigated segment area in hectares.

Table VI-1 presents an accounting of the number and cumulative size of irrigated segments less than 0.1 Ha that are removed in this filtering process. This table shows that greater than 82% of all predicted irrigated segments in both Tigray and Amhara in 2020 and 2021 are smaller than 0.1 Ha, but that less than 18% of predicted irrigated area is ignored by removing these predictions. Here, the large quantity of small predicted irrigated segments underlines the

importance of filtering final predictions by a lowest allowable segment area. Without this filtering, the large number of small predicted irrigated segments causes substantial speckle in predicted irrigation maps, distracting from areas with more likely irrigation presence.

Table VI-1: Effects of removing all predicted irrigated area segments less than 0.1 Ha.

	Fraction of irrigated segments removed	Fraction of irrigated area removed
Tigray, 2020	0.86	0.17
Tigray, 2021	0.88	0.18
Amhara, 2020	0.82	0.12
Amhara, 2021	0.84	0.14

1.2 Prediction confusion due to varying background vegetation phenologies

A key insight of the methodology presented in Chapter V is that there are spatial limits on the applicability of any model trained to detect dry season crop growth as an indication of irrigation presence. This insight is summarized in the spatiotemporal mixture map in Figure V-1, which presents the contribution of four distinct phenologies to background vegetation cycles for the entirety of Ethiopia. Here, areas of Ethiopia that appear red have vegetation cycles that closely match the red temporal endmember (tEM) shown in the bottom of Figure V-1; the same applies for blue and green areas of Ethiopia, while secondary colors indicates that the background phenology is best represented by a combination of multiple tEMs.

As different dominant vegetation patterns are driven by different climatological conditions, one can assume that areas in a spatiotemporal mixture map that present in similar color gradients have similar weather patterns. The inverse statement also holds: Areas that present in different colors will contain distinct weather patterns. In Figure V-1, the Amhara and Tigray states are dominated by red hues and can be said to largely follow a single vegetation cycle

that aligns with the long rains during the months of June through August. Accordingly, both rainy and dry season vegetation will look similar across these areas; this is the assumption that underpins the applicability of the presented irrigation detection model throughout the entirety of both states.

Outside the boundaries of Amhara and Tigray, divergent climatological conditions will yield vegetation cycles that will confuse trained irrigation detectors. For example, in the southeast of Ethiopia, persistent rains produce evergreen vegetation throughout most of the Southern Nations, Nationalities, and Peoples (SNNP) region. In such a rainy, evergreen vegetation-dominated area, a dry season crop detector trained to identify clear vegetation cycles will struggle. This confusion can also occur near the borders of Amhara and Afar, as Ethiopia's regional borders often align with dominant climatological patterns[xxii]. In particular, the Oromia zone of Amhara – not to be confused with the Oromia *region* of Ethiopia – that abuts Afar presents as magenta in Figure V-1, indicating that vegetation cycles in this zone can be modeled through combinations of red (*single)* and blue (*double)* tEMs, also shown in Figure V-1.

[xxii] Evidence of this alignment can be seen on the borders of Oromia-Amhara, Amhara-Afar, and Somali-Oromia in Figure V-1, where white administrative boundaries separate distinct color gradients.

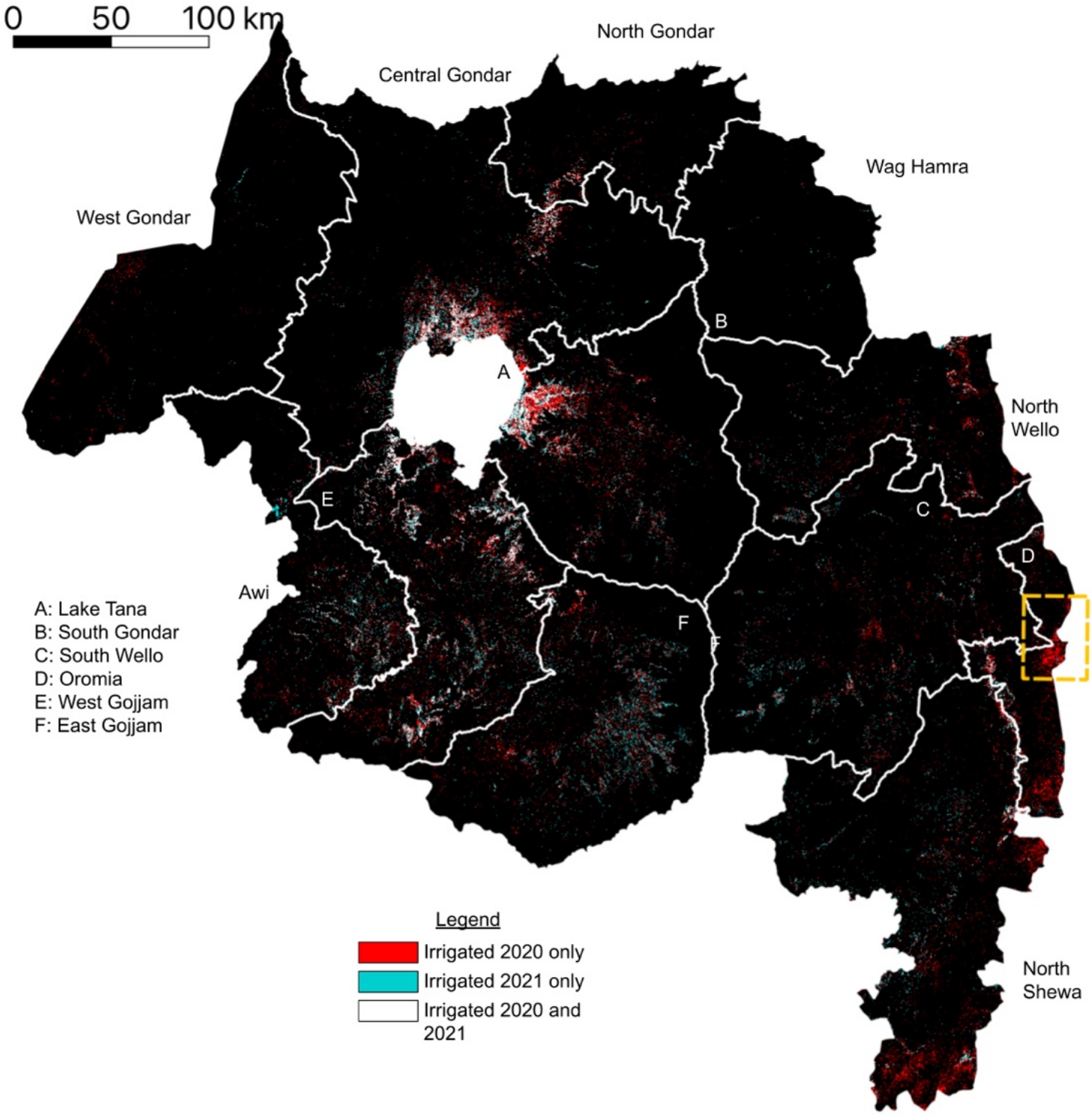

Figure VI-2: Bitemporal irrigation map for Amhara, identical to the map presented in Figure V-13. A portion of the Oromia zone in eastern Amhara containing predicted irrigation in 2020 – signified in red – is outlined with a yellow rectangle.

Figure VI-2 shows the same bitemporal irrigation map introduced in Figure V-13, with a yellow rectangle outlining a portion of the Oromia zone in eastern Amhara. Contained within this yellow rectangle are a number of pixels predicted as irrigated in 2020, but not 2021, as is evidenced by the red hue. Henceforth, this area is referred to as “Bati”, the name of a nearby town.

A closer look into the underlying land cover reveals a degree of irrigation misclassification. Figure VI-3 presents a high-resolution image of an area near Bati containing predicted irrigated pixels; this image is collected from Google Earth Pro. No cropland is immediately evident in Figure VI-3, and the vegetation that does exist appears to be either 1) evergreen riparian vegetation along the two East-West oriented stream banks; or 2) shrub vegetation in the triangular area in between. Why then were many pixels in this area of interest predicted as irrigated?

Figure VI-3: High resolution image near the town of Bati revealing non-cropped land cover for pixels that were predicted as irrigated during 2020. The image was collected in January 2019 and is located at 10.81°N, 40.15°E; the purple line indicates the border between Amhara and Afar. The image was retrieved from Google Earth Pro.

Figure VI-4 presents MODIS enhanced vegetation index (EVI) timeseries containing annual vegetation cycling for a 250m pixel that overlaps many 10m positive irrigation predictions for 2020 near the town of Bati. EVI timeseries are collected at a 16-day timestep from June 1, 2011 to June 1, 2021; these timeseries are then grouped by year. Figure VI-4 shows primary July – September vegetation cycles that align with those contained throughout Amhara and Tigray. Moreover, secondary vegetation cycles during the smaller April – May *belg* rains are present in

most years. These secondary cycles are the reason for the magenta hue of this area in the Figure V-1: The red *single cycle* tEM needs to combine with the blue *double cycle* tEM to represent this secondary greening and senescence. Differences between the magenta hue along the Amhara/Afar border and the solid red hue throughout most of Tigray and Amhara indicate this difference in background vegetation patterns.

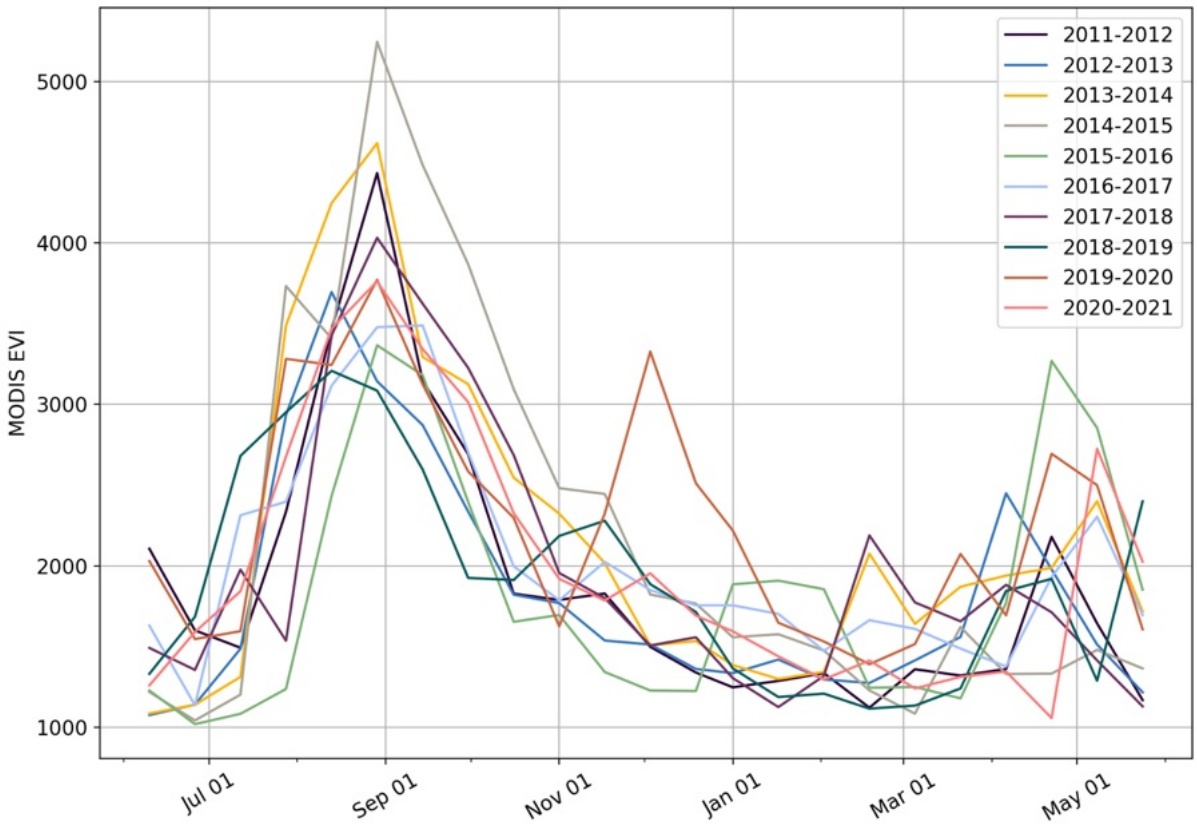

Figure VI-4: Annual MODIS enhanced vegetation index (EVI) cycles collected near the town of Bati. Vegetation phenologies indicate greening and senescence that align with the primary cropping season from July – October, along with secondary vegetation cycles during the smaller rains of April and May. Anomalous winter vegetation during the 2019-2020 dry season is shown by the salmon-colored line.

While these secondary vegetation cycles indicate growth not seen throughout most of Amhara and Tigray, they do not explain why the detection model predicted dry season (December – April) irrigation presence over non-cropped land near Bati in 2020. Figure VI-4

uncovers the reason for these positive predictions: The 2019-2020 dry season contains an anomalous vegetation cycle from November to January, as is indicated by the salmon-colored line. Here, this additional vegetation cycle is confused with irrigation, as many irrigated areas throughout Amhara and Tigray display a similar phenology.

Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) precipitation estimates explain this anomalous vegetation cycle [265]. Figure VI-5 presents monthly CHIRPS precipitation data for June 2011 to June 2011, grouped by month. In all 10 years, primary July – September and secondary March – April rains are evident. However, from 2019-2020, additional rainfall events are predicted from October to December, culminating in an estimated 100mm of rain in November 2019. The 2019-2020 dry season vegetation cycle presented in Figure VI-4 can thus be attributed to these additional rains. Given the fact that dry season rainfall/vegetation cycles are present in only one of 10 years assessed, they are deemed uncharacteristic for the area of interest near Bati.

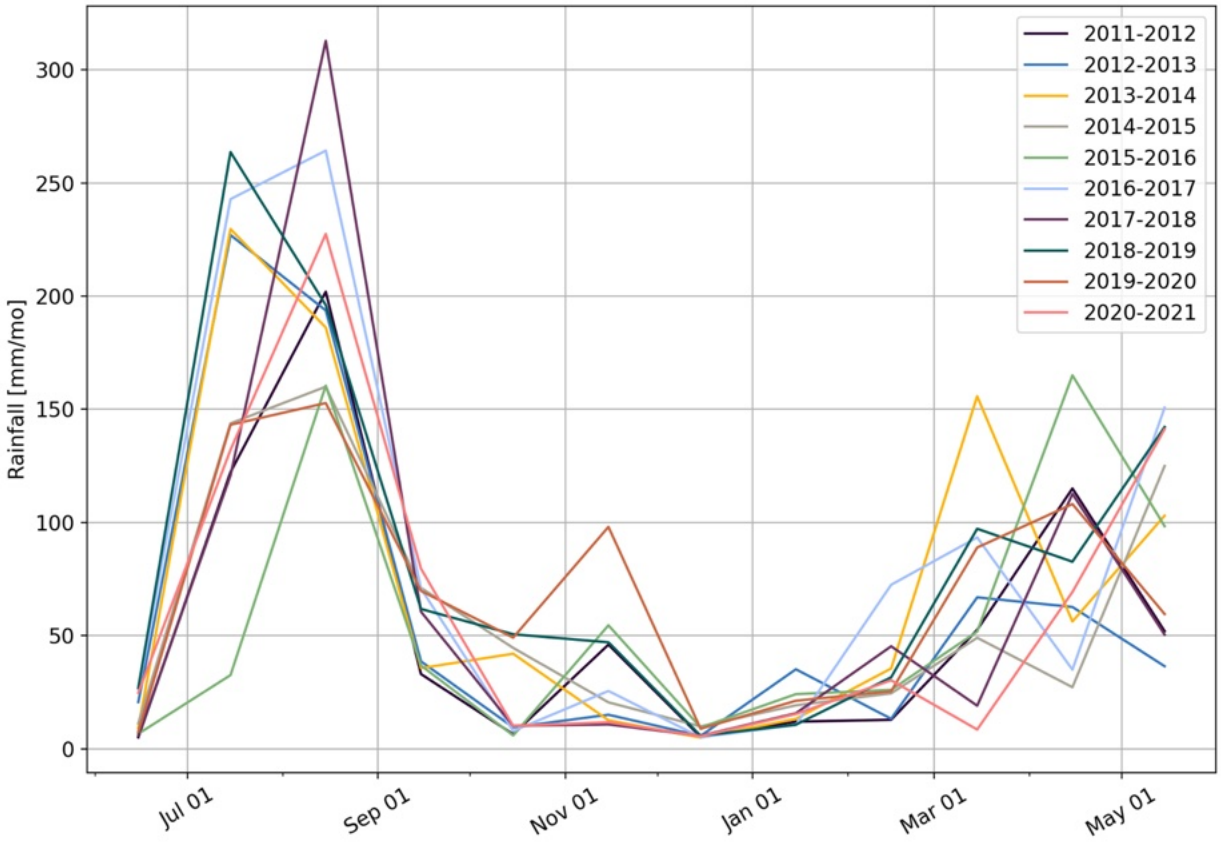

Figure VI-5: Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) monthly rainfall estimates near the town of Bati, grouped by year. Timeseries demonstrate primary rains from July – September, with secondary rains during March and April. Anomalous winter rains during the 2019-2020 dry season are shown in the salmon-colored line.

1.3 Filtering irrigation predictions using an ESA land cover layer

Chapter VI, Section 1.2 provides examples of misclassified irrigation predictions over non-cropped areas. A certain level of this type of misclassification is expected since the irrigation detection classifier does not see any non-cropped land cover samples during training, and because the imposed prediction admissibility criteria will be imperfect. However, every effort should be made to filter out these misclassifications in the final irrigation prediction maps. As

such, this subsection explores the use of a European Space Agency (ESA) land cover (LC) map to restrict positive irrigation predictions to only areas assessed to be cropland.

ESA recently published the WorldCover 2020 Map v100, a global LC map [301]. With its 10m resolution, the map joins another recent product from ESRI and Impact Observatory [302] to offer the highest resolution of any publicly available global LC map. To create the ESA LC product, researchers generated labeled pixels using GEOWIKI. These training labels were paired with Sentinel-2 (S2) Level-2A (L2A) multispectral and Sentinel-1 (S1) radiometric terrain corrected (RTC) backscatter timeseries. Cloudy, saturated, or shadowed S2 pixels are removed using the scene classification layer of the L2A product. A gradient boosted decision tree algorithm – CatBoost – is then used for model training [295]. As a final step, OpenStreetMaps [303], Global Human Settlement Layer [304,305], and Global Surface Water Explorer [306] auxiliary datasets are used in the determination of expert rules for improving prediction quality. The ESA WorldCover 2020 Map v100 achieves an overall accuracy of 74.4% across the 11 classes simulated[xxiii] on the holdout validation tiles. In Africa, the map achieves 73.6% overall accuracy. Given the map's extent, spatial resolution, and accuracy, its LC classifications are used to filter out positive irrigation predictions in the Ethiopian Highlands that fall on land not deemed to be cropland. In practical terms, this means changing all irrigated predictions over non-cropland to non-irrigated predictions.

[xxiii] The 11 classes simulated in the ESA WorldCover 2020 v100 map are: tree cover, shrubland, grassland, cropland, built-up, bare/sparse vegetation, snow and ice, permanent water bodies, herbaceous wetland, mangroves, and moss and lichen.

Table VI-2: Effects of filtering positive irrigation predictions in Tigray with the 2020 ESA WorldCover v100 land cover layer.

Zone	Total Ha.	Total Cropped Ha.	Irrigated Ha., 2020	Filtered Irrigated Ha., 2020	Percent Filtered Out, 2020	Irrigated Ha., 2021	Filtered Irrigated Ha., 2021	Percent Filtered Out, 2021
Central	954,616	239,732	3710	3517	5%	3554	3382	5%
Eastern	635,670	214,153	3068	2846	7%	2863	2679	6%
Mekelle	52,313	33,243	556	505	9%	397	367	8%
Northwestern	1,246,715	160,054	7439	6890	7%	2062	1805	12%
Southeastern	533,334	199,714	2658	2475	7%	2301	2219	4%
Southern	506,151	194,463	16,474	14,894	10%	8064	7898	2%
Western	1,331,652	263,993	2278	1720	24%	2557	1800	30%
Total	5,260,451	1,305,351	36,181	32,848	9%	21,799	20,149	8%

Table VI-2 presents the impact of cropland filtering on positive irrigation predictions in Tigray. Overall, 8-9% of irrigation predictions are filtered out in both 2020 and 2021. These prediction removals are evenly distributed across all zones in Tigray except for the sparsely cropped Western Zone, which sees 24% of predicted irrigated pixels removed in 2020 and 30% removed in 2021.

Table VI-3: Effects of filtering positive irrigation predictions in Amhara with the 2020 ESA WorldCover v100 land cover layer.

Zone	Total Ha.	Total Cropped Ha.	Irrigated Ha., 2020	Filtered Irrigated Ha., 2020	Percent Filtered Out, 2020	Irrigated Ha., 2021	Filtered Irrigated Ha., 2021	Percent Filtered Out, 2021
Awi	906,682	445,190	27,443	25,552	7%	20,547	17,388	16%
Central Gondar	2,095,018	869,274	73,450	72,671	1%	50,954	50,278	1%
East Gojjam	1,405,689	868,749	44,975	43,170	4%	33,888	32,831	3%
North Gondar	684,247	199,760	7381	6752	9%	3367	3241	4%
North Shewa (AM)	1,622,197	796,398	62,933	49,301	22%	21,362	18,709	12%
North Wello	1,110,856	421,103	21,367	17,761	17%	8250	7829	5%
Oromia	380,773	67,073	30,875	9877	68%	5285	3405	36%
South Gondar	1,406,698	860,380	72,682	71,704	1%	43,046	42,375	2%
South Wello	1,849,812	850,898	28,215	25,686	9%	16,302	15,403	6%
Wag Hamra	890,004	182,373	447	339	24%	698	601	14%
West Gojjam	1,348,158	955,505	97,206	94,429	3%	71,052	69,116	3%
West Gondar	1,529,197	74,060	6180	681	89%	1342	203	85%
Total	15,229,329	6,590,762	473,155	417,923	12%	276,093	261,350	5%

The effects of cropland filtering on irrigation predictions in Amhara is more asymmetrical, per Table VI-3. In 2020, 12% of positive irrigation predictions are filtered out; only 5% of predictions in 2021 are filtered out. A number of highly agricultural zones – such as Central Gondar, East Gojjam, and West Gojjam – experience very little prediction filtering, on the order of 1-3%. In contrast, the zones with the least amount of cropland (as determined by ESA) see

large reductions in predicted irrigated area, due to a stricter filtering mask: the Oromia zone sees a 68% reduction in positive irrigations predictions in 2020 and a 36% reduction in 2021, while the Wag Hamra zone sees an 89% reduction in 2020 and an 85% reduction in 2021.

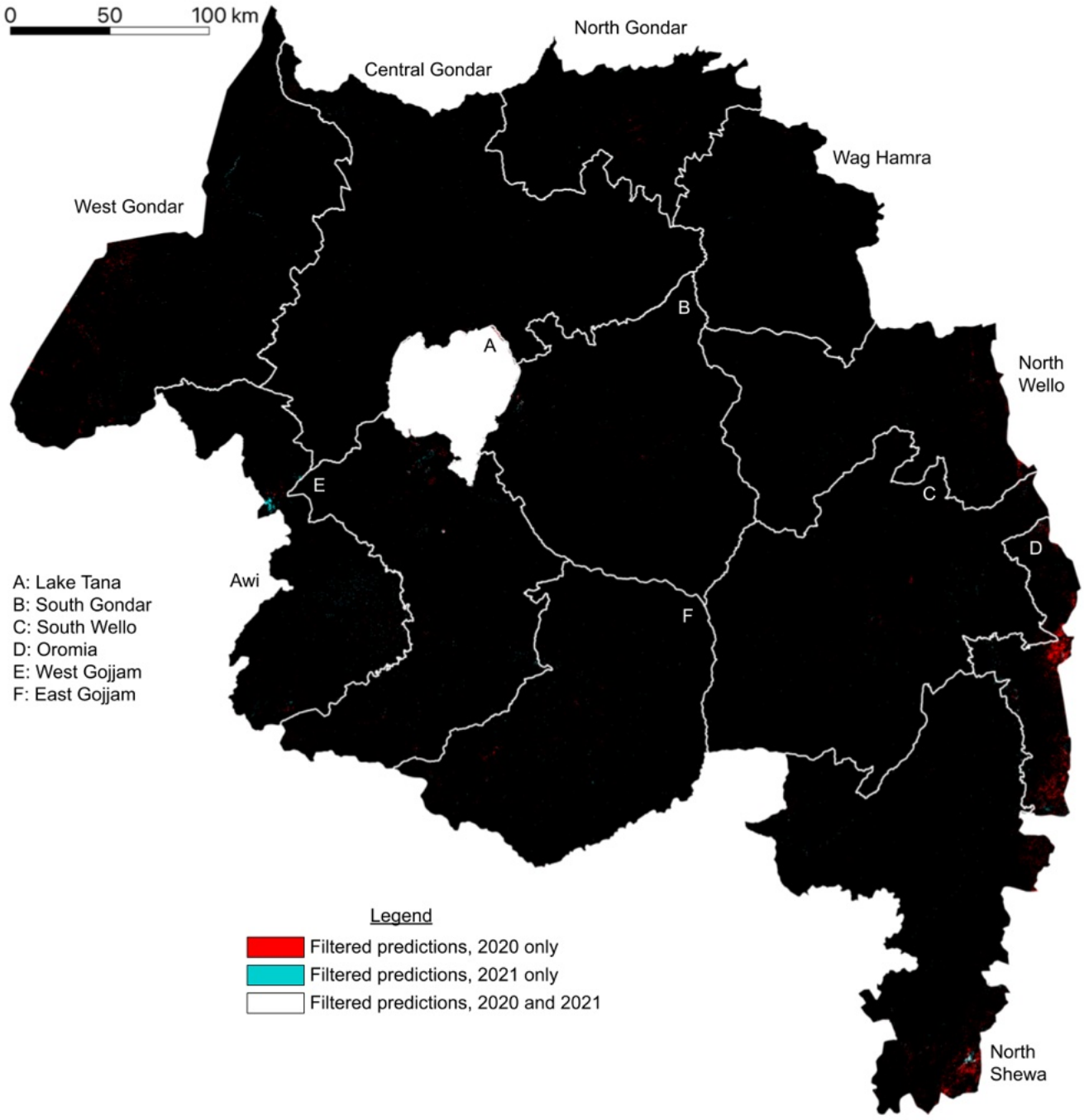

Figure VI-6: Effects of prediction filtering using the European Space Agency WorldCover 2020 v100 map. Many 2020 irrigation predictions in the Oromia and North Shewa zones are removed after filtering.

Figure VI-6 displays the distribution of filtered predictions in Amhara. While image resolution limitations occlude the locations of many filtered predictions, removed positive predictions in the Oromia and North Shewa zones on the Amhara/Afar border are clearly visible. As a reminder, the portion of the Oromia zone that appears as red in Figure VI-6 was the area explored in Chapter VI, Section 1.2. Accordingly, a method of removing these previously discussed misclassified positive predictions is demonstrated. However, a tradeoff will always remain when using auxiliary LC products for irrigation prediction filtering. While cropland filtering can increase user confidence in predictions over diverse land cover types not seen by a classification model, filtered predictions will only be as accurate as the LC products. If an LC map used for filtering has very low accuracy for the classes in question, the resultant filtered predictions will be highly inaccurate themselves.

1.4 Extension of irrigation predictions for Tigray in 2019

To assess whether the significant decrease in irrigated area in Tigray from 2020 to 2021 was in part due to inflated irrigated area predictions in 2020, the transformer-based irrigation detection model from Chapter V is deployed for Tigray in 2019. An identical image collection and prediction process was deployed: S2 enhanced vegetation index (EVI) imagery for the entirety of Tigray was collected at a 10-day timestep between June 1, 2018 and June 1, 2019. The resultant imagery timeseries were then inputted to the trained classifier model to predict irrigation presence. All predicted values above 0.5 are considered irrigated; all those below are considered non-irrigated. Irrigated segments less than 0.1 Ha in size are similarly discarded.

Figure VI-7 contains a tritemporal irrigation map for Tigray, with the new 2019 dry-season predicted irrigation map presented in the red channel of the map, the 2020 predicted irrigation

map in the green channel, and the 2021 predicted irrigation map in the blue channel. This tritemporal map contains substantial color, indicating that the locations of irrigation change across the three years modeled[xxiv]. Yellow and green pixels in the Northwest zone reveal irrigation around the city of Shire, particularly in the 2020 dry season. Red pixels in the South and East zones indicate areas where irrigation existed in 2019 and then shifted to other locations in 2020 and 2021. In fact, the only part of Tigray with consistent irrigation from 2019 to 2021 is an area near the large-scale irrigation project outside of Rama in the north part of the Central zone.

[xxiv] If the locations of irrigation remained constant throughout 2019-2021, predictions would appear as white.

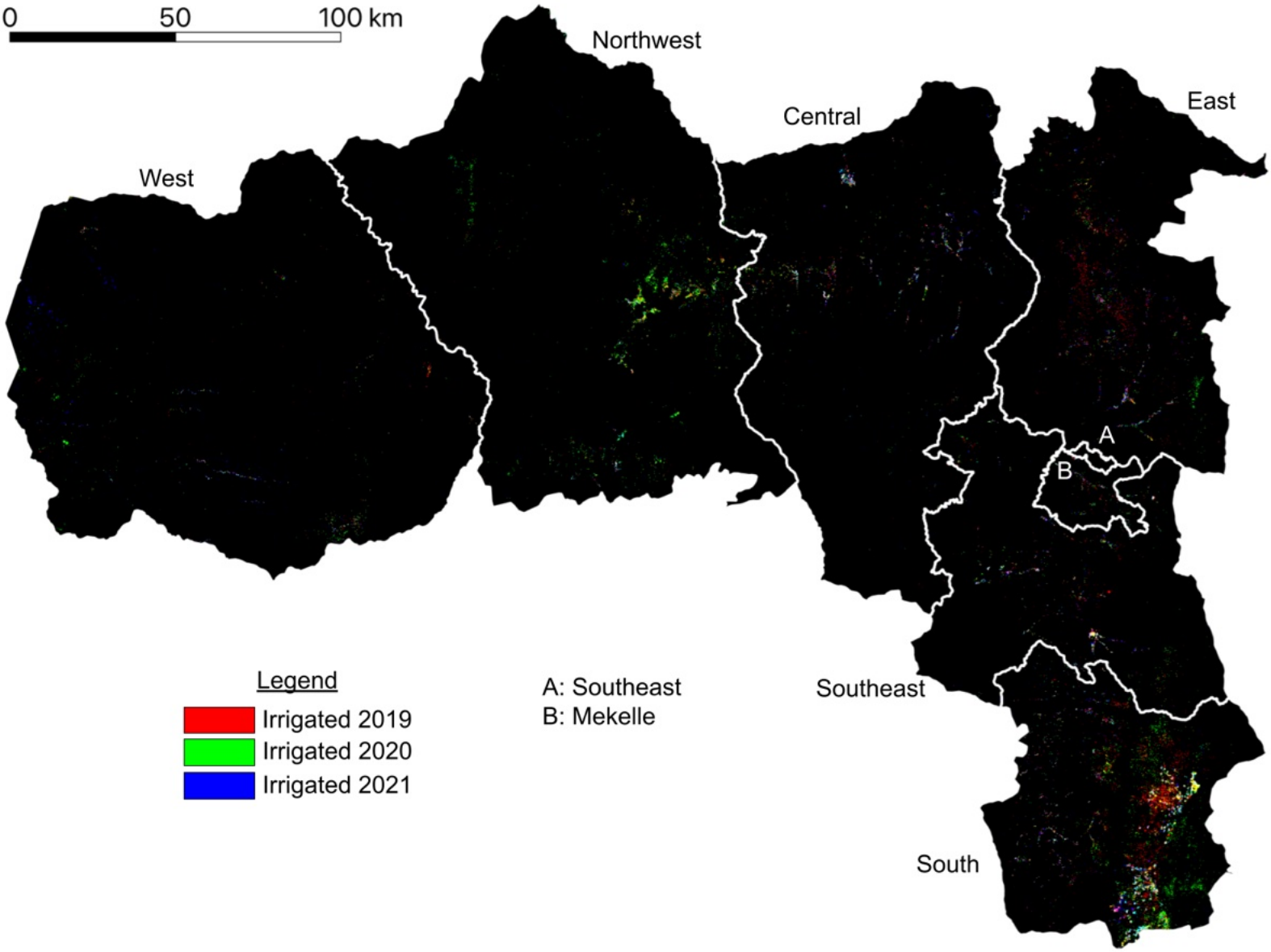

Figure VI-7: Tritemporal irrigation map for Tigray. Irrigation predictions for 2019 are shown in the red channel; predictions for 2020 are shown in the green channel; and predictions for 2021 are shown in the blue channel. Secondary colors indicate positive predictions in two of the three presented years.

Table VI-4: Predicted irrigated area statistics in Tigray for 2019-2021, organized by zone.

Zone	Irrigated Ha., 2019	Irrigated Ha., 2020	Irrigated Ha., 2021	Total Ha.	Percent Change, 2019 to 2020	Percent Change, 2020 to 2021
Central	3454	3710	3554	954,616	7.4%	-4.2%
Eastern	5808	3068	2863	635,670	-47.2%	-6.7%
Mekelle	627	556	397	52,313	-11.4%	-28.5%
Northwestern	3291	7439	2062	1,246,715	126.0%	-72.3%
Southeastern	2876	2658	2301	533,334	-7.6%	-13.4%
Southern	16,625	16,474	8064	506,151	-1.0%	-51.1%
Western	1452	2278	2557	1,331,652	56.9%	12.3%
Total	34,133	36,181	21,799	5,260,451	6.0%	-39.8%

To quantify the amount of irrigated area in Tigray from 2019 to 2021, Table VI-4 presents predicted irrigated areas statistics for the three years modeled. Table VI-4 shows that irrigated area in Tigray increased 6.0% from 2019 to 2020, followed by a -39.8% decline from 2020 to 2021. The quantities contained in Table IV-4 supports insights gleaned from interpreting the predominance of color in Figure VI-6: many zones in Tigray see marked change in irrigated area from 2019 to 2020, and then again from 2020 to 2021. Together, Table IV-4 and Figure VI-6 support the finding that the location of irrigation in Tigray shifts from dry-season to dry-season, similar to the crop rotation patterns shown in the inset of Figure V-13.

Lastly, the modest 6% increase in total irrigated area from 2019 to 2020 corroborates the conclusion that the substantial decrease in irrigated area in Tigray from 2020 to 2021 was not a result of excessive positive predictions in 2020. While the locations of irrigated area changes significantly from 2020 to 2021, the similarity in total statewide quantities implies that these

estimates are closer to the expected amount of irrigation during years when Tigray is not wracked by civil conflict, as was the case in 2021 (and remains to this day).

1.5 Irrigation predictions outside Amhara and Tigray

Spatiotemporal characterization of dominant vegetation phenologies in Ethiopia per the methodology first introduced in [183] yields a map of dry-season irrigation detector applicability. Due to different input vegetation timeseries, a dry-season irrigation classifier trained in the Ethiopian Highlands (Amhara and Tigray states) will not perform well in areas dominated by other types of native vegetation patterns, such as in the Southern Nations, Nationalities, and People's (SNNP) region, which is primarily covered by evergreen vegetation. A straightforward way of summarizing this assumption is to say that irrigation detectors trained over areas in Ethiopia's phenology map that appear as one set of color gradients will only perform well over other areas that present in those same color gradients.

Figure VI-8 displays the same phenology map first presented in Figure V-1. As a reminder, this phenology map is produced by extracting *single cycle*, *double cycle*, *evergreen*, and *non-vegetated* temporal endmembers (tEMs) from a feature space representation of 250m MODIS EVI imagery collected between 2011 and 2021. A temporal mixture model is then applied using these four tEMs, a process that yields tEM abundances that are presented as the phenology map. Figure VI-8 also highlights a valley in southern Oromia near the town of Asasa that contains red color gradients, similar to those in Amhara and Tigray where the previously discussed irrigation detection model was trained, indicating that this area contains vegetation phenologies that approximate the *single cycle* tEM shown in Figure V-1. The inset of Figure VI-8 presents a magnified view of this part of the country: red gradients are clearly visible throughout much of

the valley that is formed by Mount Batu to the south and K'ech'a Terara to the north, with the exception of green pixels surrounding a reservoir directly east of Asasa.

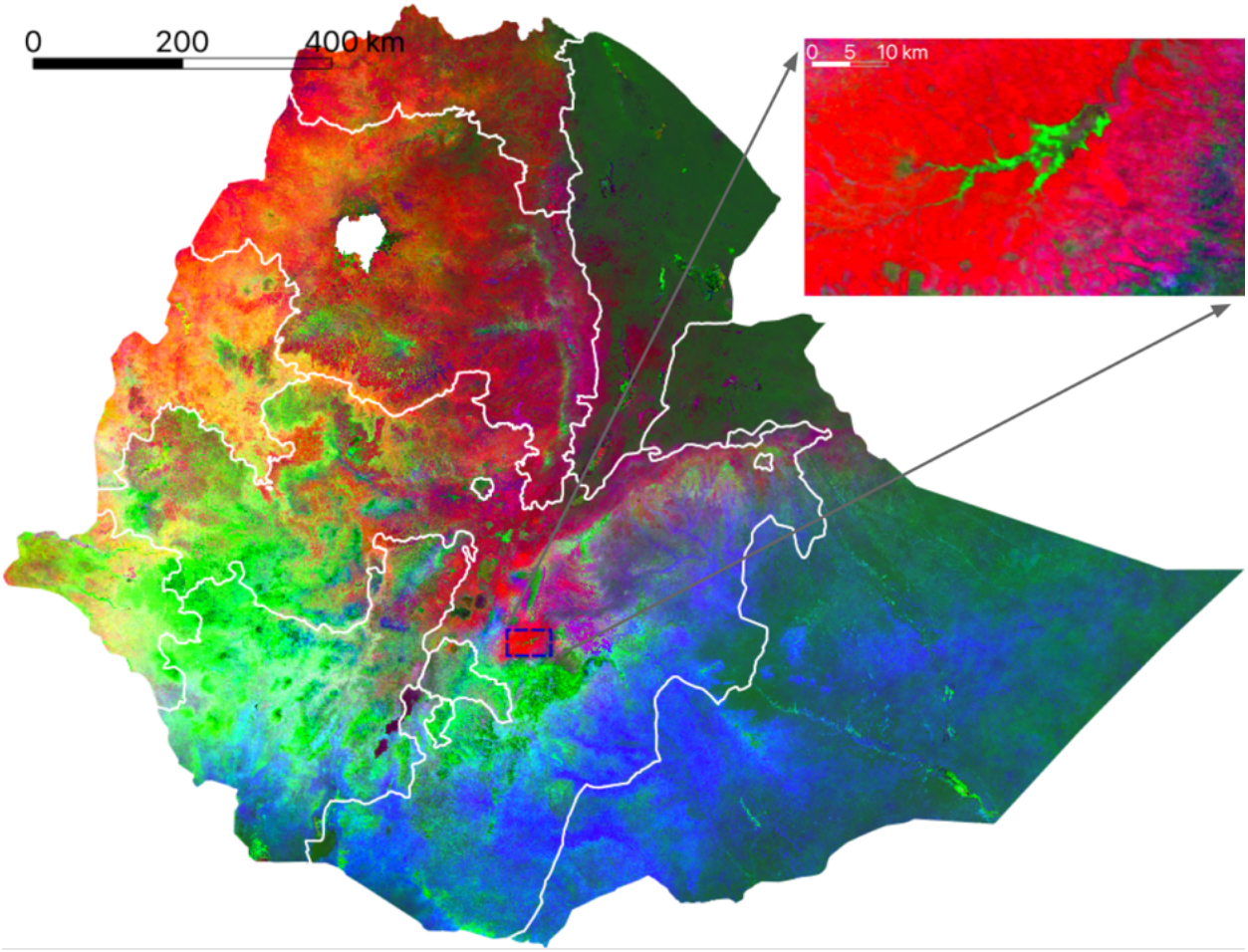

Figure VI-8: Phenology map for Ethiopia, identical to the one first introduced in Figure V-1. Figure inset contains a magnified look at an area near Asasa, Oromia, which primarily contains vegetation cycles that align with the single cycle temporal endmember shown in Figure V-1, as is the case in Amhara and Tigray. Administrative boundaries are shown in white.

Figure VI-9 shows the CHIRPS rainfall and MODIS EVI timeseries for the area contained within the inset of Figure VI-8. Here, CHIRPS rainfall is presented as the average of all values within the inset, and EVI is presented for a single 250m pixel that appears bright red[xxv]. Figure VI-

[xxv] Inspection of other EVI timeseries near Asasa confirms that all pixels that appear red in the inset of Figure VI-8 demonstrate similar phenologies to the one presented in Figure VI-9.

9 displays that the vegetation and precipitation conditions near Asasa are highly similar to those in Tigray and Amhara: Background vegetation cycles once annually with the long *belg* rains that occur from July to September, with smaller kiremt rains taking place beforehand in April and May. As such, the area near Asasa provides another part of Ethiopia in which the irrigation detection classifier can be deployed.

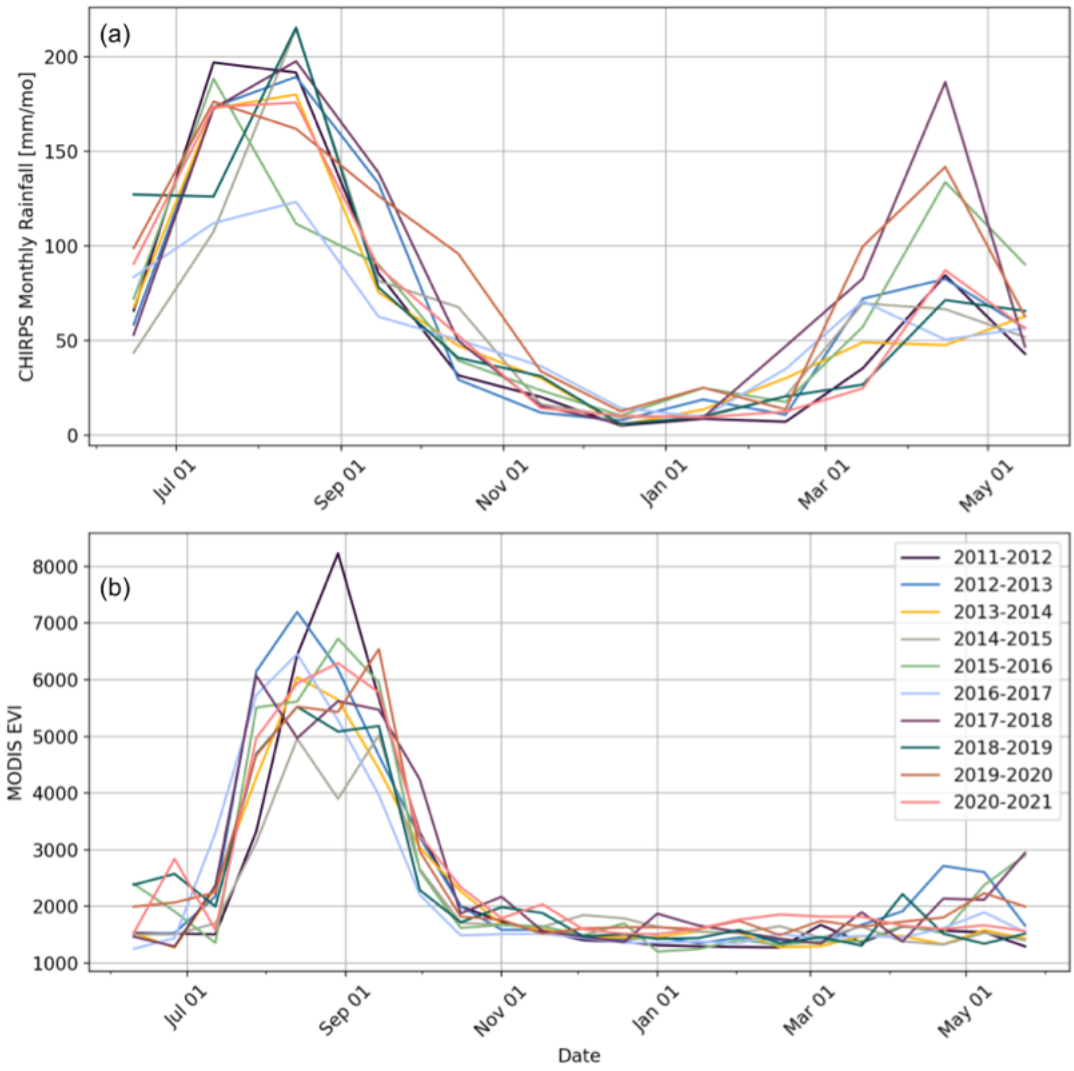

Figure VI-9: (a) Monthly Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) rainfall estimates, and (b) 16-day MODIS enhanced vegetation index (EVI) timeseries for an area near Asasa, Ethiopia, shown in the inset of Figure IV-8. The CHIRPS rainfall timeseries are calculated as the average of the full spatial extent shown in the inset of Figure IV-8, while the MODIS EVI timeseries is presented for a single 250m pixel in the inset that appears bright red.

Figure VI-10(a) presents irrigation predictions for the area contained in the inset of Figure VI-8 for the 2021 dry season. Predictions are made on S2 EVI timeseries with the same irrigation detection model that was 1) introduced in Chapter V and used for inference across Tigray and Amhara, and 2) used to predict irrigation across Tigray in 2019 per Section VI-1.4. In this figure, one can see predicted irrigation along the extent of the Asasa reservoir, and along canals that extend to the south of the reservoir. Such spatial distribution of irrigation is highly similar to that predicted around Lake Tana (see Figure V-13). Figure VI-10(b) plots a series of 8 Sentinel-2 EVI timeseries predicted as irrigation. These plots reveal phenologies highly similar to those predicted as irrigated in Tigray and Amhara, as they contain a full vegetation cycle during the dry months February – April.

Despite the promising initial results, one caveat about applying the irrigation detection model to this part of Ethiopia should be mentioned. Asasa experiences substantial cloud cover for most of the year. Persistent cloud cover prevented any irrigation predictions for the 2020 dry season, as not enough cloud-free imagery could be collected to create the classifier inputs. Moreover, the nearly straight portions of the plots in Figure VI-10(b) from June to December 2020 indicate that many of the 10-day timestep values in this time range were interpolated from a few cloud-free images. These findings serve as a helpful reminder that despite algorithmic sophistication and increasing access to computational resources, any machine learning model that relies on optical imagery will always depend on the availability of cloud-free imagery.

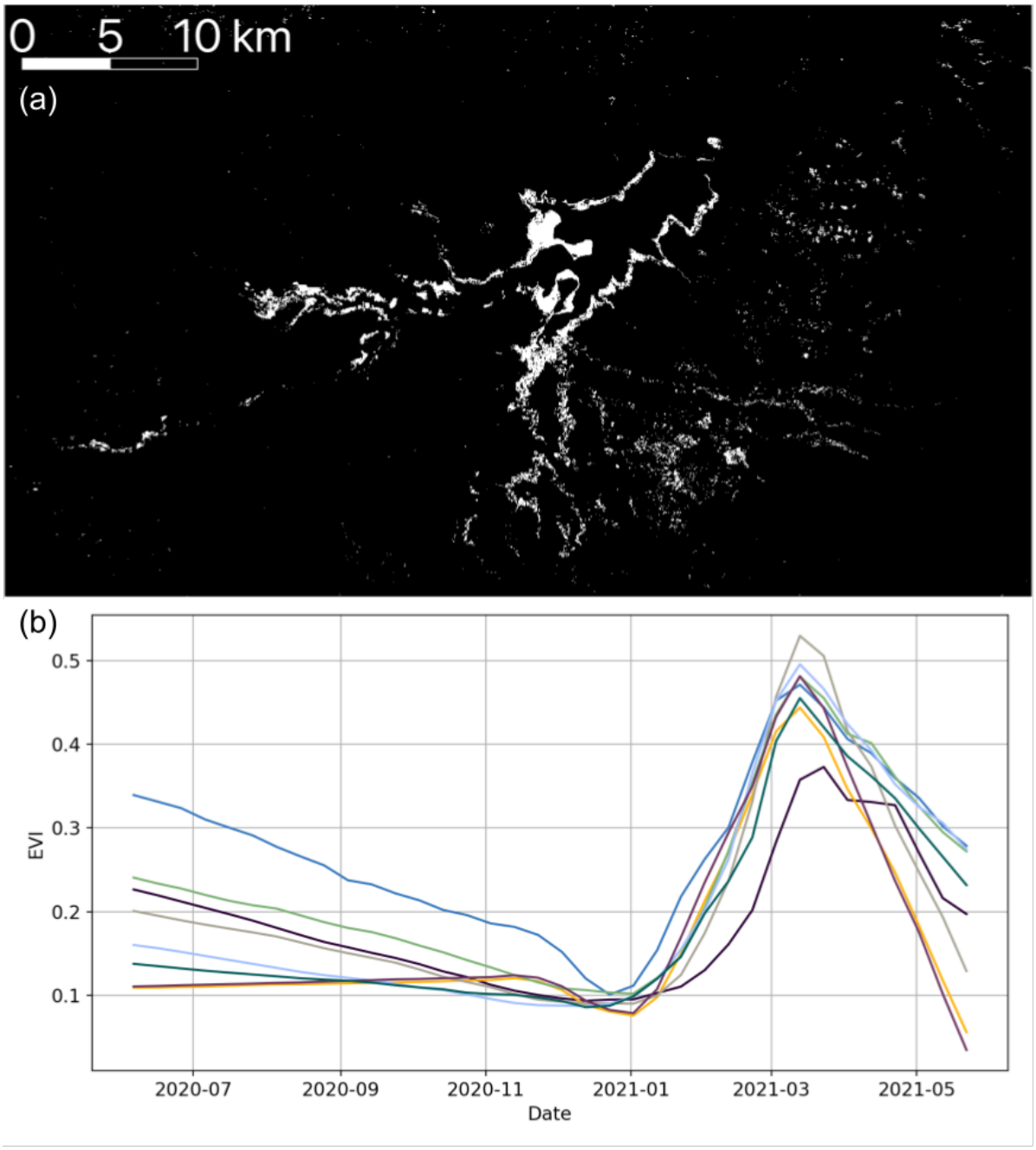

Figure VI-10: (a) Irrigation predictions during the 2021 dry season near Asasa, Oromia. (b) Eight sample EVI timeseries for pixels predicted as irrigated in (a).

1.6 Expansion of irrigation from 2017 to 2021 near Hawzen, Tigray

The S2 mission started collecting imagery in June 2015 after operationalization of the S2A satellite; with the launch of the S2B satellite, dual imaging operations began in June 2017. As such, S2 EVI timeseries can be constructed for the 2017-2021 dry seasons in Ethiopia, providing insight into 5 years of irrigation conditions. It is worth noting that despite imaging capabilities beginning in June 2015, creation of June 2015 – June 2016 EVI timeseries for 2016 dry season irrigation detection in Ethiopia is not possible due to a lack of collected imagery during the first months of operation. Figure VI-11 presents a timeseries of S2 imagery availability, courtesy of Descartes Labs.

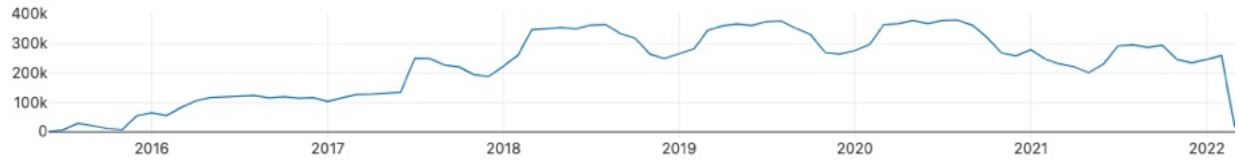

Figure VI-11: Sentinel-2 (S2) image availability as of March 3, 2022, per Descartes Labs.

Given that the developed irrigation detection model requires only EVI timeseries as inputs, imagery from other satellite missions could be used for detection, after interpolation to the correct timesteps. The Landsat missions are of particular interest for this purpose, as their 30m resolution and regular revisit closely approximate the spatial and temporal resolution of the S2 mission. However, Landsat-5, Landsat-7, and Landsat-8 each have unique limitations that preclude them from being used alongside S2 imagery to construct a longer analysis of irrigation presence in Ethiopia. In the case of Landsat-5, a longer gap between image collection timesteps and a lack of regular collection over Ethiopia prevents the download of EVI timeseries at a 10-day

timestep over Tigray or Amhara. Scan line corrector failure on the Landsat-7 missions means that 22% of each collected scene contains invalid information, making the task of pixel interpolation between timesteps even more difficult. Lastly, the Landsat-8 mission began collecting imagery in March 2017, meaning that it provides no new temporal information compared to S2.

Previously established methodology can therefore be applied to assess irrigation change in the Ethiopian Highlands between 2017 and 2021 using S2 imagery. To this end, a 14km-by-22km location near Hawzen in Tigray is selected for longitudinal analysis of irrigation presence. Figure VI-12(a) presents a tritemporal irrigation map for this area, with predicted 2017 irrigation in the red channel, predicted 2019 irrigation in the green channel, and 2021 irrigation in the blue channel; Figure VI-12(b) shows a high-resolution true color image of the landscape taken from Google Satellite Hybrid in QGIS. Table VI-5 contains the predicted irrigated hectares for this area for the 5 dry seasons modeled, demonstrating a general increase in irrigated area from 482.1 Ha in 2017 to 613.3 Ha in 2021.

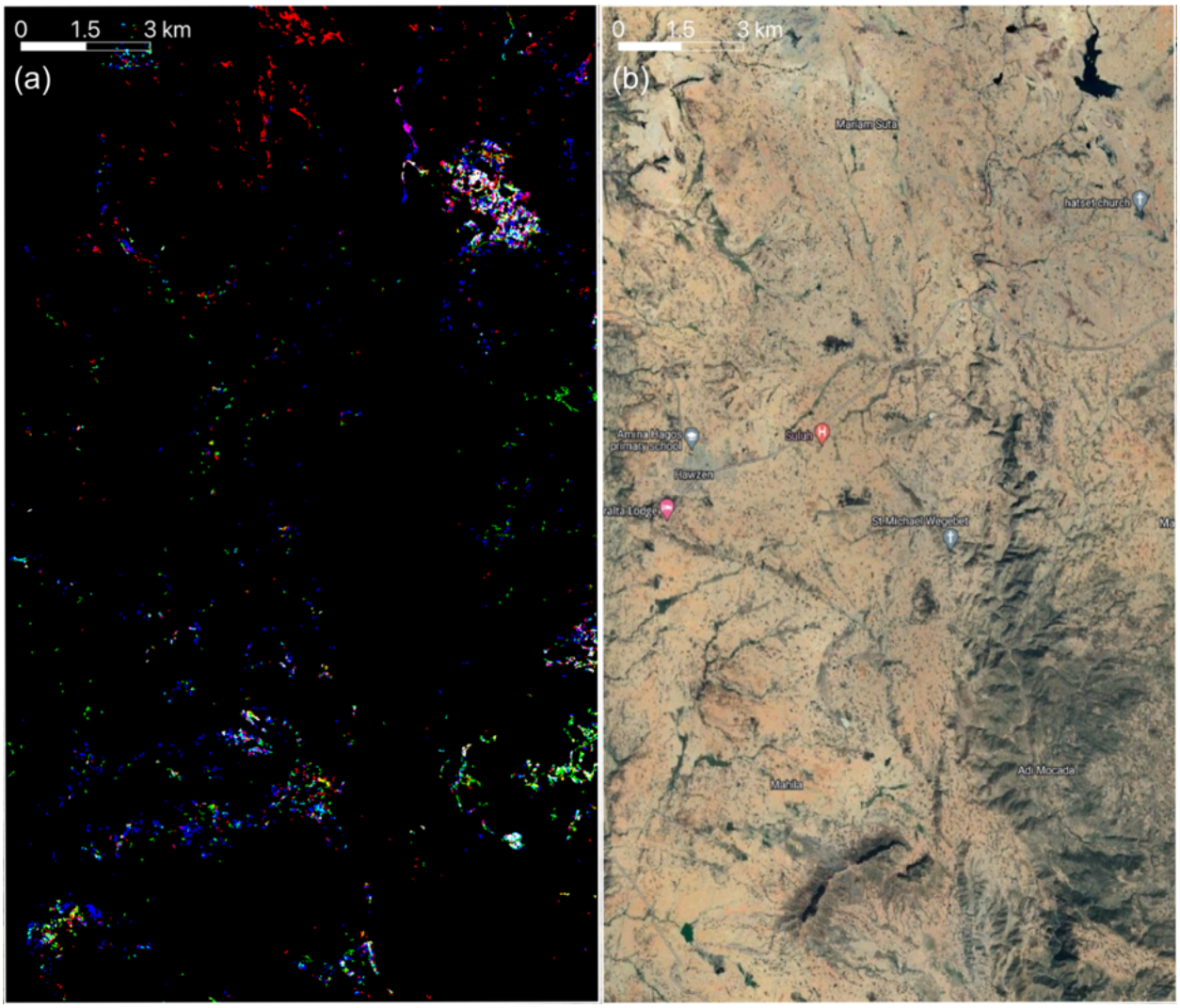

Figure VI-12: (a) Tritemporal irrigation map near Hawzen, Tigray, with 2017 predictions in the red channel, 2019 predictions in the green channel, and 2021 predictions in the blue channel. (b) Background Google Satellite Hybrid imagery collected from QGIS for the same area.

Table VI-5: Predicted annual irrigated hectares for the specified 14km-by-22km area outside Hawzen, Ethiopia.

	2017	2018	2019	2020	2021
Irrigated Hectares	482.1	479.8	523.7	543.2	613.3

While Figure VI-12(a) and Table VI-6 communicate an overall increase in irrigated area outside of Hawzen, a couple caveats should be mentioned. First, as indicated by the predominance of color in Figure VI-12(a), the locations of irrigation in Tigray shift substantially from year to year. It is therefore possible, albeit unlikely, that the increase in irrigation seen in Table VI-5 is not representative of irrigation conditions throughout the entirety of Tigray for this period of time, and instead represent the migration of irrigation from external areas to land nearer to Hawzen. In fact, there are only two portions of Figure VI-12(a) that demonstrate consistent irrigation across the 5 years modeled. The first area is just south of the reservoir in the Northwest of the area of interest, an intuitive finding given that the reservoir will likely ensure water availability for nearby plots year over year. The second area is south of the city Adi Mocada and is much smaller, with no apparent standing body of water nearby. Altogether, Figure VI-12 reinforces a finding first communicated by Figure VI-7: Ignoring the steep decline in irrigated area in 2021 that was likely due to ongoing civil conflict, irrigation in Tigray has generally increased over the modeled years, with large interannual shifts in the locations of irrigation.

1.7 Confirmation of prediction performance in the 2020 irrigation season in Tigray

As model training data and validation polygons collected by independent enumerators are both gathered across Amhara for the 2021 irrigation season, the same irrigation detection classifier first introduced in Chapter V needs to be separately evaluated for the 2020 irrigation season, ideally for a different part of the Ethiopian Highlands. While growing seasons in Ethiopia demonstrate interannual consistency, with irrigated crops being reliably grown during the December through April dry months, confirmation of model performance in another year will further reinforce our findings.

Accordingly, 159 new labeled polygons are collected across Tigray, denoting the presence or absence of irrigation in 2020. Of these 159 polygons, 93 cover irrigated areas and 66 cover non-irrigated areas. These new labeled polygons are not used for model training, but to assess the performance of the existing, trained irrigation detection model. Table VI-6 presents an overview of these polygons.

Table VI-6: Overview of labeled 2020 irrigation polygons in Tigray.

Class	Polygon Count	Total Area	Average Polygon Size
Non-Irrigated	66	13473 Ha	204.8 Ha
Irrigated	93	1399 Ha	15.0 Ha

Table VI-6 shows that while there are fewer collected non-irrigated polygons, these polygons cover nearly an order of magnitude more area. This is due to the nature of non-irrigated agriculture being the dominant method of crop cultivation in Tigray: there is simply more adjacent, non-irrigated cropland in the state than irrigated cropland, thus facilitating the collection of larger non-irrigated polygons.

Figure VI-13 shows the distribution of the collected polygons throughout Tigray. Many of the polygons are gathered in the highly agricultural region of the Northwest Zone near Shire. Other polygons are collected near a large-scale irrigation scheme north of Axum in the Central Zone, and still others are acquired in portion of the rift valley in Tigray's Southern Zone, near the borders with Amhara and Afar.

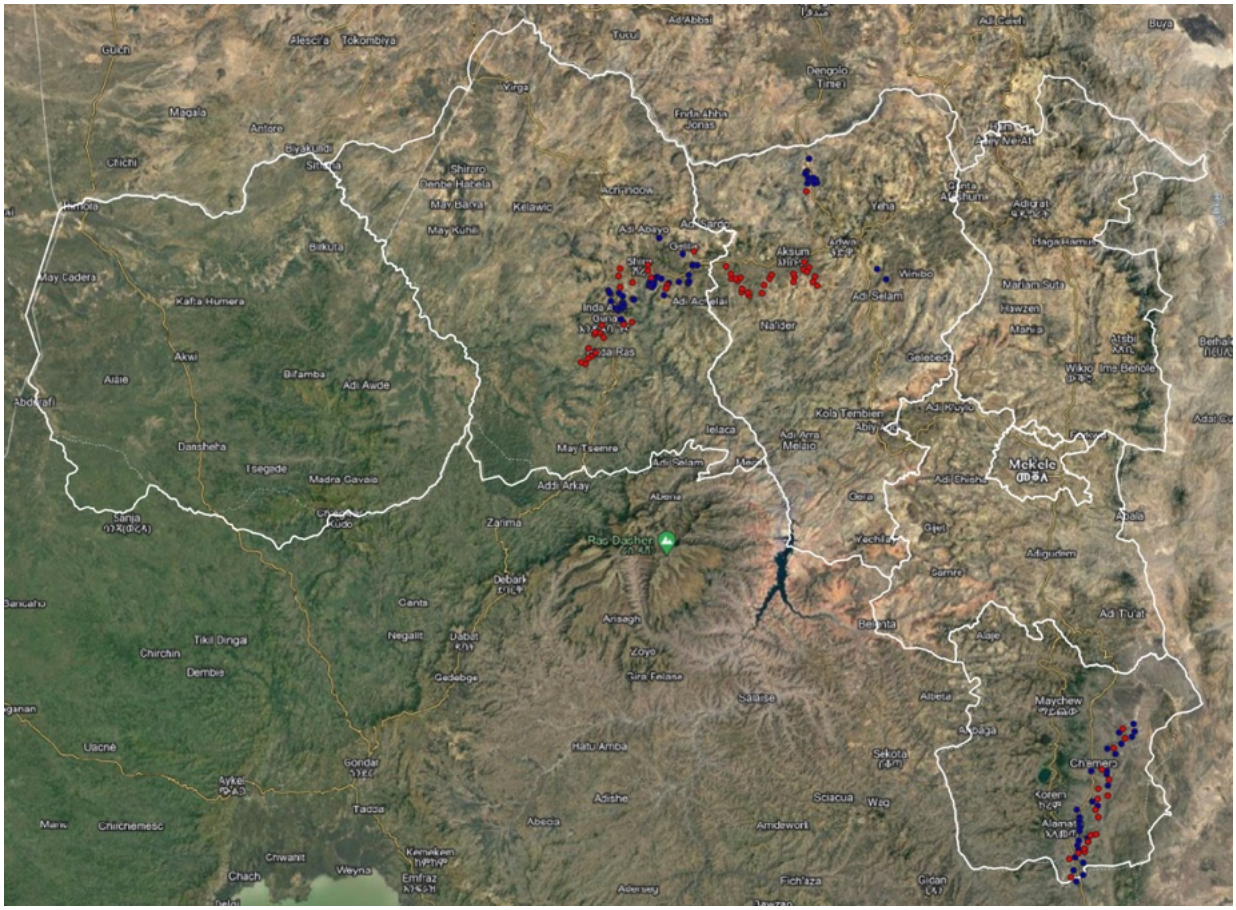

Figure VI-13: Distribution of collected polygons for model validation in 2020. Colored points show the centroids of each polygon, red for non-irrigated and blue for irrigated.

After collecting the labeled polygons, the contained 10m Sentinel-2 EVI timeseries are cluster-cleaned. This process is completed in a manner identical to that introduced in Chapter V. The resulting, cleaned timeseries are presented in Figure VI-14, split into each class's 20 dominant clusters. As a result of the cluster cleaning, 0.2% and 9.5% of the pixel samples in the non-irrigated and irrigated polygons are removed, respectively.

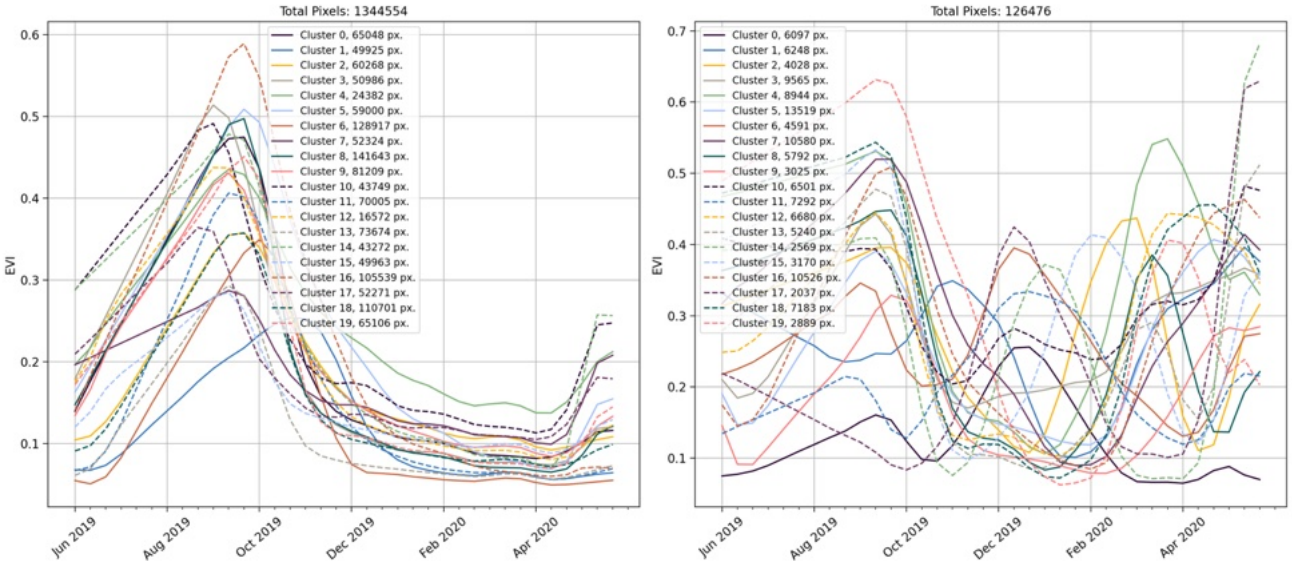

Figure VI-14: Cluster cleaner non-irrigated (left) and irrigated (right) timeseries for the model validation in 2020 across Tigray.

Performance is then assessed over the cluster-cleaned samples for the same model used for inference over Tigray and Amhara in Chapter V, a model which is only trained on labeled data collected across Amhara in 2021. For irrigated samples, the model achieves 96.8% accuracy (122387/126476); for non-irrigated samples, it achieves 98.7% accuracy (1327069/1344554). These performance metrics are consistent with other reported metrics for the model, both over held-out regions during training and over the validation set collected by independent enumerators across the entirety of Amhara. From this additional test, we can conclude that the presented irrigation detection model performs well across Tigray and in the 2020 irrigation season – a year from which no training data was derived. Here, the interannual consistency of irrigated and non-irrigated signatures across Amhara and Tigray (as evidenced by similar red gradients across these states in the phenological map presented in Figure V-1) result in an irrigation detection model that is robust across the Ethiopian Highlands from year to year.

2. Addendum

To ensure replicability and accessibility, I have created open-source alternatives for all portions of the irrigation detection methodology first introduced in Chapter V. In Chapter V, the Descartes Labs platform was used to create the raw imagery timeseries for model training and inference. Subsequently, I replicated this functionality using Google Earth Engine, a platform that is freely available to researchers. To test this new functionality, I performed an end-to-end training and prediction process for a region near Lake Tana in Amhara for the 2020 irrigation season. This process is detailed below, with the resultant irrigation detection models demonstrating up to 99% alignment with predictions introduced in Chapter V. While some prediction variance is expected – the models are trained on different samples – the similarity between predictions 1) on imagery acquired from the Descartes Labs Platform using a model trained on 2021 irrigation data (Chapter V), and 2) on imagery acquired from Google Earth Engine using a model trained on 2020 irrigation data (this addendum) indicates that our irrigation detection methodology remains robust over differences in training data and imagery collection platform.

Application in Amhara for the 2020 irrigation season

To demonstrate the utility of the fully open-source alternative code, irrigation detection is performed for an AOI measuring approximately 100km-by-100km just east of Lake Tana. This AOI is shown in Figure VI-15.

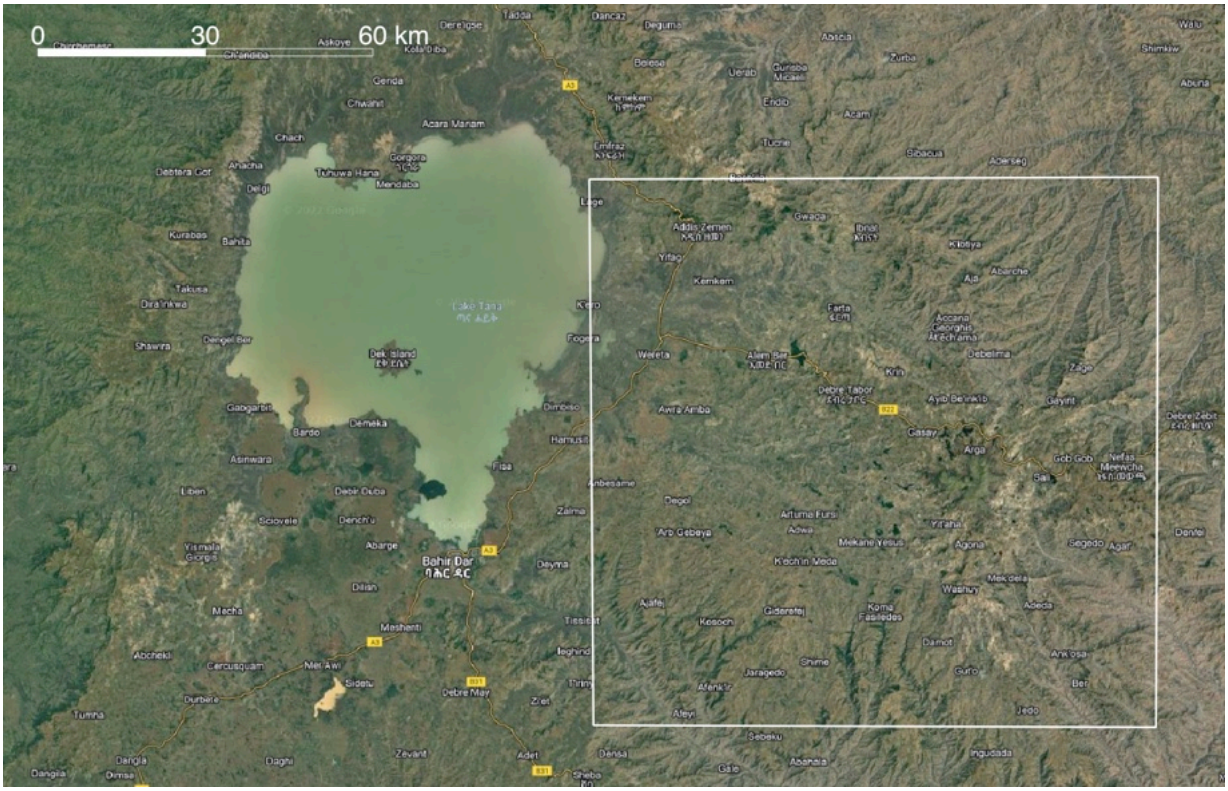

Figure VI-15: AOI for deploying open-source irrigation detection methodology. Imagery provided by Google Satellite Hybrid.

Another member of the Sustainable Engineering Lab collected polygons for model training within this AOI. Here, all polygons were collected for the 2020 irrigation season (December 1, 2019, to April 1, 2020). In total, 122 irrigated polygons measuring an average 0.95 Ha and 101 non-irrigated polygons measuring an average 5.15 Ha are acquired. After generating Sentinel-2 (S2) imagery for these polygons and then cluster-cleaning the resultant timeseries, we are left with 7099 irrigated pixel timeseries and 52,095 non-irrigated pixel timeseries for model training. Figures VI-16 and VI-17 present the clusters for the non-irrigated and irrigated cleaned timeseries. Readers are referred to Chapter V, Section 3.3.2 for full details on the cluster-cleaning process.

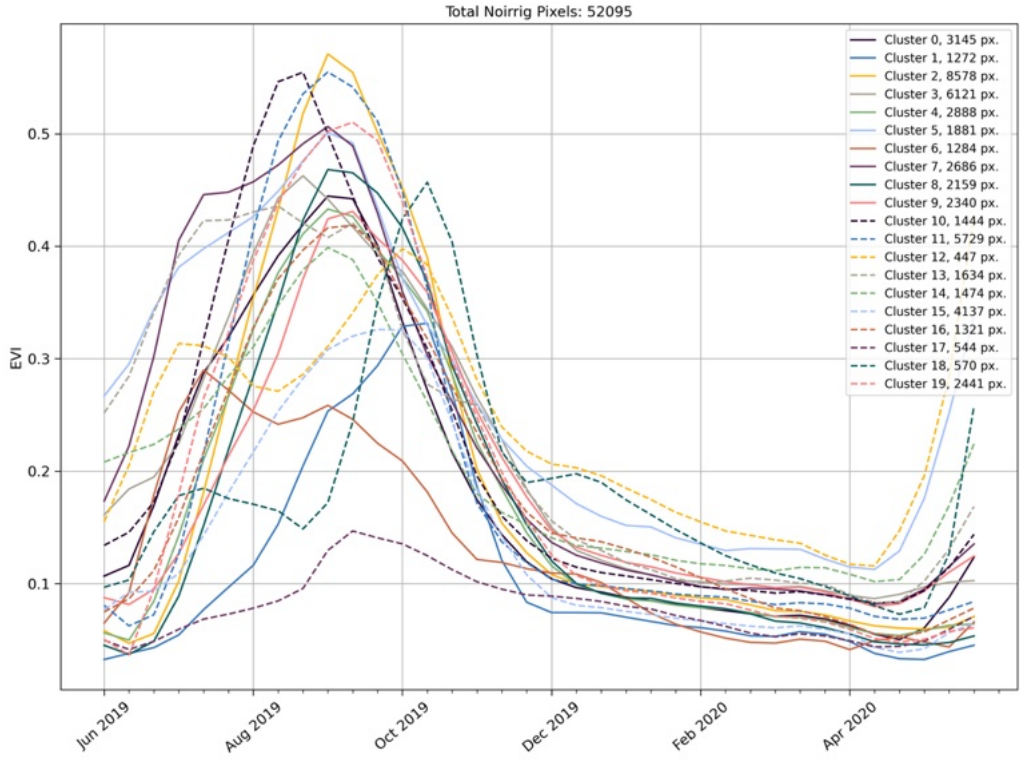

Figure VI-16: Cluster-cleaned non-irrigated EVI timeseries derived from 101 hand-labeled polygons.

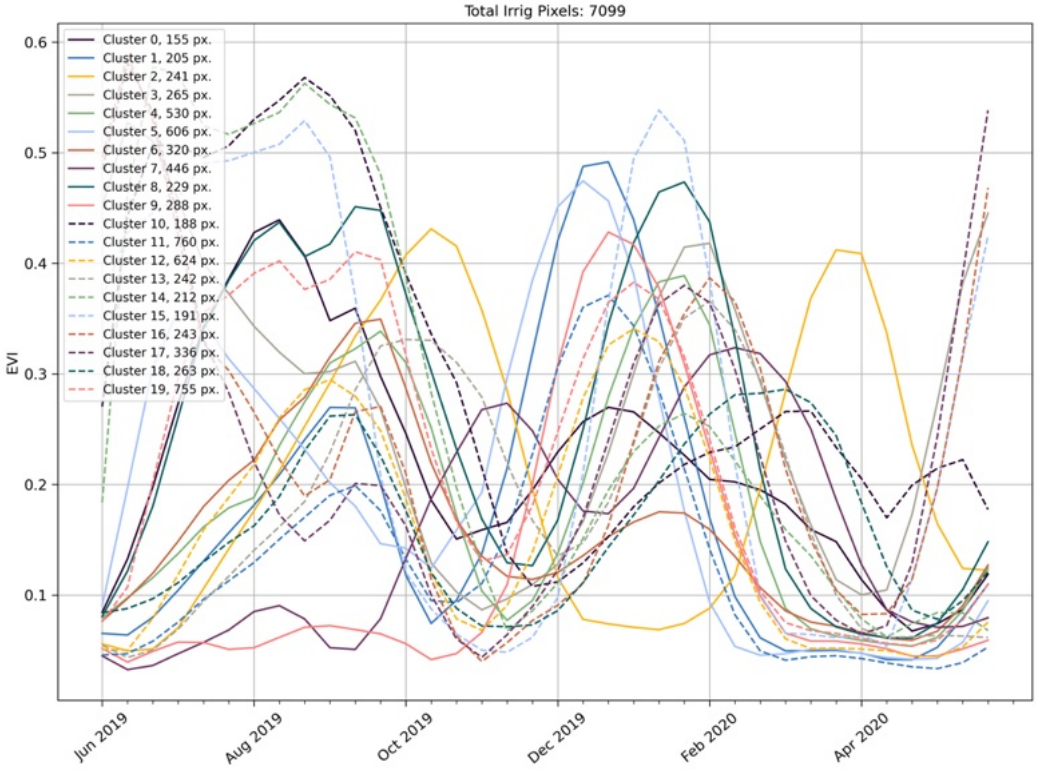

Figure VI-17: Cluster-cleaned irrigated EVI timeseries derived from 122 hand-labeled polygons.

The breakdown of polygons and pixels across model training/validation/testing configurations is presented in Table VI-7. Here, polygons are split by a 70/15/15 ratio, with each configuration's polygons yielding the cleaned pixel timeseries for that grouping.

Table VI-7: Summary of labeled data, split by model training configuration. Pixel data is derived from cluster cleaning the labeled polygons.

	Number of Labeled Polygons							
Label Type	Training		Validation		Testing		Total	
	Non-Irrigated	Irrigated	Non-Irrigated	Irrigated	Non-Irrigated	Irrigated	Non-Irrigated	Irrigated
Polygons	70	85	16	19	15	18	101	122
Cleaned Pixels	37580	4363	6751	1633	7764	1103	52095	7099

The cleaned pixels are then used for model training. Training is performed on both the Catboost and Transformer architectures, the two highest performing model types explored in Chapter V. As in Chapter V, each model is trained for 30 epochs, with validation and test accuracy assessed over pixel timeseries from polygons withheld from the training set. Only data from the new Tana AOI are used for model training and evaluation. Table VI-8 presents the training results from for these two model architectures.

Table VI-8: Model training performance over the new Tana AOI.

	Testing Dataset Performance		
	Non-Irrigated Sample Accuracy	Irrigated Sample Accuracy	F1 Score
Catboost	0.99 (7662/7764)	0.98 (1081/1103)	0.95
Transformer	0.97 (7545/7764)	0.97 (1071/1103)	0.90

From Table VI-8, we see that both models accurately predict irrigation presence over the reserved test samples, with all class accuracies above 0.97. The Catboost model performs slightly better than the Transformer model, although this margin is small.

Once trained, we deploy these irrigation detection models over the entirety of the new Tana AOI and compare predictions to those generated for the 2020 irrigation season in Figure V-14. Here we observe that Catboost predictions have 94.5% agreement over all pixels (114.7M/121.4M) and that Transformer predictions have 99.1% agreement (120.3M/121.4M). This finding reveals that the transformer predictions will be more robust over a wider extent, as they are more similar to predictions made by a model trained on data from across the Ethiopian Highlands (Chapter V).

Figure VI-18 shows the 2020 irrigation predictions from this addendum's Transformer model next to a high-resolution image of the AOI from Bing Aerial. This figure reveals a high concentration of irrigation predictions in the flood plains directly east of Lake Tana, as is expected.

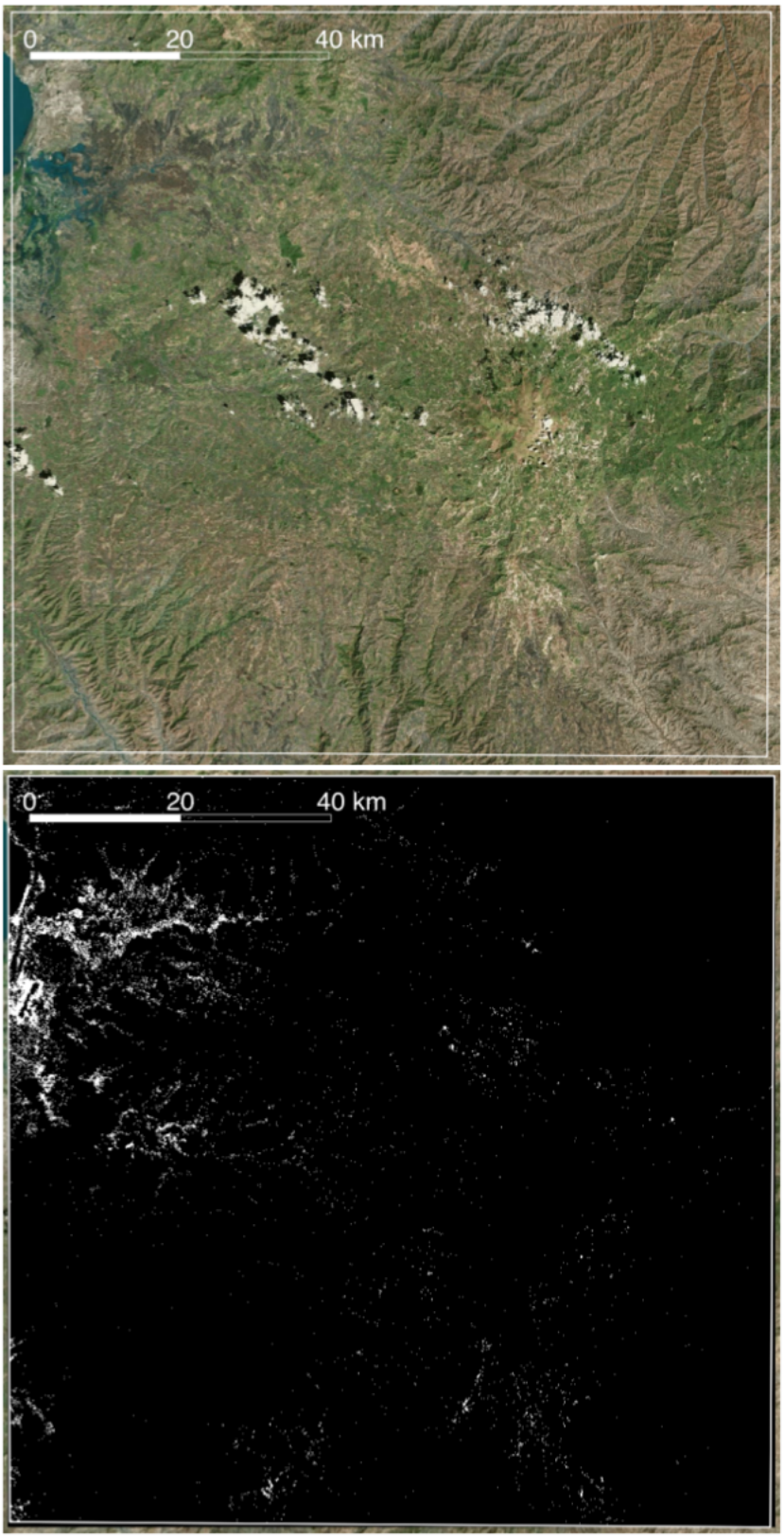

Figure VI-18: (a) Bing Aerial high resolution imagery, and (b) 2020 irrigation predictions from this addendum's Transformer model for the new Tana AOI.

Summary of new scripts

There are five new Google Colaboratory notebooks that contain the code necessary to replicate the irrigation detection methodology in a fully open-source manner. All code, imagery, predictions, and saved model files are stored at *gs://gee_irrigation_detection.*[xxvi]

The first notebook – *0_s2_imagery_upload* – creates the 36-timestep timeseries of S2 imagery using Google Earth Engine. These uploaded images are neither smoothed nor interpolated, and accordingly have many missing pixels due to cloud cover. This notebook generated imagery for both training (i.e., for smaller, hand-labeled polygons), and for inference (i.e., for a larger continuous extent). For inference, the notebook will split a larger AOI into tiles approximately 2500px by 2500px.

The second notebook – *1_clean_labeled_data* – takes imagery uploaded over hand-labeled polygons for training, converts the S2 timeseries within the polygons to CSVs, and then performs cluster-cleaning. There is also a function at the end of the notebook for converting these cleaned CSVs to tfrecords for model training. Next, *2_smooth_inference_imagery* smooths and interpolates the S2 imagery timeseries uploaded for inference.

Model inference is performed by *3_irrigation_detection_inference*. This notebook loads a pretrained model, and then uses it to predict over the smoothed, interpolated inference timeseries. Lastly, *4_utils* contains helper functions for the prediction process, including 1) a function that converts saved Transformer models to a format that can be more easily uploaded and accessed, and 2) a function that compares model predictions.

[xxvi] As this is a Google Storage Bucket affiliated with a Columbia project, it cannot be made public. However, access to this bucket can be granted upon request.

Together with the training scripts on a Google Cloud Platform virtual machine and the Earth Engine script for collecting hand-labeled polygons (also saved in *gs://gee_irrigation_detection*), these 5 notebooks provide all the functionality necessary for replicating the irrigation detection methodology in an open-source manner.

Notes on Google Earth Engine compute timing

Google Earth Engine has internal compute restrictions to ensure that resources are distributed fairly across users. Unfortunately, these restrictions are opaque, and it is difficult to accurately estimate the time jobs will take. To provide timing estimates to future researchers, I recorded how long portions of the new GEE methodology took for this addendum's application:

- Generating 0.25 degree by 0.25 degree (approx. 2500px by 2500px) Sentinel-2 EVI timeseries for inference takes approximately 45 minutes each.
- Generating Sentinel-2 EVI timeseries over hand-labeled polygons for training takes approximately 1:40 minutes per polygon.
- Predicting over a 0.25 degree by 0.25 degree Sentinel-2 EVI timeseries using the Catboost model takes approximately 1:20 minutes each.
- Predicting over a 0.25 degree by 0.25 degree Sentinel-2 EVI timeseries using the Transformer model (with GPU acceleration enabled on Google Colaboratory) takes approximately 11 minutes each (without the GPU, this takes 5x as long).

Chapter VII: Contextualizing the impact of productive electricity demand on energy system planning

Abstract

In all electricity system planning efforts, the shape and nature of electricity demands remain paramount. Among the different types of electricity demand, productive demand – demand for agricultural, commercial, or industrial activities – has substantial potential for growth across sub-Saharan Africa generally and Ethiopia specifically. Moreover, the size and potential flexibility of this productive demand ensures its importance for future electrification efforts.

To connect predicted irrigated area back to its energy system planning implications, a linear program called the Irrigated System Electrification (ISE) model is introduced for exploring least-cost methods of satisfying combined residential and productive electricity demand for various system configurations. Model simulations indicate that interconnected systems with access to grid power yield the lowest levelized costs of electricity at $0.164/kWh. In the absence of grid power, a system of solar, battery, and diesel that facilitates irrigation over a 300m radius area corresponds to a levelized cost of electricity of $0.272/kWh; analogous plot-level (0.2 ha) systems can meet demand at a cost of $0.397/kWh.

1. Electricity demand types and energy system planning

Energy services, the benefits that energy carriers produce for human well-being, are a critical foundation for sustainable growth and the maintenance of social welfare [307]. Higher quality, more accessible energy services provide significant benefits on an individual level and in the aggregate: A 2002 study found the total benefit of providing electricity to a typical,

nonelectrified Philippine household to be $105–$195 (USD 2019) per month [308], while higher annual electricity consumption correlates to increased national GDP per-capita, as is shown in Figure VII-1 [309].

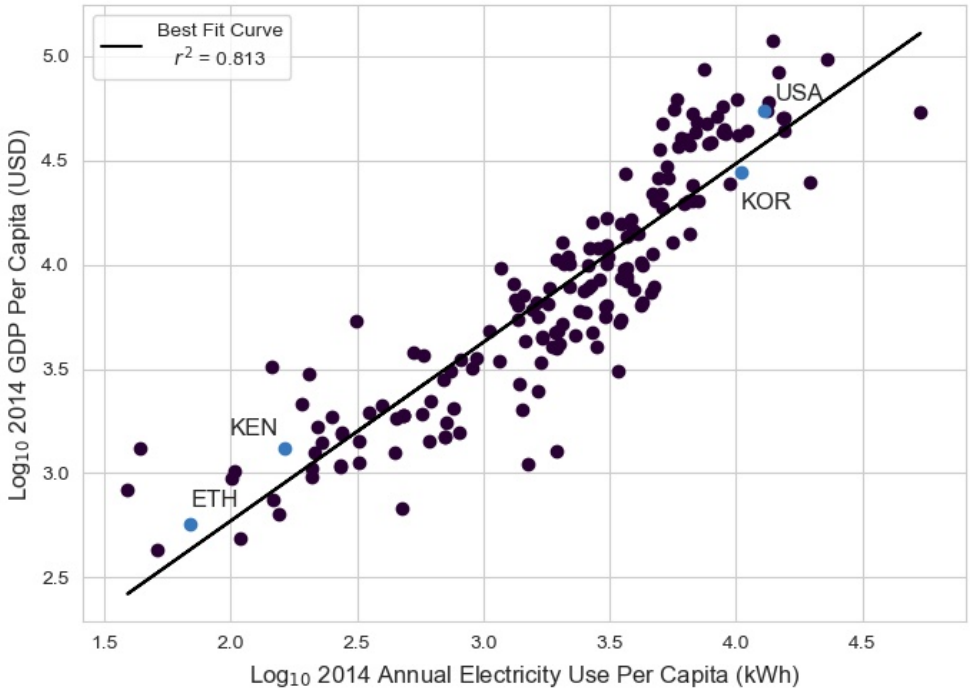

Figure VII-1: Log-log plot of GDP per capita (USD) vs. annual electricity consumption (kWh), 2014 [309].

Energy systems – defined as all components related to the production, conversion, delivery, and use of energy – provide energy services [310]. Electricity systems, components of energy systems more broadly, are of particular importance in delivering energy services: Electricity grids supply consumers with affordable energy, which, depending on the modernity of the system and the fuel mix of grid-connected generators, can be clean, reliable, and delivered at the flick of a switch. Given that electricity systems are complex, massive organisms upon whose reliable function billions of people depend, they cannot be experimented upon in real time; as a

result, computational models provide a valuable tool to assess future energy system scenarios under a wide range of assumptions [311]. However, electricity system modeling necessarily involves a large amount of simplification, and one must be careful that at the end of the modeling process, the model faithfully represents the original system [312]. Because models will, at best, accurately simulate a version of real-life electricity systems, their value must be in giving insights, not strict quantitative predictions.

Nevertheless, system models can facilitate informed policy and planning strategies, a prerequisite for sustainable development. Electricity models have been created for a vast array of settings; they vary in terms of their overall objective, considered available technologies, temporal and spatial resolutions, time-periods modeled, inclusion of existing infrastructure and physical constraints, and considered costs [313]. As a result of differences in scoping and intent, there exists a significant gulf in electricity system planning approaches and tools used in higher-income and lower-income settings, distinctions which are discussed below.

1.1 Electricity system planning in higher-income settings

As explored in Part I of this dissertation, electricity system planning in higher-income settings is focused primarily on decarbonization [314]. More specifically, this means realizing ambitious renewable electricity targets [1], meeting renewable generation capacity goals [141], and prioritizing electrification of liquid fuel-intensive energy end-uses, including heating and transport [315]. In planned transitions away from carbon-intensive electricity systems, researchers have identified the crucial role of system flexibility [316]. Future systems will likely depend on non-dispatchable energy sources such as wind and solar generation; accordingly, there will be a need for battery storage, demand response, fast-ramping generation capacity, and

other flexibility measures to integrate these low-carbon alternatives [317]. As a result of the efforts of planners, researchers, policy-makers, business leaders, and a suite of other engaged individuals, there has been a rich, public dialogue over the tradeoffs of certain decarbonization strategies, one that is currently shaping energy system policy in higher-income settings: In the US alone, seven states have recently passed major climate legislation [3].

1.2 Electricity system planning in lower-income settings

Electricity system development will have a significant impact on the growth of lower-income settings. Despite many parts of the world requiring better access to electricity, the following discussion of system planning in lower-income settings will focus on the African context, as this is where Professor Modi and the Sustainable Engineering Lab's (SEL) recent work has focused. With 20% of the world's population, Africa accounts for only 6% of global energy consumption and 3% of electricity demand [177]. In the coming decades, electricity networks are likely to undergo wholesale transformation, as they will both need to provide power to the 700 million people across sub-Saharan Africa currently living without electricity access and to accommodate the 4-fold increase in overall demand that is predicted for 2040 [177].

Excluding South Africa, the entire installed electricity generation capacity of sub-Saharan Africa is 28 GW, equivalent to that of Argentina [318]. Electricity access also lags stated goals: As of 2018, sub-Saharan Africa has an electrification rate of 45% [177] and an average per-capita electricity consumption of 488 kWh per year, or 5% of average US consumption [319]. While this is a comparatively low access rate – 94% of the population in developing countries in Asia has an electricity connection – significant progress has been made in recent years: The number of people

gaining access to electricity has doubled from 9 million per year between 2000 and 2013 to 20 million per year between 2014 and 2018.

The reasons for incomplete electricity access in Africa include a lack of generation capacity; an absence of proper grid infrastructure; regulatory impediments to maintaining and investing in infrastructure; and a large portion of the population living in remote areas [319]. An inability to provide reliable electricity has also led to the growth of expensive, on-site diesel generators for commercial, industrial, and even residential end-uses. On the grid side, a lack of planning and operational maintenance has resulted in power systems that average 18% transmission and distribution losses, which in turn has increased dependence on large dams and expensive diesel plants. Furthermore, growing demand for transport fuels and inefficient refineries have made Africa the largest importer of refined fuel products in the world; despite recent discoveries of substantial natural gas reserves, especially in East Africa, African natural gas production remains 6% of the global total [177]. This high dependence on imported fuel creates the dual problems of supply and price variability: Electricity producers will curtail supply under low-price conditions, while consumers will suffer economic losses during high-price periods [319].

1.3 Types of electricity demand by end-use

Electricity demand is usually divided into residential, commercial, industrial, and transportation sector contributions. Within these sectors, multiple electricity end-uses exist that differ in terms of their scale, shape, flexibility, and penetration in various regions of the world. We can also subdivide types of electricity consumption into existing and latent demands: Existing demand, representing current consumption; and latent demand, a quantification of future

electricity consumption that will exist once other energy end-uses are electrified, or electricity access is improved. Some electricity end-uses are particularly relevant to this dissertation are characterized below.

1.3.1 Lighting, refrigeration, and household appliances

Lighting, refrigeration, and household appliances make up a significant portion of residential and commercial electricity consumption. In the US, the EIA estimates that lighting accounts for 8% of total electricity consumption [320]; refrigeration accounts for 6% of residential electricity, which as a sector, constitutes 38% of total electricity consumption [321]. In the absence of the rapid advancement of appliance and demand-side management technologies, these electricity demands will remain non-schedulable, requiring other portions of the electricity system to respond flexibly to their presence.

In lower-income settings, these basic energy services are often the first met by a new electricity connection, as Professor Modi has shown in previous projects [322,323]. Accordingly, substantial latent demand is predicted for residential electricity exists in sub-Saharan Africa; however, until adoption of high-consumption household appliances like air-conditioners is widespread, residential demand will likely remain relatively small, and in many cases, will not economically justify a grid connection.

1.3.2 Heating and cooling

A recent report from SEL alumnus Mike Waite and Professor Modi found that 30% of US heating energy demands are being satisfied by electricity, either through resistive space heating or high-efficiency electric heat pumps; fossil fuels supply the remaining 70% of demand [324].

Given that emissions from electric heating will track the overall emissions rate of the grid, pairing renewable electricity generation with heat pumps offers the only realistic path to substantial decarbonization of the heating sector. Transitioning current heating infrastructure to electric options will require significant capital investments; managing the shape of the demand profile will also be critical for system planning considerations, as a large increase in peak demand will require significant buildouts of peaker plants and grid infrastructure. In the US, full electrification will increase aggregate peak loads by 70%, while retaining some fossil fuel backup supports 97% of heating electrification without new peak loads [324].

In 2019, electricity use for space cooling in US residential and commercial sectors was 10% of overall electricity consumption [320]. Cooling demand is climate and weather-driven: Hotter regions of the world measure more cooling degree days (CDDs), an index for overall cooling energy requirements. While the number of annual population-weighted CDDs in the US has generally increased since 1950, in part due to a warming planet, electricity use for cooling has skyrocketed in tropical countries that have recently achieved middle-income status: Residential electricity consumption has increased 2500% since 1990 [309], while air conditioners sales grew by 160% from just 2012-2016 [325]. It is likely that sub-Saharan Africa will see similar increases in electricity demand for cooling in the coming decades as economic growth continues and the region becomes even warmer.

1.3.3 Productive uses

Productive demand refers to electricity requirements for agricultural, commercial, or industrial activities. More than other types of electricity load, productive demand is used to enhance health and income [326]. In sub-Saharan Africa, where there is substantial need for

modern energy services to improve health outcomes and provide economic opportunity, the IEA predicts productive demand will quadruple by 2040, growth which has serious implications for future electricity systems.

Of the many productive electricity end-uses, Part II of this dissertation focuses on electricity demand for irrigation, which plays an important role in improving agricultural productivity and increasing incomes [177]. As expected, irrigation is more beneficial during periods of drought or inconsistent rainfall; as climate conditions continue to become more extreme and variable, reliable access to water is an important step in improving food security [327]. A study in India found that agricultural electricity demand is price elastic: Affordable, grid-priced power can facilitate a large increase in consumption, while higher cost, less reliable options can severely depress the adoption and use of electricity for productive purposes [328]. Relatedly, researchers in Ethiopia founds that the average income of non-irrigating households was half that of irrigating households [181]. These studies reinforce the need for rigorous planning of grid extensions and other modern energy services that deliver electricity inexpensively, for only then will individuals be able to realize health and income gains from increased productive electricity consumption.

1.4 Impact of productive electricity demand on energy system economics

A more concrete example illustrates how productive demand impacts the viability of electricity provision services. Figure VII-2 presents a representative residential electricity demand curve for a consumer in a lower-income setting – this is not a formal electricity demand curve, but an estimated one based on previous SEL experience. Here, the intersections of the dashed lines represent costs and consumptions associated with various electrification technologies. At

the top of the figure, a small solar home system can produce 1 kWh of electricity per month at a cost of 4 dollars. Consumption increases as price drops along the curve until the last displayed point, which shows the consumption and cost associated with a grid connection.

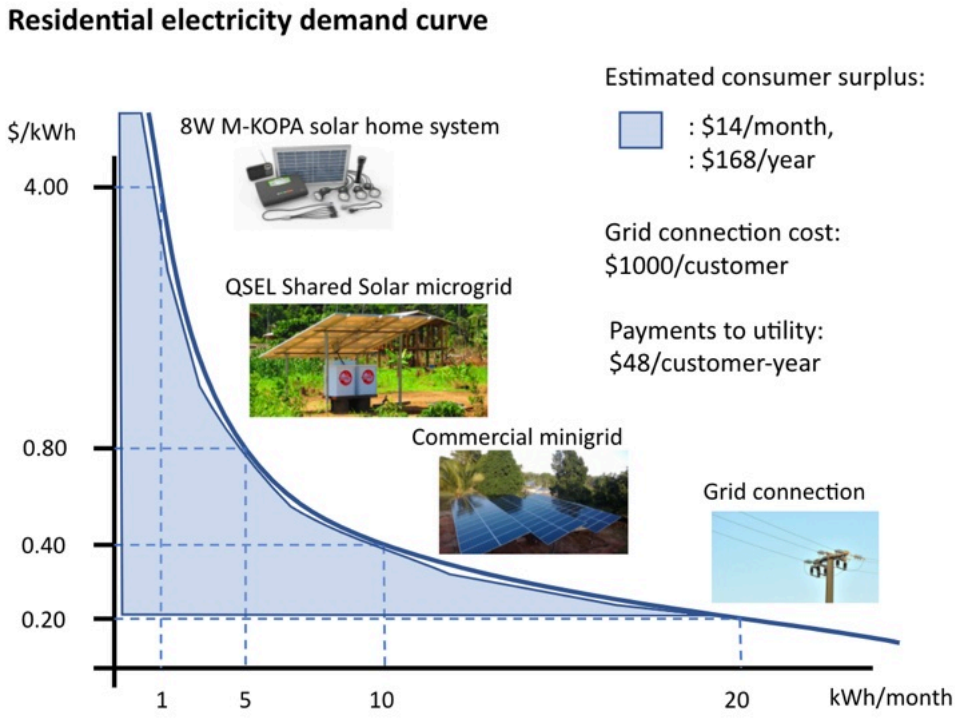

Figure VII-2: Simulated residential electricity demand curve. Likely electricity service mechanisms for paired electricity costs ($/kWh) and electricity consumption quantities ($/kWh) are shown next to the demand curve, along with estimated consumer surplus, grid connection costs, and annual payments to the utility.

From Figure VII-2, the consumer surplus of having grid-price electricity available is $168/year. This quantity can be compared to a grid connection cost of approximately $1000/customer, and the electricity payments received by the utility, $48 dollars per customer per year. As such, extending the grid to a consumer with this electricity demand profile is a risky proposition for utilities, who under the best of circumstances, will recoup their investment in 20 years.

In contrast, Figure VII-3 displays the demand curve for a customer with both residential and productive electricity requirements. Compared to a hypothetical customer with only a residential electricity load, this demand curve shows that consumption increases more quickly as electricity cost decreases. With increased demand, grid-priced electricity results in a consumer surplus of $288/year and payments to the utility of $120/customer-year, quantities that are 2x and 2.5x larger than their respective values in Figure VII-2. Accordingly, the utility's initial $1000 investment in a grid-connection will be repaid in approximately 8-9 years instead of 20. By first providing grid connections to areas with productive demand, utilities can recover investments more quickly, a process that allows for continued grid expansion.

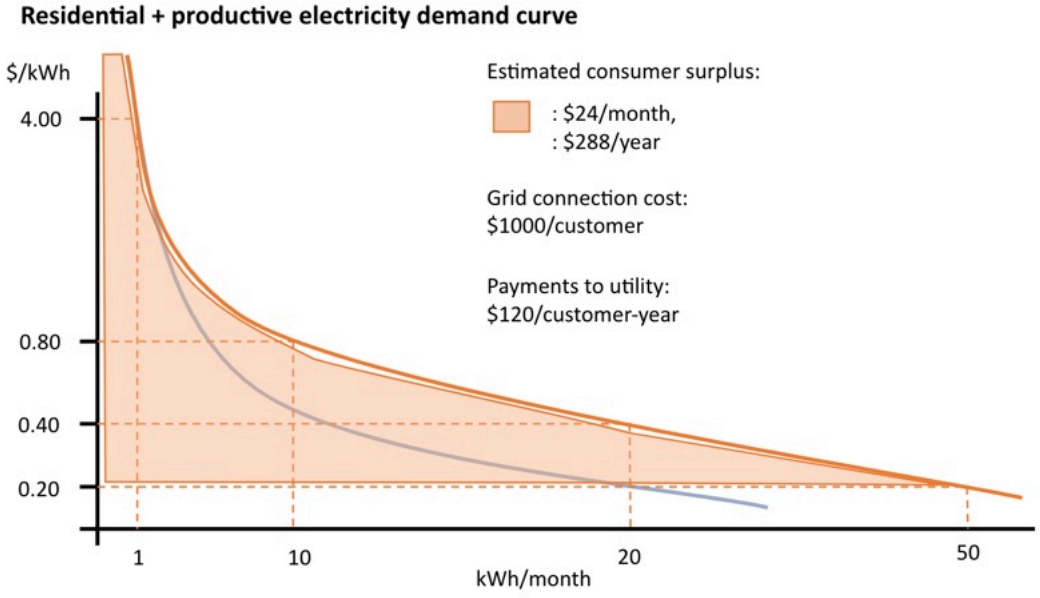

Figure VII-3: Simulated residential and productive electricity demand curve. Estimated consumer surplus, grid connection costs, and annual payments to the utility are presented alongside the demand curve.

2. The Effects of Productive Electricity Demand on Energy System Design

As previous chapters have established methods of detecting irrigation in the Ethiopian Highlands, predicted irrigated areas can now be inputted to energy system planning models to understand the effect of productive electricity demand. To this end, the Irrigated System Electrification (ISE) model is introduced, a linear program that solves for lowest-cost combinations of energy capacity and dispatch required to meet combined household and productive demand. The ISE model is parameterized based on SEL experience with sub-Saharan energy system costs and constraints, and applied to a case study near Gebedge, Ethiopia, where a specific orientation of irrigated area is predicted. The following subsections detail the irrigated area prediction process, ISE assumptions and governing constraints, and model results for a series of configurations that range from isolated household to grid-connected systems.

2.1 Predicting irrigation presence near Gebedge, Ethiopia and defining grid network topology

The irrigation detection model introduced in Chapter V is deployed over a 5.12km-by-2.56km area near Gebedge, Ethiopia, located at 12.87°N, 37.74°E. Irrigation predictions are made at 10m using Sentinel-2 (S2) imagery, and the resulting predictions are polygonized with all polygons smaller than 3 Ha being removed. Figure VII-4(a) shows the polygonized predictions in red for the area of interest on top of high-resolution Airbus imagery collected during the 2019 dry season. For the most part, these polygons outline vegetated areas among a largely barren/non-vegetated background. In total, 257 Ha of the 1310 ha are predicted as irrigated.

Next, predicted irrigated areas are split into 300m radius irrigation zones. Here, 300m represents the longest distance that can be connected with low-cost low voltage (LV) wire. The exact locations of the 300m irrigation zones are determined to maximize the overlap between

zones and predicted irrigated areas; 19 zones are selected, as any additional zone will not cover more than 3 ha in supplementary predicted irrigated area. Figure VII-4(b) shows the orientation of irrigation zones for the case study, with zonal centroids assumed to be the locations of on farm electricity generation/storage systems as well as the sites of interzonal grid connection.

Figure VII-4(c) displays a grid network topology that connects the predicted irrigation zones through a combination of medium voltage (MV) and LV wire, shown in orange and blue, respectively. The grid network is determined by a two-level network design algorithm, a SEL-developed tool that has been deployed to estimate optimal electrification strategies in Kenya [329]. From the modeled grid network, total MV and LV wire lengths are calculated, which contribute to overall system cost in specific simulated energy system configurations.

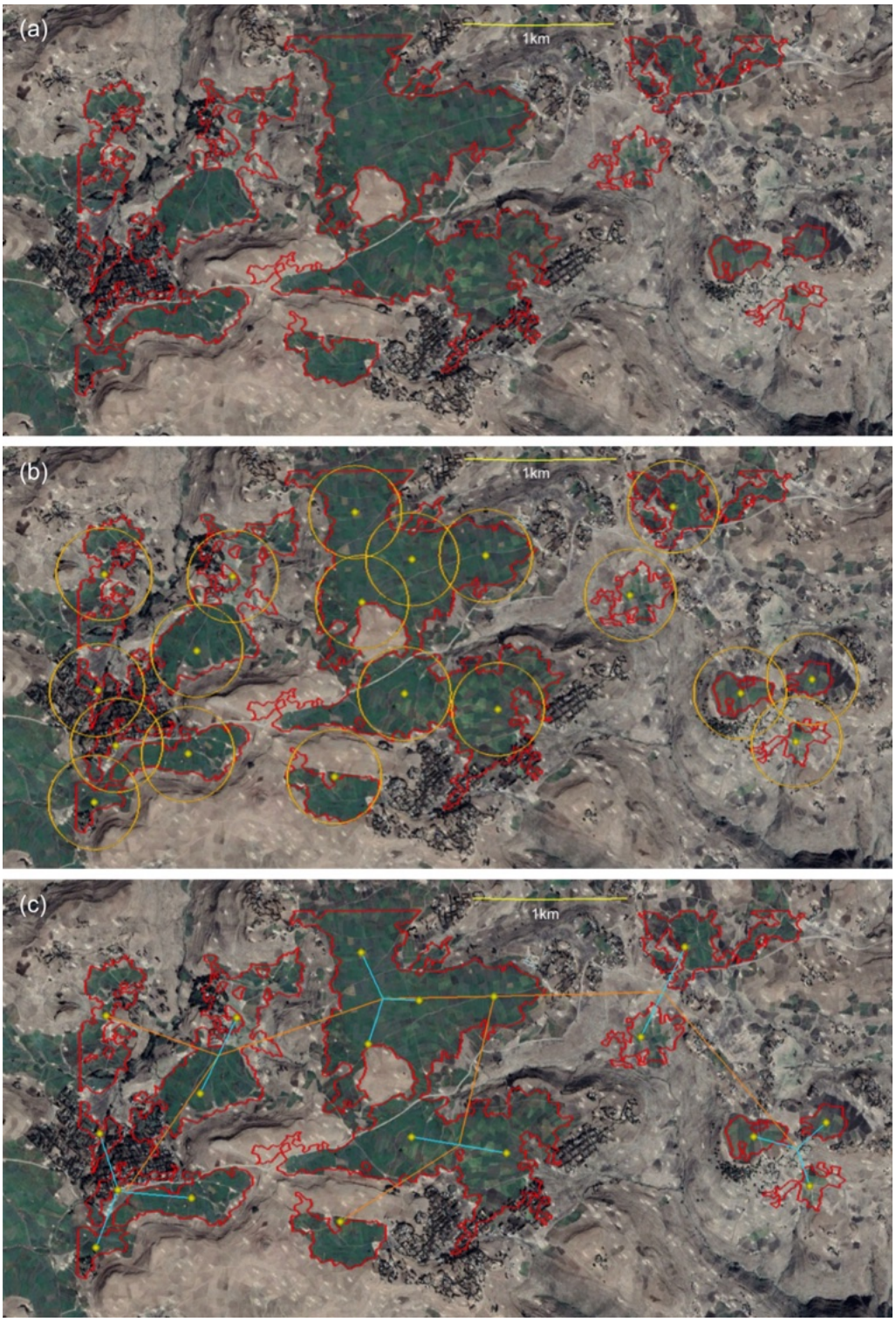

Figure VII-4: (a) Polygonized irrigation predictions (red) near Gedebge, Ethiopia (12.87°N, 37.74°E). Predictions span a 2.56km x 5.12km rectangle and are presented on top of Airbus imagery from December 21, 2018, hosted in Google Earth Pro. (b) Irrigation zones (300m radius orange circles) for polygonized irrigation predictions. (c) Proposed electric grid network with MV (orange) and LV (blue) lines serving irrigation zone centroids (yellow).

2.2 Energy system configurations and modeling assumptions

To understand the cost-effectiveness of various methods of meeting combined domestic and productive demand, four separate energy system configurations are tested.

1. Each farm, assumed to be 0.2 Ha in size, has its own generation system (solar, diesel, and battery), with no LV wire requirements.
2. Each irrigation zone has its own generation system (solar, diesel, and battery) shared by all farms in the zone. Additional wiring is required within the zone to connect all farms to the central generation system.
3. Configuration (2), plus zones are connected by the MV and LV network shown in Figure VII-4.
4. Configuration (3), with grid electricity available to the network.

The network topology introduced in Figure VII-4 contains 19 unique irrigation zones that cover 257 ha. Of this total area, 10% is considered irrigated. Moreover, for zone-level systems (Configurations (2), (3), and (4)), 150m of LV wire per irrigated hectare is assumed for connecting boreholes located at each farm to centralized solar generation. Given 10% irrigation, this corresponds to 3855m of LV wire. Connecting all irrigation zones via a grid network, as is the case in Configurations (3) and (4), requires an additional 7904m of MV wire and 4691m of LV wire.

In parameterizing ISE, assumptions are drawn from previous SEL experience. These values are found to be roughly in line with recent development reports and values specified by lab collaborators; however, a full sourcing of system assumptions is left for future work.

ISE is formulated as a mixed integer linear program, where optimal power flow and solar, diesel, and battery capacities are determined over a rain-free 5-day period, T. The model has an hourly time resolution, is programmed in Python, and is solved with Gurobi. Similar to the energy

system models introduced in Part I, ISE is spatially heterogenous, with the definition of spatial nodes depending on the configuration: For configuration 1, nodes are defined as individual farms plots and are not connected. For configurations (2)-(4), nodes are defined as the 300m radius irrigation zones pictured in Figure VII-4; in configurations (3) and (4), these zones are connected via an LV and MV grid network.

2.2.1 Objective function

The objective function of ISE is defined as follows:

$$obj = minimize(\sum_{i \in I} \Big[C_{solar} * X_{solar,i} + C_{batt} * X_{batt,i} + C_{diesel} * X_{diesel,i} + \sum_{t=0}^{T} c_{diesel} * G^{t}_{diesel,i} + c_{grid} * G^{t}_{grid,i} \Big] \quad \text{(VII-1)}$$

where C_Y is the annualized cost of technology *Y*; $X_{Y,i}$ is the capacity of technology *Y* installed at node *i;* c_{diesel} is the cost of diesel fuel; $G^{t}_{diesel,i}$ is diesel-generated electricity; c_{grid} is the grid electricity tariff; and $G^{t}_{grid,i}$ is the amount of grid electricity consumed at time *t* in node *i.*

Annualized costs for technology *Y*, C_Y, are determined for interest rate *j*; technology-specific annualization period P_Y, and per-unit capacity cost CAP_Y. Eq. (VII-2) presents this formulation, where the $\frac{5}{365}$ coefficient represents the adjustment of annual costs to those for a 5-day simulation period:

$$C_Y = \frac{5}{365} * \frac{j * (1+j)^{P_Y}}{((1+j)^{P_Y} - 1)} * CAP_Y \quad \text{(VII-2)}$$

We note that adjustment of the annualized capacity costs for the 5-day simulation period will result in the lowest possible LCOEs for generated electricity, as these costs are being spread over a period of time in which all capacities will be utilized. If instead the ISE optimization model is solved and costed over the entire year, LCOE will increase, as there will be times during the simulation period when the solar/diesel/battery system is not fully utilized, as will be the case in the rainy season.

2.2.2 Energy balance

Each node in the system contains hourly timeseries of fixed domestic electricity load, $D^t_{fixed,i}$, and a productive electricity load, $D^t_{irrig,i}$, both proportional to the amount of irrigated land present in the node. Each node also contains an identical solar potential timeseries, W^t_{solar}. The fixed domestic load and solar potential timeseries are presented in Figure VII-5; these timeseries repeat for every day in the 5-day simulation period. In contrast, nodal hourly productive electricity demand is a flexible quantity, constrained to sum to the full nodal productive electricity demand, $D^{full}_{irrig,i}$, over the entire simulation period per Eq. (VII-3). This designed flexibility allows the model to meet the productive electricity demand during times when energy is abundant, as is the case during afternoon hours when fixed demand is low but solar generation potential is high.

$$\sum_{t=0}^{T} D^t_{irrig,i} = D^{full}_{irrig,i} \tag{VII-3}$$

Eq. (VII-4) presents the full ISE energy balance equation:

$$X_{us-solar,i} * W^t_{solar} + \delta^t_{batt,i} - \gamma^t_{batt,i} + G^t_{diesel,i} + G^t_{grid,i} \geq D^t_{fixed,i} + D^t_{irrig,i} \quad \text{(VII-4)}$$

where $\delta^t_{batt,i}$ is battery discharge, and $\gamma^t_{batt,i}$ is battery charge. Excess generation in Eq. VII-4 is attributed to solar generation and is calculated by the slack in the constraint. $G^t_{grid,i}$ is only available for Configuration (4) simulations.

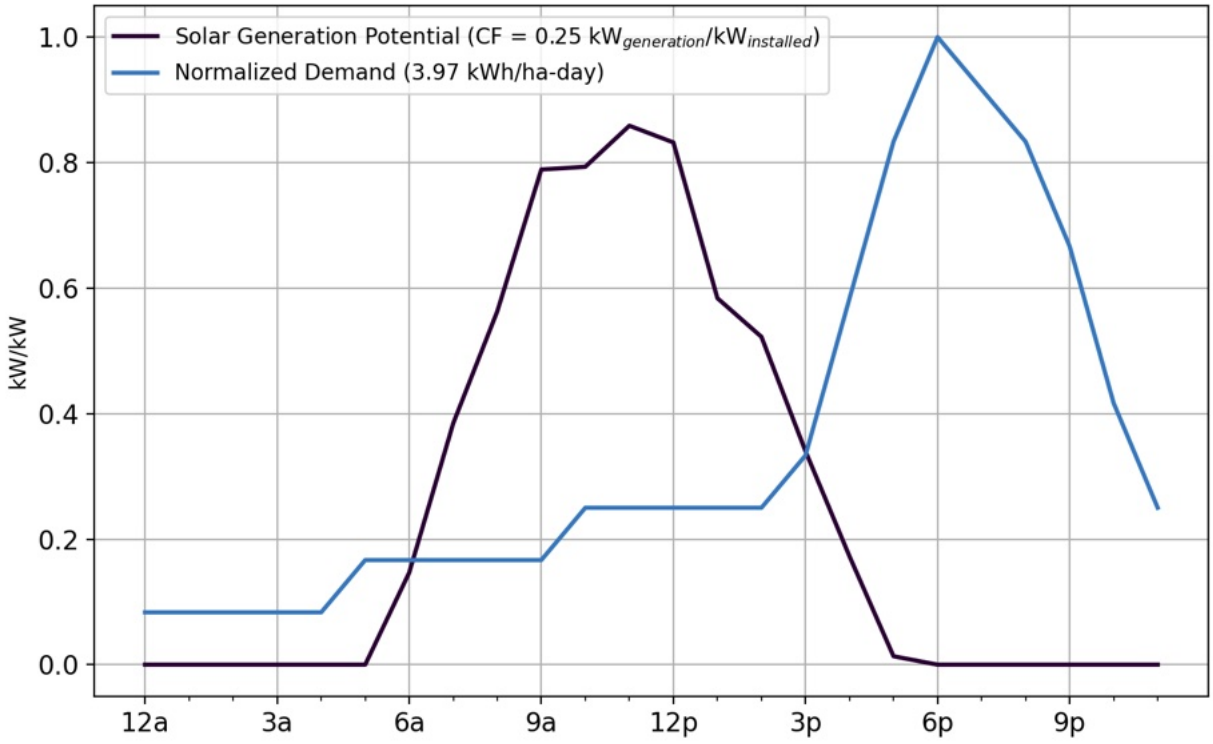

Figure VII-5: Daily timeseries of solar generation potential and normalized fixed household demand for the ISE optimization model. Daily timeseries are repeated every day within the 5-day simulation period.

2.2.3 Solar

Solar power is a decision variable in the optimization model. It is assumed to have an annualization rate of 0.171 – corresponding to 15% interest paid over 15 years – with decreasing, piecewise per-unit capacity costs, as is shown in Table VII-1:

Table VII-1: Piecewise solar capacity costs.

kW	0-1	1-3	3-10	10-100	>100
$	3000	2500	2000	1500	1000

The solar generation potential timeseries is extracted for North Gorda, Ethiopia, from the European Commission Photovoltaic Geographic Information System [330]. The timeseries is scaled to have a capacity factor of 25%, and is repeated for each of the 5 days in the simulation period *T*.

2.2.4 Battery storage

The optimization model also includes storage capacity as a decision variable, $E^t_{batt,i}$. Here, storage characteristics are loosely based on available lithium-ion battery capabilities. The following assumptions are specified: An annualization rate of 0.298 – corresponding to 15% interest paid over 5 years; a $250/kWh per-unit cost; a power-to-energy ratio of 1:4; and a one-way efficiency, η_{batt}, of 95%. Battery storage constraints are presented in Eqs. (VII-5) – (VII-8).

$$\frac{\delta^t_{batt,i}}{\eta_{batt}} - \eta_{batt} * \gamma^t_{batt,i} = E^T_{batt,i} - E^t_{batt,i} \ , \qquad \forall t = 0 \tag{VII-5a}$$

$$\frac{\delta^t_{batt,i}}{\eta_{batt}} - \eta_{batt} * \gamma^t_{batt,i} = E^{t-1}_{batt,i} - E^t_{batt,i} \ , \qquad \forall t > 0 \tag{VII-5b}$$

$$E_{batt,i}^{t} \leq X_{batt,i} \tag{VII-6}$$

$$\gamma_{batt,i}^{t} \leq \frac{X_{batt,i}}{4} \tag{VII-7}$$

$$\delta_{batt,i}^{t} \leq \frac{X_{batt,i}}{4} \tag{VII-8}$$

2.2.5 Diesel generation

In the absence of grid power or battery dispatch, diesel generation is used to meet electricity demand for the simulated energy systems. Here, the following assumptions are stipulated: An annualization rate of 0.199 – corresponding to 15% interest paid over 10 years; a per-unit cost of $400/kW; a fuel cost, c_{fuel} of $1.25/L; a fuel energy content of 10.6 kWh/L; and an efficiency, η_{diesel}, of 30%. Eq. (VII-9) presents the relationship between c_{fuel} and the per-kWh cost of diesel-generation electricity, c_{diesel}:

$$c_{diesel} = \frac{1}{10.6 * \eta_{diesel}} * c_{fuel} \tag{VII-9}$$

2.2.6 Grid costs

For model configurations in which either low voltage (LV) or medium voltage (MV) wiring is required, the following assumptions are used: An annualization rate of 0.16 – corresponding to 15% interest paid over 20 years; a per-unit MV cost of $10/m; a per-unit LV cost of $5/m, and a grid electricity tariff of $0.06/kWh.

2.2.7 Irrigation and pumping system

To characterize the irrigation and pumping systems that deliver water to the irrigated areas, a 5-day water requirement of 35mm/m^2 is adopted. Each farm is assumed to have a size of 0.2 ha. On-site pumps are specified to have a head of 50m, an efficiency of 60%, a minimum pump operation time of 2 hours, and a minimum pump size of 1 kW. Altogether, these assumptions yield an electricity load for irrigation equivalent to 15.89 kWh/ha/day; over the 5-day simulation period, the full flexible productive load, $D_{irrig,i}^{full}$, equals 79.45 kWh/ha.

2.2.8 Fixed household load

Fixed household load, $D_{fixed,i}^{t}$, is determined based on the average of the daily energy patterns for the 45 lowest usage circuits in [331]. Figure VII-6 shows the household demand timeseries for these lowest 45 consumers, with the average in bold. $D_{fixed,i}^{t}$ is repeated for each of the 5 days in the simulation period *T*.

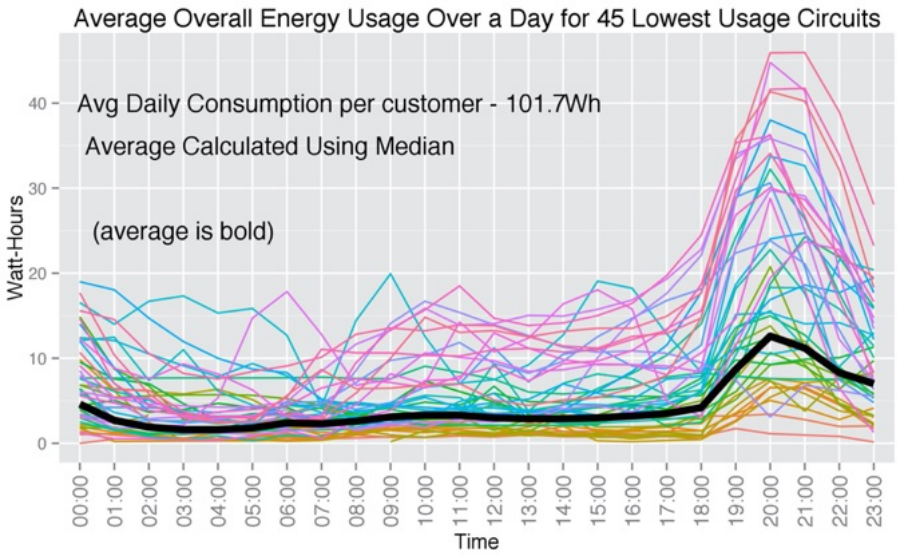

Figure VII-6: Daily energy consumption timeseries for the 45 lowest consumers in [331]. The average timeseries in bold constitutes the fixed household load timeseries in ISE simulations.

2.3 Energy system optimization results

Table VII-2 contains results for the four configurations considered in the case study. The table details the configuration description, the average electricity demand, the amount of medium and low voltage wire present, and in the final four columns, results returned by the least-cost optimization model. The last column presents the levelized cost of electricity, or LCOE, for each configuration, which can be understood as the all-in cost of electricity generation and distribution divided by the total electricity produced.

Table VII-2 shows that Configuration (1) – with farm level generation systems – offers the most expensive method of meeting the combined electricity demand with an LCOE of $0.397/kWh. Transitioning to an irrigation-zone based generation system, as is achieved in Configuration (2), brings cost savings of nearly 30% ($0.272/kWh), as 177 kWh less battery capacity is needed to meet the electricity demand, along with a 12 kW reduction in the required solar capacity. Configuration (3) does not result in any cost reductions; in fact, the cost of interconnecting the independent irrigation zones increases LCOE to $0.289/kWh, as the interconnections do not bring enough load balancing benefits to offset the increased wiring costs. Lastly, Configuration (4) demonstrates the value of irrigation zone interconnection if grid power is available: Connecting all zones with grid power results in the lowest LCOE of $0.164/kWh, as no solar, diesel, or battery capacity is required.

While presented results apply only to a single batch of tests, this case study provides a framework for future energy system modelling of productive and non-productive demands. Many additional questions remain for future work – e.g., determining the value of productive

demand flexibility – however these efforts are left the next generation of SEL energy system researchers.

Table VII-2: Energy system optimization results for the 4 different configurations modeled.

		Additional configuration details			Optimal model-returned capacities[1]			
Configuration	Description	Total Avg. Demand [kW]	MV line [m]	LV line [m]	Solar [kW]	Diesel [kW]	Battery [kWh]	LCOE [$/kWh][2,3]
1	Each farm has its own generation system (solar, diesel, and battery)	21.27	0	0	89	8	208	0.397
2	Each irrigation zone has its own generation system shared by all farms in the zone. Addl. LV wiring is required within the zone.	21.27	0	3855	77	9	31	0.272
3	Configuration (2), plus irrigation zones are connected with an MV and LV network.	21.27	7904	8546	75	9	191	0.289
4	Configuration (3), plus grid electricity is available to the network.	21.27	7904	8546	0	0	0	0.164

[1] Returned quantities are rounded to their nearest unit.
[2] Levelized cost of electricity (LCOE): the all-in costs of electricity generation and distribution divided by the total electricity produced.
[3] The presented LCOEs represent the cost of electricity during the 5-day irrigation period only. If costs are considered on an annual basis – i.e., over a period of time when the irrigation system is not in operation – the LCOE will increase, especially in capital-intensive scenarios 1, 2, and 3.

Conclusions

This dissertation presents two series of analyses focused on energy system planning, the first set in New York State and second set in Ethiopia. In New York State, a collection of linear programs is produced to evaluate methods of energy-sector decarbonization. In Ethiopia, approaches are introduced to first detect smallholder irrigation at a spatial resolution of approximately 30m by 30m, and then to tie these irrigated areas and their attendant energy demands back to system planning considerations.

The culmination of developed energy system models for New York State, the System Electrification and Capacity Transition (SECTR) model, is described in Chapter I. SECTR is spatially explicit, and includes decision variables for onshore wind, offshore wind, solar, battery storage, and transmission capacities. Once parameterized with realistic costs for these decision variables, SECTR enables the investigation of decarbonization pathways with varying percentages of low-carbon electricity (LCPs) and heating and vehicle electrification (HVE) rates.

Applying SECTR to the NYS setting (SECTR-NY) reveals a series of findings regarding optimal transitions to a low-carbon economy. In a system with 18.7 GW of electricity demand, deep decarbonization targets are only feasible with both significant LCPs and large increases in HVE. However, on the path to these deep targets, results indicate that intermediary climate goals (corresponding to those defined for 2030 to 2040) can be met more affordably by prioritizing electrification with 40-70% low-carbon electricity supply instead of aiming for full grid decarbonization. Furthermore, as decarbonization progresses, SECTR-NY results identify a shift in optimal energy infrastructure expansion strategies. First, low-cost onshore generation is built, followed by higher cost offshore wind generation that feeds directly into downstate load pockets.

Beyond 80% low-carbon electricity, paired solar and batteries are shown to be the most cost-effective method of displacing remaining fossil fuel-based generation, due to the natural cycling of solar that facilitates regular battery charging and discharging.

Altogether, SECTR-NY pathways reveal three primary cost drivers for a decarbonizing regional energy system: (1) decreasing per-unit costs of existing infrastructure with increasing electrified demand, (2) increasing in-state generation costs from low-carbon sources relative to gas-based and hydropower generation, and (3) increasing integration costs at high percentages of low-carbon electricity. Comparison of these pathway results to analyses presented to the New York State Climate Action council lends further confidence to our findings: After accounting for differences in upstream natural gas leakage assumptions and among timeseries used to represent wind/solar generation and electricity demand, the two sets of proposed decarbonization strategies largely agree, with similar energy resource capacities recommended to meet the state's 2030 climate goals.

Chapters II and III present more contained studies of decarbonization in New York State, with more limited simulation of electrified vehicle loads and no inclusion of heating electrification. The Renewable Target Model (RTM) introduced in Chapter II also parameterizes the NYS grid with realistic costs and existing capacities, although it only assesses the least cost methods of achieving renewable generation targets between 50% and 80%. Here, achieving 50% renewable energy supply requires only a buildout of new generation capacity – approximately 20 GW in total – at a levelized cost of electricity in line with current supply values ($52/MWh).

RTM results yield a series of other insights regarding the NYS energy transition. First, expanded transmission will play a minimal role in meeting renewable generation targets between

50-60%, and the costs of expanding transmission to reach higher percentages of renewable electricity supply will be dwarfed by the costs of new renewable generation and battery storage at these stages in the decarbonization pathway. Second, similar to findings from SECTR-NY, the need for battery storage increases substantially at higher proportions of renewable electricity supply: No battery storage is required to meet a 50% target, while storage equal to 2 hours of average demand (37.4 GWh) is installed at 65%, a quantity that increases to 16 hours of storage at 80%.

Exploration of different types of energy system flexibility in New York (Chapter III) illustrates the comparative advantages of additional hydropower reservoirs (supply-side flexibility), pumped hydropower storage (bidirectional flexibility), and electric vehicle (EV) demand (demand-side flexibility). Flexibility is measured in terms of additional wind generation integration, with either 10 GW or 30 GW of capacity simulated. After devising a method to compare similar amounts of these flexibility measures, a linear program developed to represent a simplified version of the NYS grid determines that additional EV demand allows for the greatest increase in wind utilization (16% increase over a baseline of 5.3 GW) when 30 GW of wind capacity is installed, followed by bidirectional storage (12% increase over the baseline), and then supply-side flexibility (9% increase over the baseline).

Energy system planning considerations in Ethiopia vary significantly from those in New York, as the second half of this dissertation makes clear. In Ethiopia, additional electricity loads facilitate the expansion of higher-quality energy provision services, namely expansion of the grid. Electricity demands for irrigation – a type of productive energy demand – are of particular importance due to 1) their scale, and 2) flexibility in when they can be met across a period of

days. Accordingly, methodologies that locate irrigation in Ethiopia can provide substantial value to energy system planners, as these are areas that either already contain productive electricity demands or can sustain additional demand through the expansion of existing irrigation.

Chapter IV summarizes relevant literature on land cover mapping and irrigation detection, in particular diving deep into a 2012 paper by Professor Christopher Small that introduces a method of characterizing dominant vegetation phenologies within a region of interest. This method is then applied to the Amhara state in the Ethiopian Highlands to visualize areas that contain vegetation cycles in-phase and out-of-phase with rainfall. A separate approach then calculates the misalignment between timeseries of Moderate Resolution Imaging Spectroradiometer (MODIS) imagery and Climate Hazards Group InfraRed Precipitation with Station Data (CHIRPS) estimates to offer insight into where vegetation growth cannot be attributed to rainfall and must be enabled through some other mechanism, potentially irrigation.

Like many other parts of sub-Saharan Africa, Ethiopia lacks public data on the locations of irrigation within the country. Because any irrigation prediction model will only perform as well as the quality of data used to train it, methods that collect high-accuracy, verifiable, spatially diverse irrigation labels provide immense value to researchers. Chapter V presents such an approach, whereby limited ground-collected labels from a 2021 survey effort are supplemented by visually collected labels derived from enhanced vegetation index (EVI) timeseries inspection on the Descartes Labs platform, a commercial geospatial analytics tool. Given the immense agrological and climatological diversity within Ethiopia, the previously introduced method of spatiotemporal vegetation characterization is applied to guide supplementary label collection and indicate the extent of classifier applicability.

To develop a robust irrigation classifier, many neural network architectures and training strategies are tested, all for input Sentinel-2 imagery with a spatial resolution of 10m. These experiments reveal the insight that EVI timeseries allow for greater classification accuracy compared to a full suite of spectral bands; and that randomly shifting inputs by up to ±30 days improves classifier performance. The highest performing irrigation detection model achieves a >95% accuracy across samples withheld for testing, and across a series of samples independently collected via the introduced method of label supplementation. This model is then deployed over the entirety of Tigray and Amhara, finding that irrigated area in these two states has decreased by 40% from 2020 to 2021, likely due to ongoing civil conflict in the region.

Chapter VI extends the introduced irrigation detection methodology across a number of separate applications. First, the size distribution of predicted irrigated areas in Chapter V is assessed, followed by an investigation into regions that contain model misclassification of irrigation presence. These results show that erroneous irrigation predictions over non-cropped areas near the Amhara/Afar border can be attributed to anomalous rains during the 2020 dry season; an auxiliary 10m land cover map provided by the European Space Agency (ESA) is then used to filter out irrigation predictions that do not fall on cropped areas. Furthermore, a longitudinal study of irrigation in Tigray reveals that the plot-level location of irrigated areas changes substantially from year to year, and that the 40% decline in irrigated area from 2020 to 2021 is not due to excessive positive predictions in 2020, as irrigation increases marginally from 2019 to 2020 in the state. Lastly, performance evaluation over additional collected labels in 2020 in Tigray show that the presented irrigation detection model remains robust even when applied across a year and region not included in the training process.

Further theoretical grounding for the importance of considering productive electricity demands – especially those for irrigation – is provided in Chapter VII. In tying irrigation predictions back to the original goal of informed energy system planning, Chapter VII also presents the Irrigated System Electrification (ISE) model, developed to determine least-cost methods of satisfying combined residential and productive electricity demands. Using irrigation predictions from an area near Gebedge, Ethiopia, and model parameters gleaned from previous Sustainability Engineering Lab (SEL) experience, simulations indicate that grid power offers the cheapest method of meeting residential and productive demand at a cost of $0.16/kWh; if a grid network is not present, a minigrid produces an electricity cost of $0.27/kWh, while a system configuration with plot level generators results in a cost of $0.40/kWh.

The pages herein contain a variety of studies across many different locations and research directions. While these efforts comprise a great deal of work – six years to be exact – there are many natural extensions and related questions that have not been addressed. These applications are discussed briefly below and are left to the next cohort of SEL PhD students. I wish them the best of luck.

Understanding distribution system needs in New York State during the decarbonization transition. Simulated cost-optimal decarbonization strategies in New York include substantial amounts of heating and vehicle electrification. These electrified loads will have significant impact on distribution system needs, and the System Electrification and Capacity Transition model for New York (SECTR-NY) should be expanded to account for these considerations.

Higher spatial resolution for energy system modeling efforts in New York. Currently, SECTR-NY splits the NYS electricity grid into four nodes that align with major transmission lines

within the state. To achieve a finer-grained understanding of transmission bottlenecks and upgraded network needs, future work should improve the spatial resolution of SECTR-NY by increasing the number of modeled nodes, particularly in densely populated downstate node pockets.

Further evaluation of the methods and value of managing increased demand peaks due to electrification, primarily those from heating. New electrified demands will induce significant increases in peak electricity demand in New York State, mostly due to electrified heating requirements during cold weather events. Managing these peaks will be crucially important to controlling distribution system upgrade requirements; future work should explore the various methods of peak shaving, their costs, and the feasibility of their implementation at scale.

Understanding the transferability of the irrigation detection methodology developed for Ethiopia. The irrigation detection methodology introduced in Chapter V demonstrated excellent performance across the Ethiopian Highlands, a highly cropped region dominated by a single main rainy season with clear vegetation phenologies. It remains to be seen whether this approach is applicable in other parts of sub-Saharan Africa, particularly those like Uganda with persistent rains and cloud cover that do not contain easily differentiable phenologies.

Determining the plot-level boundaries of irrigated areas. Irrigation predictions for the Ethiopian Highlands are generated pixelwise. Although this approach is methodologically simpler than convolutional or object-based alternatives, it does not account for spatial context and accordingly cannot generate field boundaries for individual plots. Additional work should look into filling this gap, potentially via sub-meter resolution imagery.

Developing an irrigation detection model with leaner imagery requirements. The trained classifier produced for irrigation detection in Ethiopia intakes 36-timestep timeseries, each representing the annual vegetation cycle for a single 10m by 10m pixel. These timeseries are likely oversampled in time: yearly vegetation phenologies can be represented with fewer timesteps. Moreover, methods that do not require input imagery to be collected at the exact same timesteps, instead encoding temporal information along with the input timeseries, will save researchers substantial time spent preprocessing data.

Assessing the value of flexible electricity demands for irrigation. Model simulations in Chapter VII produce electricity cost estimates for an energy system with combined residential and productive electricity demands. In these scenarios, productive electricity demand can be flexibly met across the 5-day simulation period. Determining the value of this flexibility, compared to baselines where the productive electricity load is fixed in time, is another useful direction of research.

Determining the likely extent of future irrigated areas. The predicted irrigated areas in Chapter V represent the best-known extent of current irrigation within the Ethiopian Highlands. While these predictions yield insights into where productive energy demands presently exist, efforts to estimate the distribution of future irrigated areas within the country will have substantial value for energy system planners.

Incorporating electricity demands for irrigation in regional or national planning strategies. Lastly, predicted irrigated areas and their attendant energy demands are fed into the Irrigated System Electrification (ISE) model in Chapter VII. The extent of the system area considered is 2.5km-by-5 km, and model results produce insights relevant to the portion of Ethiopia in question

and to other similarly oriented groups of irrigated plots. However, productive demands for irrigation should be considered for the entire country, and then fed into regional and national planning strategies that can assess infrastructure expansion decisions in a wholistic manner.

References

[1] New York State Senate. Climate Leadership and Community Protection Act — Final Bill Text 2019. https://www.nysenate.gov/legislation/bills/2019/s6599.

[2] Clack CTM, Choukulkar A, Cote B, McKee SA. Technical Report: Economic & Clean Energy Benefits of Establishing a Competitive Wholesale Electricity Market in the Southeast United States 2020:153. https://vibrantcleanenergy.com/wp-content/uploads/2020/08/SERTO_WISdomP_VCE-EI.pdf.

[3] Clean Air Task Force. State and utility climate change targets shift to carbon reductions 2019:1–9. https://www.catf.us/wp-content/uploads/2019/05/State-and-Utility-Climate-Change-Targets.pdf.

[4] Scovronick N, Budolfson M, Dennig F, Errickson F, Fleurbaey M, Peng W, et al. The impact of human health co-benefits on evaluations of global climate policy. Nat Commun 2019;10:1–12. https://doi.org/10.1038/s41467-019-09499-x.

[5] Fitzroy F. A Green New Deal: The Economic Benefits of Energy Transition. Substantia 2019;3(2) Suppl:55–67. https://doi.org/10.13128/Substantia-276.

[6] Tyson A, Kennedy B. Two-Thirds of Americans Think Government Should Do More on Climate. Pew Res Cent 2020. https://www.pewresearch.org/science/2020/06/23/two-thirds-of-americans-think-government-should-do-more-on-climate/.

[7] Guelpa E, Bischi A, Verda V, Chertkov M, Lund H. Towards future infrastructures for sustainable multi-energy systems: A review. Energy 2019;184:2–21. https://doi.org/10.1016/j.energy.2019.05.057.

[8] Grubler A, Wilson C, Bento N, Boza-Kiss B, Krey V, McCollum DL, et al. A low energy demand scenario for meeting the 1.5 °C target and sustainable development goals without negative emission technologies. Nat Energy 2018;3:515–27. https://doi.org/10.1038/s41560-018-0172-6.

[9] Bistline JET. Roadmaps to net-zero emissions systems: Emerging insights and modeling challenges. Joule 2021;5:2551–63. https://doi.org/10.1016/j.joule.2021.09.012.

[10] Conlon T, Waite M, Modi V. Assessing new transmission and energy storage in achieving increasing renewable generation targets in a regional grid. Appl Energy 2019;250:1085–98. https://doi.org/10.1016/j.apenergy.2019.05.066.

[11] Barth R, Brand H, Meibom P, Weber C. A stochastic unit-commitment model for the evaluation of the impacts of integration of large amounts of intermittent wind power. 2006 9th Int Conf Probabilistic Methods Appl to Power Syst PMAPS 2006. https://doi.org/10.1109/PMAPS.2006.360195.

[12] Kern JD, Patino-Echeverri D, Characklis GW. An integrated reservoir-power system model for evaluating the impacts of wind integration on hydropower resources. Renew Energy 2014. https://doi.org/10.1016/j.renene.2014.06.014.

[13] An Y, Zeng B. Exploring the modeling capacity of two-stage robust optimization: Variants of robust unit commitment model. IEEE Trans Power Syst 2015;30:109–22. https://doi.org/10.1109/TPWRS.2014.2320880.

[14] Jenkins J, Sepulveda N. Enhanced Decision Support for a Changing Electricity Landscape: the GenX Configurable Electricity Resource Capacity Expansion Model. MIT Energy Initiat Work Pap 2017:1–40. https://energy.mit.edu/wp-content/uploads/2017/10/Enhanced-

Decision-Support-for-a-Changing-Electricity-Landscape.pdf.
[15] Levi PJ, Kurland SD, Carbajales-Dale M, Weyant JP, Brandt AR, Benson SM. Macro-Energy Systems: Toward a New Discipline. Joule 2019;3:2282–6. https://doi.org/10.1016/j.joule.2019.07.017.
[16] Denholm P, Arent DJ, Baldwin SF, Bilello DE, Brinkman GL, Cochran JM, et al. The challenges of achieving a 100% renewable electricity system in the United States. Joule 2021;5:1331–52. https://doi.org/10.1016/j.joule.2021.03.028.
[17] Howells M, Rogner H, Strachan N, Heaps C, Huntington H, Kypreos S, et al. OSeMOSYS: The Open Source Energy Modeling System. An introduction to its ethos, structure and development. Energy Policy 2011;39:5850–70. https://doi.org/10.1016/j.enpol.2011.06.033.
[18] Gil E, Aravena I, Cardenas R. Generation Capacity Expansion Planning Under Hydro Uncertainty Using Stochastic Mixed Integer Programming and Scenario Reduction. IEEE Trans Power Syst 2015;30:1838–47. https://doi.org/10.1109/PESGM.2015.7285838.
[19] Short W, Sullivan P, Mai T, Mowers M, Uriarte C, Blair N, et al. Regional Energy Deployment System (ReEDS) - Technical Report. NREL 2011. https://www.nrel.gov/docs/fy12osti/46534.pdf.
[20] Karlsson K, Meibom P. Optimal investment paths for future renewable based energy systems-Using the optimisation model Balmorel. Int J Hydrogen Energy 2008;33:1777–87. https://doi.org/10.1016/j.ijhydene.2008.01.031.
[21] DeCarolis JF, Jaramillo P, Johnson JX, McCollum DL, Trutnevyte E, Daniels DC, et al. Leveraging Open-Source Tools for Collaborative Macro-energy System Modeling Efforts. Joule 2020;4:2523–6. https://doi.org/10.1016/j.joule.2020.11.002.
[22] Jenkins JD, Mayfield EN, Larson ED, Pacala SW, Greig C. Mission net-zero America: The nation-building path to a prosperous, net-zero emissions economy. Joule 2021;5:2755–61. https://doi.org/10.1016/j.joule.2021.10.016.
[23] Murphy C, Mai T, Sun Y, Jadun P, Muratori M, Nelson B, et al. Electrification Futures Study: Scenarios of Power System Evolution and Infrastructure Development for the United States. Natl Renew Energy Lab NREL/TP-6A20-72330 2021;1.
[24] Murphy C, Mai T, Sun Y, Jadun P, Donohoo-Vallett P, Muratori M, et al. High electrification futures: Impacts to the U.S. bulk power system. Electr J 2020;33:106878. https://doi.org/10.1016/j.tej.2020.106878.
[25] Steinberg D, Dave Bielen, Eichman J, Eurek K, Logan J, Mai T, et al. Electrification and Decarbonization: Exploring U.S. Energy Use and Greenhouse Gas Emissions in Scenarios with Widespread Electrification and Power Sector Decarbonization. Natl Renew Energy Lab 2017:43. https://doi.org/doi:10.2172/1372620.
[26] Zhou E, Mai T. Electrification Futures Study: Operational Analysis of U.S. Power Systems with Increased Electrification and Demand-Side Flexibility n.d. https://doi.org/10.2172/1785329.
[27] Mai T, Steinberg D, Logan J, Bielen D, Eurek K, McMillan C. An electrified future: Initial scenarios and future research for U.S. Energy and electricity systems. IEEE Power Energy Mag 2018;16:34–47. https://doi.org/10.1109/MPE.2018.2820445.
[28] Bellocchi S, Manno M, Noussan M, Prina MG, Vellini M. Electrification of transport and residential heating sectors in support of renewable penetration: Scenarios for the Italian

energy system. Energy 2020;196:117062. https://doi.org/10.1016/j.energy.2020.117062.
[29] Eichman JD, Mueller F, Tarroja B, Schell LS, Samuelsen S. Exploration of the integration of renewable resources into California's electric system using the Holistic Grid Resource Integration and Deployment (HiGRID) tool. Energy 2013;50:353–63. https://doi.org/10.1016/j.energy.2012.11.024.
[30] Tarroja B, Chiang F, AghaKouchak A, Samuelsen S, Raghavan S V., Wei M, et al. Translating climate change and heating system electrification impacts on building energy use to future greenhouse gas emissions and electric grid capacity requirements in California. Appl Energy 2018;225:522–34. https://doi.org/10.1016/j.apenergy.2018.05.003.
[31] Environmental Economics and Energy. Pathways to Deep Decarbonization in New York State 2020. https://climate.ny.gov/-/media/Project/Climate/Files/2020-06-24-NYS-Decarbonization-Pathways-CAC-Presentation.ashx.
[32] New York State Climate Action Council. Technical Advisory Group, Integration Analysis -- Initial Results Presentation (updated November 21, 2021) 2021. https://climate.ny.gov/-/media/Project/Climate/Files/2021-11-18-Integration-Analysis-Initial-Results-Presentation.ashx.
[33] Giarola S, Molar-Cruz A, Vaillancourt K, Bahn O, Sarmiento L, Hawkes A, et al. The role of energy storage in the uptake of renewable energy: A model comparison approach. Energy Policy 2021;151:112159. https://doi.org/10.1016/j.enpol.2021.112159.
[34] Waite M, Modi V. Electricity Load Implications of Space Heating Decarbonization Pathways. Joule 2020;4:376–94. https://doi.org/10.1016/j.joule.2019.11.011.
[35] Waite M. Analyses of Energy Infrastructure Serving a Dense Urban Area: Opportunities and Challenges for Wind Power, Building Systems, and Distributed Generation 2016:1–282. https://academiccommons.columbia.edu/doi/10.7916/D8FT8M6H.
[36] US Energy Information Administration (EIA). Cost and Performance Characteristics of New Generating Technologies 2020:1–4. https://www.eia.gov/outlooks/aeo/assumptions/pdf/table_8.2.pdf.
[37] Bloom A, Townsend A, Palchak D, Novacheck J, King J, Barrows C, et al. Eastern Renewable Generation Integration Study 2016:TP-6A20-64. https://www.nrel.gov/grid/ergis.html.
[38] New York Independent System Operator (NYISO). Market and Operational Data 2020. https://www.nyiso.com/energy-market-operational-data.
[39] NOAA Centers for Environmental Information. Integrated Surface Dataset 2001. https://www.ncei.noaa.gov/products/land-based-station/integrated-surface-database.
[40] US Energy Information Administration (EIA). Updated Buildings Sector Appliance and Equipment Costs and Efficiencies 2018. https://www.eia.gov/analysis/studies/buildings/equipcosts/.
[41] Northeast Energy Efficiency Partnerships (NEEP). Northeast Energy Efficiency Partnerships Cold Climate Air Source Heat Pump Product Listing 2019. https://ashp.neep.org.
[42] Shapiro C, Puttagunta S. Field Performance of Heat Pump Water Heaters in the Northeast Consortium for Advanced Residential Buildings 2016. https://www.nrel.gov/docs/fy16osti/64904.pdf.

[43] US Energy Information Administration (EIA). State Energy Consumption Estimates 2018. https://www.eia.gov/state/seds/sep_use/notes/use_print.pdf.
[44] US Bureau of Transportation Statistics. Average Age of Automobiles and Trucks in Operation in the United States 2020. https://www.bts.gov/content/average-age-automobiles-and-trucks-operation-united-states.
[45] US Environmental Protection Agency (EPA). The 2020 EPA Automotive Trends Report: Greenhouse Gas Emissions, Fuel Economy, and Technology since 1975 2020:151. https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P1010U68.txt%0A.
[46] New York Open Data. New York State 2016 Gasoline Sales by County 2016. https://data.ny.gov/Energy-Environment/Estimated-Gasoline-Sales-Beginning-1995/cwrk-j5nn.
[47] National Renewable Energy Laboratory (NREL). EVI-Pro Lite API 2021. https://developer.nrel.gov/docs/transportation/evi-pro-lite-v1/.
[48] New York Independent System Operator (NYISO). New York ISO Reliability Needs Assessment 2020. https://www.nyiso.com/documents/20142/2248793/2020-RNAReport-Nov2020.pdf/64053a7b-194e-17b0-20fb-f2489dec330d.
[49] Goldberg M, Keyser D. Transmission Line Jobs and Economic Development Impact (JEDI) Model User Reference Guide 2016. https://www.nrel.gov/analysis/jedi/transmission-line.html.
[50] MacDonald AE, Clack CTM, Alexander A, Dunbar A, Wilczak J, Xie Y. Future cost-competitive electricity systems and their impact on US CO2 emissions. Nat Clim Chang 2016. https://doi.org/10.1038/nclimate2921.
[51] New York Independent System Operator (NYISO). Western New York Public Policy Transmission Planning Report 2017.
[52] Transmission Developers Inc (TDI). Champlain Hudson Power Express: Project Development Portal 2019. http://ideas.mowerinteractive.com/clients/tdi/60394-website/site/index.php (accessed August 14, 2019).
[53] Starwood Energy Group. The Neptune Regional Transmission System 2007. http://starwoodenergygroup.com/wp-content/uploads/2014/06/6_NeptuneAnnouncement.pdf.
[54] US Energy Information Administration (EIA). Annual Energy Outlook 2020 with projections to 2050 2020. https://www.eia.gov/outlooks/aeo/pdf/aeo2020.pdf.
[55] U.S. Energy Information Administration (EIA). Electric Power Annual 2019 2020. https://www.eia.gov/electricity/annual/.
[56] New York Independent System Operator (NYISO). 2019 Load and Capacity Data: Gold Book 2019. https://www.nyiso.com/documents/20142/2226333/2019-Gold-Book-Final-Public.pdf/.
[57] New York Independent System Operator (NYISO). 2020 Load and Capacity Data: Gold Book 2020. https://www.nyiso.com/documents/20142/2226333/2020-Gold-Book-Final-Public.pdf/.
[58] US Energy Information Administration (EIA). Electric Power Annual Report 2019. Table 5.12. Accessed 10/21/2020. 2020. https://www.eia.gov/electricity/annual/.
[59] US Energy Information Administration (EIA). Electric Power Annual Report 2019. Table 3.11. Accessed 10/21/2020. 2020. https://www.eia.gov/electricity/annual/.

[60] New York Independent System Operator (NYISO). Locational Minimum Installed Capacity Requirements Study for the 2020-2021 Capability Year. 2020. https://www.nyiso.com/documents/20142/8583126/LCR2020-Report.pdf/4c9309b2-b13e-9b99-606a-7af426d93a47.
[61] New York State Energy Research and Development Authority (NYSERDA). New Efficiency: New York Analysis of Residential Heat Pump Potential and Economics 2019. https://www.nyserda.ny.gov/-/media/Files/Publications/PPSER/NYSERDA/18-44-HeatPump.ashx.
[62] Draxl C, Clifton A, Hodge BM, McCaa J. The Wind Integration National Dataset (WIND) Toolkit. Appl Energy 2015;151:355–66. https://doi.org/10.1016/j.apenergy.2015.03.121.
[63] Draxl C, Hodge B-M, Clifton A, McCaa J. Overview and Meteorological Validation of the Wind Integration National Dataset Toolkit. NREL 2015:87. https://www.nrel.gov/docs/fy15osti/61740.pdf.
[64] Waite M, Modi V. Modeling wind power curtailment with increased capacity in a regional electricity grid supplying a dense urban demand. Appl Energy 2016. https://doi.org/10.1016/j.apenergy.2016.08.078.
[65] Musial W, Heimiller D, Beiter P, Scott G, Draxl C. 2016 Offshore Wind Energy Resource Assessment for the United States 2016. https://doi.org/NREL/TP-5000-66599.
[66] Wiser R, Bolinger M, Barbose G, Millstein D. 2016 Wind Technologies Market Report 2017. https://www.energy.gov/eere/wind/downloads/2016-wind-technologies-market-report.
[67] Wiser R, Bolinger M. 2017 Wind Technologies Market Report 2018:1–98. https://www.energy.gov/eere/wind/downloads/2017-wind-technologies-market-report.
[68] Wiser R, Bolinger M. 2018 Wind Technologies Market Report 2019:1–98. https://www.energy.gov/eere/wind/downloads/2018-wind-technologies-market-report.
[69] Wiser R, Jenni K, Seel J, Baker E, Hand M, Lantz E, et al. Expert elicitation survey on future wind energy costs. Nat Energy 2016;1. https://doi.org/10.1038/nenergy.2016.135.
[70] Bloomberg New Energy Finance (BNEF). 2018 Wind O&M Price Index. 2018.
[71] Stehly T, Beiter P, Heimiller D, Scott G, Stehly T, Beiter P, et al. 2017 Cost of Wind Energy Review 2018. https://www.nrel.gov/docs/fy18osti/72167.pdf.
[72] Walter M, Philipp B, Spitsen P, Nunemake J, Gevorgian V. 2018 Offshore Wind Technologies Market Report 2019:1–94. https://www.osti.gov/biblio/1572771-offshore-wind-technologies-market-report.
[73] Bosch J, Staffell I, Hawkes AD. Global levelised cost of electricity from offshore wind. Energy 2019;189:116357. https://doi.org/10.1016/j.energy.2019.116357.
[74] Hummon M, Ibanez E, Brinkman G, Lew D. Sub-Hour Solar Data for Power System Modeling From Static Spatial Variability Analysis. 2nd Int Work Integr Sol Power Power Syst 2012. https://www.nrel.gov/docs/fy13osti/56204.pdf.
[75] Blair N, Dobos AP, Freeman J, Neises T, Wagner M, Ferguson T, et al. System Advisor Model, 2014.1.14: General Description. Natl Renew Energy Lab 2014. http://www.nrel.gov/docs/fy14osti/61019.pdf.
[76] Brooks B, Xenergy K, Whitaker C. Guideline for the use of the Performance Test Protocol for Evaluating Inverters Used in Grid-Connected Photovoltaic Systems. Sandia Natl Lab 2005. https://www.energy.ca.gov/sites/default/files/2020-06/2004-11-

22_Sandia_Test_Protocol_ada.pdf.
[77] Fu R, Feldman D, Margolis R. U.S. Solar Photovoltaic System Cost Benchmark: Q1 2018. NREL 2018:1–47. https://doi.org/10.7799/1325002.
[78] USDA National Agricultural Statistics Service. 2017 Census of Agriculture 2017. https://www.nass.usda.gov/Publications/AgCensus/2017/index.php.
[79] Ong S, Campbell C, Denholm P, Margolis R, Heath G. Land-Use Requirements for Solar Power Plants in the United States. 2013. https://doi.org/10.1016/j.rapm.2006.08.004.
[80] New York Independent System Operator (NYISO). Solar Impact on Grid Operations: An Initial Assessment 2016:1–57. https://www.transmissionhub.com/wp-content/uploads/2018/12/New-York-ISO-JUN-30-2016-Solar-Report.pdf.
[81] New York State Energy Research and Development Authority (NYSERDA). Solar Electric Programs Data (Accessed 06/01/2020) 2020. https://data.ny.gov/Energy-Environment/Solar-Electric-Programs-Reported-by-NYSERDA-Beginn/3x8r-34rs.
[82] Sengupta M, Habte A, Gotseff P, Weekley A, Lopez A, Molling C, et al. A Physics-Based GOES Satellite Product for Use in NREL's National Solar Radiation Database. Natl Renew Energy Lab 2014. https://www.nrel.gov/docs/fy14osti/62237.pdf.
[83] US Energy Information Administration (EIA). Energy Mapping System 2020. https://www.eia.gov/state/maps.php.
[84] Cole W, Frazier AW. Cost Projections for Utility-Scale Battery Storage: 2020 Update. Natl Renew Energy Lab 2020. https://www.nrel.gov/docs/fy20osti/75385.pdf.
[85] Tesla. Powerpack System Specifications 2019. https://www.tesla.com/powerpack.
[86] Smith K, Saxon A, Keyser M, Lundstrom B, Cao Z, Roc A. Life prediction model for grid-connected Li-ion battery energy storage system. Proc Am Control Conf 2017:4062–8. https://doi.org/10.23919/ACC.2017.7963578.
[87] Aggreko. Aggreko delivers grid stability to New York State with 2MW/3.8MWh energy storage system for National Grid. GlobeNewswire 2019. https://www.globenewswire.com/news-release/2019/06/03/1863350/0/en/Aggreko-delivers-grid-stability-to-New-York-State-with-2MW-3-8MWh-energy-storage-system-for-National-Grid.html.
[88] Lockheed Martin. 1MWh GridStar® Lithium Energy Storage Installation in Syracuse, New York 2019. https://www.lockheedmartin.com/content/dam/lockheed-martin/mfc/documents/energy/energy-syracuses-project-summary.pdf.
[89] Key Capture Energy. KCE NY 1 Breaks Ground on 20 MW 2018. https://www.keycaptureenergy.com/kce-ny-1-breaks-ground-on-20-mw/.
[90] Clean Coalition. Long Island Community Microgrid Project (LICMP) 2019. https://clean-coalition.org/community-microgrids/long-island-community-microgrid-project/.
[91] Energy Storage Association. US Energy Storge: 2019 Year in Review 2019. https://energystorage.org/wp/wp-content/uploads/2020/04/ESA_AR_2020_FINAL.pdf.
[92] Guerra OJ, Zhang J, Eichman J, Denholm P, Kurtz J, Hodge BM. The value of seasonal energy storage technologies for the integration of wind and solar power. Energy Environ Sci 2020;13:1909–22. https://doi.org/10.1039/d0ee00771d.
[93] Steward D, Saur G, Penev M, Ramsden T, Steward D, Saur G, et al. Lifecycle Cost Analysis of Hydrogen Versus Other Technologies for Electrical Energy Storage. Natl Renew Energy Lab 2009. https://www.nrel.gov/docs/fy10osti/46719.pdf.

[94] Houchins C, James B. 2020 DOE Hydrogen and Fuel Cells Program Review 2020. https://www.hydrogen.energy.gov/pdfs/review20/st100_houchins_2020_o.pdf.
[95] Eichman J, Townsend A, Melaina M. Economic Assessment of Hydrogen Technologies Participating in California Electricity Markets. Natl Renew Energy Lab 2016. https://www.nrel.gov/docs/fy16osti/65856.pdf.
[96] Penev M, Rustagi N, Hunter C, Eichman J. Energy Storage: Days of Service Sensitivity Analysis 2019. https://www.nrel.gov/docs/fy19osti/73520.pdf.
[97] Walker SB, Mukherjee U, Fowler M, Elkamel A. Benchmarking and selection of Power-to-Gas utilizing electrolytic hydrogen as an energy storage alternative. Int J Hydrogen Energy 2016;41:7717–31. https://doi.org/10.1016/j.ijhydene.2015.09.008.
[98] Johnson S. New York's Indian Point nuclear power plant closes after 59 years of operation. US Energy Inf Adm Today Energy 2021. https://www.eia.gov/todayinenergy/detail.php?id=47776#.
[99] New York State Energy Research and Development Authority (NYSERDA). Clean Energy Standard, 2020 Compliance Year: Load Serving Entity (LSE) Zero Emission Credit (ZEC) Rate. 2020. https://www.nyserda.ny.gov/All-Programs/Clean-Energy-Standard/LSE-Obligations/2020-Compliance-Year.
[100] US Energy Information Administration (EIA). Form EIA-923 detailed data with previous form data (EIA-906/920) 2020. https://www.eia.gov/electricity/data/eia923/.
[101] Mukerji R. NYISO Market Operations Report. New York Indep Syst Oper 2019. https://www.nyiso.com/documents/20142/8196990/Market+Operations+Report_+BIC_09.11.19.pdf.
[102] Whitney N. Amount of Capacity Qualified to Offer. New York Indep Syst Oper 2019. nyiso.com/documents/20142/3036383/4_Amt of Capacity Qualified to Offer.pdf/57f56a99-3293-d795-8584-21a70c495a5a.
[103] New York City Council. Local Laws of the City of New York for the Year 2019. No. 97. 2019. https://www1.nyc.gov/assets/buildings/local_laws/ll97of2019.pdf.
[104] NYSERDA. New York State Greenhouse Gas Inventory: 1990-2016 2019. https://www.nyserda.ny.gov/-/media/Files/EDPPP/Energy-Prices/Energy-Statistics/greenhouse-gas-inventory.pdf.
[105] Myrhe G, Shindell D, Huang J, Mendoza B, Daniel JS, Nielsen CJ, et al. Anthropogenic and natural radiative forcing. Clim Chang 2013 Phys Sci Basis Work Gr I Contrib to Fifth Assess Rep Intergov Panel Clim Chang 2013;9781107057:659–740. https://doi.org/10.1017/CBO9781107415324.018.
[106] Howarth RW. Methane Emissions and Greenhouse Gas Accounting: A Case Study of a New Approach Pioneered by the State of New York 2019:14. https://documents.dps.ny.gov/public/Common/ViewDoc.aspx?DocRefId=%7B3498AB82-B671-451E-A556-A917A61F939A%7D.
[107] Hayhoe K, Kheshgi HS, Jain AK, Wuebbles DJ. Substitution of natural gas for coal: Climatic effects of utility sector emissions. Clim Change 2002;54:107–39. https://doi.org/10.1023/A:1015737505552.
[108] Howarth RW, Santoro R, Ingraffea A. Methane and the greenhouse-gas footprint of natural gas from shale formations. Clim Change 2011;106:679–90. https://doi.org/10.1007/s10584-011-0061-5.

[109] Intergovernmental Panel on Climate Change (IPCC). Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories. Table 2, CH4 Emissions: Coal Mining and Handling. 1996. https://www.ipcc-nggip.iges.or.jp/public/gp/english/.
[110] National Energy Technology Laboratory (NETL). Petroleum-Based Fuel Life Cycle Greenhouse Gas Analysis - 2005 Baseline Model 2008. https://www.nata.aero/data/files/gia/environmental/bllcghg2005.pdf.
[111] Alvarez RA, Zavala-Araiza D, Lyon DR, Allen DT, Barkley ZR, Brandt AR, et al. Assessment of methane emissions from the U.S. oil and gas supply chain. Science (80-) 2018;361:186–8. https://doi.org/10.1126/science.aar7204.
[112] Schneising O, Burrows JP, Dickerson RR, Buchwitz M, Reuter M, Bovensmann H. Remote sensing of fugitive methane emissions from oil and gas production in North American tight geologic formations. Earth's Futur 2014;2:548–58. https://doi.org/10.1002/2014ef000265.
[113] Engineering ToolBox. Fuels - Higher and Lower Calorific Values 2003. https://www.engineeringtoolbox.com/fuels-higher-calorific-values-d_169.html.
[114] EPA. Emission Factors for Greenhouse Gas Inventories 2019:1–5. https://www.ecfr.gov/current/title-40/chapter-I/subchapter-C/part-98#ap40.23.98_19.1.
[115] US Energy Information Administration (EIA). Annual Energy Outlook 2020 with projections to 2050. 2020. https://www.eia.gov/outlooks/aeo/pdf/aeo2020.pdf.
[116] Patton DB, LeeVanSchaick P, Chen J, Naga RP. 2019 State of the Market Report for the New York ISO Markets. Potomac Ecnonomics 2019:1–24. https://www.nyiso.com/documents/20142/2223763/NYISO-2019-SOM-Report-Full-Report-5-19-2020-final.pdf/.
[117] US Energy Information Administration (EIA). Electricity Data Browser 2021. https://www.eia.gov/electricity/data/browser/.
[118] Waite M, Modi V. Existing and projected infrastructure capacities motivate alternatives to all-electric heating decarbonization. Joule-D-19-00676 2019.
[119] International Energy Agency. Getting Wind and Sun onto the Grid A Manual for Policy Makers 2017:64.
[120] International Energy Agency (IEA). Status of Power System Transformation 2017 n.d.
[121] Cole W, Frew B, Mai T, Bistline J, Blanford G, Young D, et al. Variable Renewable Energy in Long-Term Planning Models : A Multi-Model Perspective Variable Renewable Energy in Long-term Planning Models : A Multi-model Perspective 2017:35. https://doi.org/NREL/TP-6A20-70528.
[122] Cochran J, Bird L, Heeter J, Arent DJ. Integrating Variable Renewable Energy in Electric Power Markets: Best Practices from International Experience 2014.
[123] Bird L, Lew D, Milligan M, Carlini EM, Estanqueiro A, Flynn D, et al. Wind and solar energy curtailment: A review of international experience. Renew Sustain Energy Rev 2016;65:577–86. https://doi.org/10.1016/j.rser.2016.06.082.
[124] Günther S, Bensmann A, Hanke-Rauschenbach R. Theoretical dimensioning and sizing limits of hybrid energy storage systems. Appl Energy 2018;210:127–37. https://doi.org/10.1016/j.apenergy.2017.10.116.

[125] Auer H, Haas R. On integrating large shares of variable renewables into the electricity system. Energy 2016. https://doi.org/10.1016/j.energy.2016.05.067.
[126] Jorgensen J, Denholm P, Mai T. Analyzing storage for wind integration in a transmission-constrained power system. Appl Energy 2018;228:122–9. https://doi.org/10.1016/j.apenergy.2018.06.046.
[127] Jorgensen J, Mai T, Brinkman G. Reducing wind curtailment through transmission expansion in a Wind Vision future 2017:38.
[128] Strbac G. Demand side management: Benefits and challenges. Energy Policy 2008;36:4419–26. https://doi.org/10.1016/j.enpol.2008.09.030.
[129] Jacobson MZ, Howarth RW, Delucchi MA, Scobie SR, Barth JM, Dvorak MJ, et al. Examining the feasibility of converting New York State's all-purpose energy infrastructure to one using wind, water, and sunlight. Energy Policy 2013;57:585–601. https://doi.org/10.1016/j.enpol.2013.02.036.
[130] Shaner MR, Davis SJ, Lewis NS, Caldeira K. Geophysical constraints on the reliability of solar and wind power in the United States. Energy Environ Sci 2018. https://doi.org/10.1039/C7EE03029K.
[131] Enernex Corporation. Eastern Wind Integration and Transmission Study 2010. https://doi.org/CP-550-46505.
[132] Zhou Y, Scheller-Wolf AA, Secomandi N, Smith S. Managing Wind-Based Electricity Generation in the Presence of Storage and Transmission Capacity. Ssrn 2011. https://doi.org/10.2139/ssrn.1962414.
[133] Tuohy A, O'Malley M. Pumped storage in systems with very high wind penetration. Energy Policy 2011;39:1965–74. https://doi.org/10.1016/j.enpol.2011.01.026.
[134] Johnson JX, De Kleine R, Keoleian GA. Assessment of energy storage for transmission-constrained wind. Appl Energy 2014;124:377–88. https://doi.org/10.1016/j.apenergy.2014.03.006.
[135] Castillo A, Gayme DF. Grid-scale energy storage applications in renewable energy integration: A survey. Energy Convers Manag 2014;87:885–94. https://doi.org/10.1016/j.enconman.2014.07.063.
[136] Ghofrani M, Arabali A, Etezadi-Amoli M, Fadali MS. A framework for optimal placement of energy storage units within a power system with high wind penetration. IEEE Trans Sustain Energy 2013;4:434–42. https://doi.org/10.1109/TSTE.2012.2227343.
[137] Wogrin S, Gayme DF. Optimizing Storage Siting, Sizing, and Technology Portfolios in Transmission-Constrained Networks. IEEE Trans Power Syst 2015;30:3304–13. https://doi.org/10.1109/TPWRS.2014.2379931.
[138] National Conference of State Legislatures. State Renewable Portfolio Standards and Goals 2018.
[139] Waite M, Modi V. Impact of Deep Wind Power Penetration on Variability at Load Centers n.d.
[140] New York Independent System Operator (NYISO). Growing Wind: Final Report of the NYISO 2010 Wind Generation Study 2010.
[141] New York State Energy Research and Development Authority (NYSERDA). New York State Offshore Wind Master Plan 2016.
[142] Patton DB, LeeVanSchaick P, Chen J. 2016 State of the Market Report for the New York

ISO Markets. Potomac Econ n.d.
[143] New York Independent System Operator (NYISO). Locational Minimum Installed Capacity Requirements Study for the 2019-2020 Capability Year. 2019.
[144] New York Independent System Operator (NYISO). 2018 Load and Capacity Data: Gold Book 2018.
[145] Powerbridge. The Hudson Transmission Project. 2021.
[146] New York State Department of Public Service. Renewable Portfolio Standard, Case 03-E-0188 n.d.
[147] New York Independent System Operator (NYISO). 2018 Power Trends 2018.
[148] New York State. Reforming the Energy Vision n.d.
[149] New York Independent System Operator (NYISO). NYISO Operating Study Summer 2018 2018.
[150] Conlon T, Modi V, Waite M. The Value of Energy Flexibility: Integrating Wind Resources in New York State. Proc. ASME 2018 Int. Mech. Eng. Conf. Expo., Pittsburgh: 2018.
[151] US Energy Information Administration (EIA). Form EIA-860 detailed data with previous form data (EIA-860A/860B) 2022.
[152] New York State Energy Research and Development Authority (NYSERDA). Estimated Gasoline Sales, Beginning 1995. 2018.
[153] US Environmental Protection Agency. Light-Duty Vehicle CO2 and Fuel Economy Trends. 2017.
[154] US Department of Energy. Chevrolet Volt Fuel Economy Specifications. n.d.
[155] Jin C, Sheng X, Ghosh P. Optimized Electric Vehicle Charging With. IEEE J Sel Top Signal Process 2014;8:1063–72. https://doi.org/10.1109/JSTSP.2014.2336624.
[156] Wiser R, Jenni K, Seel J, Baker E, Hand M, Lantz E, et al. Expert elicitation survey on future wind energy costs. Nat Energy 2016;1. https://doi.org/10.1038/nenergy.2016.135.
[157] McLaren J, Gagnon P, Anderson K, Elgqvist E, Fu R, Remo T. Battery Energy Storage Market: Commerical Scale, Lithium-ion Projects in the US. Natl Renew Energy Lab 2016.
[158] International Renewable Energy Agency. Electricity Storage and Renewables: Costs and Markets to 2030 2017. https://www.irena.org/publications/2017/oct/electricity-storage-and-renewables-costs-and-markets.
[159] US Bureau of Economic Analysis. Total Gross Domstic Product for New York [NYNGSP], retrieved from FRED, Federal Reserve Bank of St. Louis n.d.
[160] Williams JH, DeBenedictis A, Ghanadan R, Mahone A, Moore J, Morrow WR, et al. 2050: The Pivotal Role of Electricity The Technology Path to Deep Greenho. Science (80-) 2012;53.
[161] Jacobson MZ, Delucchi MA, Cameron MA, Frew BA. Low-cost solution to the grid reliability problem with 100% penetration of intermittent wind, water, and solar for all purposes. Proc Natl Acad Sci 2015. https://doi.org/10.1073/pnas.1510028112.
[162] Bistline JE, Blanford GJ. More than one arrow in the quiver: Why “100% renewables” misses the mark. Proc Natl Acad Sci 2016;113:E3988–E3988. https://doi.org/10.1073/pnas.1603072113.
[163] Bird L, Lew D, Milligan M, Carlini EM, Estanqueiro A, Flynn D, et al. Wind and solar energy curtailment: A review of international experience. Renew Sustain Energy Rev 2016;65:577–86. https://doi.org/10.1016/j.rser.2016.06.082.

[164] Huertas-Hernando D, Farahmand H, Holttinen H, Kiviluoma J, Rinne E, Söder L, et al. Hydro power flexibility for power systems with variable renewable energy sources: an IEA Task 25 collaboration. Wiley Interdiscip Rev Energy Environ 2017. https://doi.org/10.1002/wene.220.
[165] Hirth L. The benefits of flexibility: The value of wind energy with hydropower. Appl Energy 2016. https://doi.org/10.1016/j.apenergy.2016.07.039.
[166] Cutululis NA, Farahmand H, Jaehnert S, Detlefsen N, Byriel IP, Sørensen P. Hydropower flexibility and transmission expansion to support integration of offshore wind. Offshore Wind Farms Technol. Des. Oper., 2016. https://doi.org/10.1016/B978-0-08-100779-2.00016-7.
[167] International Niagara Board of Control. Section 2: The 1950 Niagara Treaty n.d.
[168] Federal Energy Regulatory Commission. Draft Environmental Assessment for New Hydropower License: Blenheim-Gilboa Pumped Storage Project FERC Project No. 2685-029 2016. https://www.ferc.gov/sites/default/files/2020-06/P-2685EA.pdf.
[169] Canadian Water Office. Real Time Hydrometric Data for Niagara River at Fort Erie (02HA013) [ON] n.d.
[170] The Independent Electricity System Operator (IESO). Power Data -- Supply Overview n.d.
[171] MathWorks. MATLAB and Statistics Toolbox Release 2012b n.d.
[172] Gurobi Optimization Inc. Gurobi Optimizer Version 7.0 2016.
[173] New York State Energy Research and Development Authority (NYSERDA). Distributed Generation Integrated Data System n.d.
[174] Opalka W. Appeals Court Ratifies New York Capacity Zone. RTO Insid 2015. https://rtoinsider.com/rto/2d-cir-new-york-ferc-capacity-zone-14255/.
[175] The World Bank. Agricultural irrigated land (% of total agricultural land) - Ethiopia 2018. https://data.worldbank.org/indicator/AG.LND.IRIG.AG.ZS?locations=ET.
[176] The World Bank. Access to electricity (% of population) - Ethiopia n.d.
[177] International Energy Agency. Africa Energy Outlook 2019 2019.
[178] Fobi S, Deshpande V, Ondiek S, Modi V, Taneja J. A longitudinal study of electricity consumption growth in Kenya. Energy Policy 2018;123:569–78. https://doi.org/10.1016/j.enpol.2018.08.065.
[179] Lee K, Miguel E, Wolfram C. Experimental Evidence on the Demand for and Costs of Rural Electrification. 2016.
[180] Food and Agriculture Organization of the United Nations (FAO). Chapter 2: Crop Water Requirements n.d.
[181] Gebregziabher G, Namara RE, Holden S. Poverty reduction with irrigation investment: An empirical case study from Tigray, Ethiopia. Agric Water Manag 2009;96:1837–43. https://doi.org/10.1016/j.agwat.2009.08.004.
[182] Lee K, Miguel E, Wolfram C. Experimental evidence on the economics of rural electrification. J Polit Econ 2020;128:1523–65. https://doi.org/10.1086/705417.
[183] Small C. Spatiotemporal dimensionality and Time-Space characterization of multitemporal imagery. Remote Sens Environ 2012;124:793–809. https://doi.org/10.1016/j.rse.2012.05.031.
[184] Ozdogan M, Yang Y, Allez G, Cervantes C. Remote sensing of irrigated agriculture: Opportunities and challenges. Remote Sens 2010;2:2274–304.

https://doi.org/10.3390/rs2092274.
[185] Vogels MFA, de Jong SM, Sterk G, Addink EA. Mapping irrigated agriculture in complex landscapes using SPOT6 imagery and object-based image analysis – A case study in the Central Rift Valley, Ethiopia –. Int J Appl Earth Obs Geoinf 2019;75:118–29. https://doi.org/10.1016/j.jag.2018.07.019.
[186] Vogels MFA, de Jong SM, Sterk G, Douma H, Addink EA. Spatio-temporal patterns of smallholder irrigated agriculture in the horn of Africa using GEOBIA and Sentinel-2 imagery. Remote Sens 2019;11. https://doi.org/10.3390/rs11020143.
[187] Chandrasekharan KM, Subasinghe C, and, Haileslassie A. Mapping Irrigated and Rainfed Agriculture in Ethiopia (2015-2016) using Remote Sensing Methods. 2021.
[188] Jain M, Mondal P, DeFries RS, Small C, Galford GL. Mapping cropping intensity of smallholder farms: A comparison of methods using multiple sensors. Remote Sens Environ 2013;134:210–23. https://doi.org/10.1016/j.rse.2013.02.029.
[189] Ferrant S, Selles A, Le Page M, Herrault PA, Pelletier C, Al-Bitar A, et al. Detection of irrigated crops from Sentinel-1 and Sentinel-2 data to estimate seasonal groundwater use in South India. Remote Sens 2017;9. https://doi.org/10.3390/rs9111119.
[190] Chen Y, Lu D, Luo L, Pokhrel Y, Deb K, Huang J, et al. Detecting irrigation extent, frequency, and timing in a heterogeneous arid agricultural region using MODIS time series, Landsat imagery, and ancillary data. Remote Sens Environ 2018;204:197–211. https://doi.org/10.1016/j.rse.2017.10.030.
[191] Pageot Y, Baup F, Inglada J, Baghdadi N, Demarez V. Detection of irrigated and rainfed crops in temperate areas using sentinel-1 and sentinel-2 time series. Remote Sens 2020;12:1–19. https://doi.org/10.3390/RS12183044.
[192] Bazzi H, Baghdadi N, Amin G, Fayad I, Zribi M, Demarez V, et al. An operational framework for mapping irrigated areas at plot scale using sentinel-1 and sentinel-2 data. Remote Sens 2021;13:1–28. https://doi.org/10.3390/rs13132584.
[193] Elwan E, Page M Le, Jarlan L, Baghdadi N, Brocca L, Modanesi S, et al. Irrigation Mapping on Two Contrasted Climatic Contexts Using Sentinel-1 and Sentinel-2 Data. Water (Switzerland) 2022;14:1–13. https://doi.org/10.3390/w14050804.
[194] Government of Catalonia (GENCAT). SIGPAC Download 2021. http://agricultura.gencat.cat/ca/serveis/cartografia-sig/aplicatius-tematics-geoinformacio/sigpac/descarregues/ (accessed August 9, 2021).
[195] Bolognesi SF, Pasolli E, Belfiore OR, De Michele C, D'Urso G. Harmonized landsat 8 and sentinel-2 time series data to detect irrigated areas: An application in Southern Italy. Remote Sens 2020;12. https://doi.org/10.3390/RS12081275.
[196] Liu L, Xiao X, Qin Y, Wang J, Xu X, Hu Y, et al. Mapping cropping intensity in China using time series Landsat and Sentinel-2 images and Google Earth Engine. Remote Sens Environ 2020;239:111624. https://doi.org/10.1016/j.rse.2019.111624.
[197] Agastya C, Ghebremusse S, Anderson I, Reed C, Vahabi H, Todeschini A. Self-supervised Contrastive Learning for Irrigation Detection in Satellite Imagery 2021.
[198] Ragettli S, Herberz T, Siegfried T. An unsupervised classification algorithm for multi-temporal irrigated area mapping in Central Asia. Remote Sens 2018;10:1–23. https://doi.org/10.3390/rs10111823.
[199] BENBAHRİA Z, SEBARİ İ, HAJJİ H, SMİEJ MF. Intelligent Mapping of Irrigated Areas From

Landsat 8 Images Using Transfer Learning. Int J Eng Geosci 2021;6:40–50. https://doi.org/10.26833/ijeg.681312.
[200] Zhao H, Duan S, Liu J, Sun L, Reymondin L. Evaluation of five deep learning models for crop type mapping using sentinel-2 time series images with missing information. Remote Sens 2021;13. https://doi.org/10.3390/rs13142790.
[201] Colligan T, Ketchum D, Brinkerhoff D, Maneta M. A Deep Learning Approach to Mapping Irrigation Using Landsat: IrrMapper U-Net. IEEE Trans Geosci Remote Sens 2022;60. https://doi.org/10.1109/TGRS.2022.3175635.
[202] Ketchum D, Jencso K, Maneta MP, Melton F, Jones MO, Huntington J. IrrMapper: A machine learning approach for high resolution mapping of irrigated agriculture across the Western U.S. Remote Sens 2020;12:1–23. https://doi.org/10.3390/rs12142328.
[203] Xie Y, Lark TJ. Mapping annual irrigation from Landsat imagery and environmental variables across the conterminous United States. Remote Sens Environ 2021;260:112445. https://doi.org/10.1016/j.rse.2021.112445.
[204] Xie Y, Lark TJ, Brown JF, Gibbs HK. Mapping irrigated cropland extent across the conterminous United States at 30 m resolution using a semi-automatic training approach on Google Earth Engine. ISPRS J Photogramm Remote Sens 2019;155:136–49. https://doi.org/10.1016/j.isprsjprs.2019.07.005.
[205] Brown JF, Pervez MS. Merging remote sensing data and national agricultural statistics to model change in irrigated agriculture. Agric Syst 2014;127:28–40. https://doi.org/10.1016/j.agsy.2014.01.004.
[206] Deines JM, Kendall AD, Hyndman DW. Annual Irrigation Dynamics in the U.S. Northern High Plains Derived from Landsat Satellite Data. Geophys Res Lett 2017;44:9350–60. https://doi.org/10.1002/2017GL074071.
[207] Xu T, Deines JM, Kendall AD, Basso B, Hyndman DW. Addressing challenges for mapping irrigated fields in subhumid temperate regions by integrating remote sensing and hydroclimatic data. Remote Sens 2019;11. https://doi.org/10.3390/rs11030370.
[208] Deines JM, Kendall AD, Crowley MA, Rapp J, Cardille JA, Hyndman DW. Mapping three decades of annual irrigation across the US High Plains Aquifer using Landsat and Google Earth Engine. Remote Sens Environ 2019;233:111400. https://doi.org/10.1016/j.rse.2019.111400.
[209] Rufin P, Müller D, Schwieder M, Pflugmacher D, Hostert P. Landsat time series reveal simultaneous expansion and intensification of irrigated dry season cropping in Southeastern Turkey. J Land Use Sci 2021;16:94–110. https://doi.org/10.1080/1747423X.2020.1858198.
[210] dela Torre DMG, Gao J, Macinnis-Ng C, Shi Y. Phenology-based delineation of irrigated and rain-fed paddy fields with Sentinel-2 imagery in Google Earth Engine. Geo-Spatial Inf Sci 2021;24:695–710. https://doi.org/10.1080/10095020.2021.1984183.
[211] Zohaib M, Kim H, Choi M. Detecting global irrigated areas by using satellite and reanalysis products. Sci Total Environ 2019;677:679–91. https://doi.org/10.1016/j.scitotenv.2019.04.365.
[212] Siebert S, Döll P, Hoogeveen J, Faures JM, Frenken K, Feick S. Development and validation of the global map of irrigation areas. Hydrol Earth Syst Sci 2005;9:535–47. https://doi.org/10.5194/hess-9-535-2005.

[213] Thenkabail PS, Biradar CM, Noojipady P, Dheeravath V, Li Y, Velpuri M, et al. Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. Int J Remote Sens 2009;30:3679–733. https://doi.org/10.1080/01431160802698919.
[214] Salmon JM, Friedl MA, Frolking S, Wisser D, Douglas EM. Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data. Int J Appl Earth Obs Geoinf 2015;38:321–34. https://doi.org/10.1016/J.JAG.2015.01.014.
[215] Meier J, Zabel F, Mauser W. A global approach to estimate irrigated areas - A comparison between different data and statistics. Hydrol Earth Syst Sci 2018;22:1119–33. https://doi.org/10.5194/hess-22-1119-2018.
[216] ESA. Climate Change Initiative Land Cover Time Series 2017.
[217] Landgrebe D. Hyperspectral Image Data Analysis. IEEE Signal Process Mag 2002:17–28.
[218] Ambika AK, Wardlow B, Mishra V. Remotely sensed high resolution irrigated area mapping in India for 2000 to 2015. Sci Data 2016;3:1–14. https://doi.org/10.1038/sdata.2016.118.
[219] Cardot H, Faivre R, Goulard M. Functional approaches for predicting land use with the temporal evolution of coarse resolution remote sensing data. J Appl Stat 2003;30:1185–99. https://doi.org/10.1080/0266476032000107187.
[220] Kerdiles H, Grondona MO. NOAA-AVHRR NDVI decomposition and subpixel classification using linear mixing in the argentinean pampa. Int J Remote Sens 1995;16:1303–25. https://doi.org/10.1080/01431169508954478.
[221] Lunetta RS, Knight JF, Ediriwickrema J, Lyon JG, Worthy LD. Land-cover change detection using multi-temporal MODIS NDVI data. Remote Sens Environ 2006;105:142–54. https://doi.org/10.1016/j.rse.2006.06.018.
[222] Verbesselt J, Hyndman R, Newnham G, Culvenor D. Detecting trend and seasonal changes in satellite image time series. Remote Sens Environ 2010;114:106–15. https://doi.org/10.1016/j.rse.2009.08.014.
[223] Wardlow BD, Egbert SL. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. Remote Sens Environ 2008;112:1096–116. https://doi.org/10.1016/J.RSE.2007.07.019.
[224] Abd El-Kawy OR, Rød JK, Ismail HA, Suliman AS. Land use and land cover change detection in the western Nile delta of Egypt using remote sensing data. Appl Geogr 2011;31:483–94. https://doi.org/10.1016/j.apgeog.2010.10.012.
[225] Ma Z, Hart MM, Redmond RL. Mapping vegetation across large geographic areas: Integration of remote sensing and GIS to classify multisource data. Photogramm Eng Remote Sensing 2001;67:295–307.
[226] Chen JM, Cihlar J. Retrieving leaf area index of boreal conifer forests using Landsat TM images. Remote Sens Environ 1996;55:153–62. https://doi.org/10.1016/0034-4257(95)00195-6.
[227] Solano-Correa YT, Bovolo F, Bruzzone L, Fernandez-Prieto D. Spatio-temporal evolution of crop fields in Sentinel-2 Satellite Image Time Series. 2017 9th Int Work Anal Multitemporal Remote Sens Images, MultiTemp 2017 2017:3–6. https://doi.org/10.1109/Multi-Temp.2017.8035236.

[228] Gao Q, Zribi M, Escorihuela MJ, Baghdadi N, Segui PQ. Irrigation mapping using Sentinel-1 time series at field scale. Remote Sens 2018;10:1–18. https://doi.org/10.3390/rs10091495.
[229] Veloso A, Mermoz S, Bouvet A, Le Toan T, Planells M, Dejoux JF, et al. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. Remote Sens Environ 2017;199:415–26. https://doi.org/10.1016/j.rse.2017.07.015.
[230] Xiong J, Thenkabail PS, Tilton JC, Gumma MK, Teluguntla P, Oliphant A, et al. Nominal 30-m cropland extent map of continental Africa by integrating pixel-based and object-based algorithms using Sentinel-2 and Landsat-8 data on google earth engine. Remote Sens 2017;9:1–27. https://doi.org/10.3390/rs9101065.
[231] Phalke AR, Özdoğan M, Thenkabail PS, Erickson T, Gorelick N, Yadav K, et al. Mapping croplands of Europe, Middle East, Russia, and Central Asia using Landsat, Random Forest, and Google Earth Engine. ISPRS J Photogramm Remote Sens 2020;167:104–22. https://doi.org/10.1016/j.isprsjprs.2020.06.022.
[232] Tong X, Brandt M, Hiernaux P, Herrmann S, Rasmussen LV, Rasmussen K, et al. The forgotten land use class: Mapping of fallow fields across the Sahel using Sentinel-2. Remote Sens Environ 2020;239. https://doi.org/10.1016/j.rse.2019.111598.
[233] Buchorn M, Smets B, Bertels L, Roo B De, Lesiv M, Tsendbazar N-E, et al. Copernicus Global Land Service: Land Cover 100m: Version 3 Globe 2015-2019: Product User Manual 2020. https://doi.org/10.5281/zenodo.3938963.Dissemination.
[234] Tsendbazar NE, Herold M, de Bruin S, Lesiv M, Fritz S, Van De Kerchove R, et al. Developing and applying a multi-purpose land cover validation dataset for Africa. Remote Sens Environ 2018;219:298–309. https://doi.org/10.1016/j.rse.2018.10.025.
[235] Blickensdörfer L, Schwieder M, Pflugmacher D, Nendel C, Erasmi S, Hostert P. Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. Remote Sens Environ 2022;269. https://doi.org/10.1016/j.rse.2021.112831.
[236] Schulz C, Holtgrave AK, Kleinschmit B. Large-scale winter catch crop monitoring with Sentinel-2 time series and machine learning–An alternative to on-site controls? Comput Electron Agric 2021;186:106173. https://doi.org/10.1016/j.compag.2021.106173.
[237] Gumma MK, Tummala K, Dixit S, Collivignarelli F, Holecz F, Kolli RN, et al. Crop type identification and spatial mapping using Sentinel-2 satellite data with focus on field-level information. Geocarto Int 2020;0:1–17. https://doi.org/10.1080/10106049.2020.1805029.
[238] Adams JB, Gillespie AR. Remote sensing of landscapes with spectral images: A physical modeling approach. Cambridge, UK: Cambridge University Press; 2006.
[239] Plaza A, Martínez P, Pérez R, Plaza J. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. IEEE Trans Geosci Remote Sens 2004;42:650–63. https://doi.org/10.1109/TGRS.2003.820314.
[240] Adams JB. Imaging spectroscopy : Interpretation based on spectral mixture analysis. Remote Geochemical Anal Elem Mineral Compos 1993:145–66.
[241] Elmore AJ, Mustard JF, Manning SJ, Lobell DB. Quantifying Vegetation Change in Semiarid Environments. Remote Sens Environ 2000;73:87–102. https://doi.org/10.1016/s0034-

4257(00)00100-0.
[242] Adams JB, Smith MO, Johnson PE. Correction [to “Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site” by John B. Adams, Milton O. Smith, and Paul E. Johnson]. J Geophys Res 1986;91:10513. https://doi.org/10.1029/jb091ib10p10513.
[243] Roberts DA, Gardner M, Church R, Ustin S, Scheer G, Green RO. Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. Remote Sens Environ 1998;65:267–79. https://doi.org/10.1016/S0034-4257(98)00037-6.
[244] Dennison PE, Roberts DA, Peterson SH. Spectral shape-based temporal compositing algorithms for MODIS surface reflectance data. Remote Sens Environ 2007;109:510–22. https://doi.org/10.1016/j.rse.2007.02.009.
[245] Somers B, Delalieux S, Verstraeten WW, Verbesselt J, Lhermitte S, Coppin P. Magnitude- and shape-related feature integration in hyperspectral mixture analysis to monitor weeds in citrus orchards. IEEE Trans Geosci Remote Sens 2009;47:3630–42. https://doi.org/10.1109/TGRS.2009.2024207.
[246] Tompkins S, Mustard JF, Pieters CM, Forsyth DW. Optimization of endmembers for spectral mixture analysis. Remote Sens Environ 1997;59:472–89. https://doi.org/10.1016/S0034-4257(96)00122-8.
[247] Dudley KL, Dennison PE, Roth KL, Roberts DA, Coates AR. A multi-temporal spectral library approach for mapping vegetation species across spatial and temporal phenological gradients. Remote Sens Environ 2015;167:121–34. https://doi.org/10.1016/j.rse.2015.05.004.
[248] Rogge DM, Rivard B, Jinkai Z, Jilu F. Iterative spectral unmixing for optimizing per-pixel endmember sets. IEEE Trans Geosci Remote Sens 2006;44:3725–35. https://doi.org/10.1109/TGRS.2006.881123.
[249] Tane Z, Roberts D, Veraverbeke S, Casas Á, Ramirez C, Ustin S. Evaluating endmember and band selection techniques for multiple endmember spectral mixture analysis using post-fire imaging spectroscopy. Remote Sens 2018;10:1–21. https://doi.org/10.3390/rs10030389.
[250] Song C. Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporate endmember variability? Remote Sens Environ 2005;95:248–63. https://doi.org/10.1016/j.rse.2005.01.002.
[251] Johnson PE, Smith MO, Taylor-George S, Adams JB. Semiempirical Method for Analysis of the Reflectance Spectra of Binary Mineral Mixtures. J Geophys Res 1983;88:3557–61. https://doi.org/10.1029/JB088iB04p03557.
[252] Johnson, P. E., Smith, M. O., & Adams JB. Quantitative analysis of planetary reflectance spectra with principal components analysis. Lunar Planet Institute, NASA, Am Geophys Union, Al, Lunar Planet Sci Conf 15th, Houston, TX, Mar 12-16, 1984 J Geophys Res Suppl (ISSN 0148-0227), Vol 90, Feb 15, 1985, p C805-C810 1985.
[253] Johnson TD, Belitz K. A remote sensing approach for estimating the location and rate of urban irrigation in semi-arid climates. J Hydrol 2012;414–415:86–98. https://doi.org/10.1016/j.jhydrol.2011.10.016.
[254] Benhadj I, Duchemin B, Maisongrande P, Simonneaux V, Khabba S, Chehbouni A. Automatic unmixing of MODIS multi-temporal data for inter-annual monitoring of land

use at a regional scale (Tensift, Morocco). Int J Remote Sens 2012;33:1325–48. https://doi.org/10.1080/01431161.2011.564220.
[255] Small C. The Landsat ETM+ spectral mixing space. Remote Sens Environ 2004;93:1–17. https://doi.org/10.1016/j.rse.2004.06.007.
[256] Small C. A global analysis of urban reflectance. Int J Remote Sens 2005;26:661–81. https://doi.org/10.1080/01431160310001654950.
[257] Small C, Milesi C. Multi-scale standardized spectral mixture models. Remote Sens Environ 2013;136:442–54. https://doi.org/10.1016/j.rse.2013.05.024.
[258] World Bank. The World Bank in Ethiopia: An Overview. 2019.
[259] Mendes DM, Paglietti L. Ethiopia: Irrigation market brief. 2015.
[260] Food and Agriculture Organization of the United Nations (FAO). AQUASTAT Main Database 2016.
[261] Awulachew, SB., Erkossa, T. and Namara R. Irrigation potential in Ethiopia: Constraints and opportunities for enhancing the system. Research Report, International Water Management Institute, Addis Ababa. Int Water Manag Inst 2010:1–59.
[262] Girma MM, Awulachew SB. Irrigation Practices in Ethiopia: Characteristics of selected Irrigation Schemes. 2007.
[263] Kenea O, Balkew M, Gebre-Michael T. Environmental factors associated with larval habitats of anopheline mosquitoes (diptera: Culicidae) in irrigation and major drainage areas in the middle course of the rift valley, central ethiopia. J Vector Borne Dis 2011;48:85–92.
[264] Yami M. Irrigation projects in Ethiopia: What can be done to enhance effectiveness under “challenging contexts”? Int J Sustain Dev World Ecol 2016;23:132–42. https://doi.org/10.1080/13504509.2015.1057628.
[265] Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, et al. The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. Nature 2015;2:1–21. https://doi.org/10.1038/sdata.2015.66.
[266] Wiggins S, Glover D, Dorgan A. Agricultural innovation for smallholders in sub-Saharan Africa. 2021.
[267] Conlon T, Wu Y, Small C, Siddiqui H, Adkins E, Modi V. A Novel Method of Irrigation Detection and Estimation of the Effects of Productive Electricity Demands on Energy System Planning. AGU Fall Meet. Abstr., vol. 2020, 2020, p. GC034-08.
[268] Shahriar Pervez M, Budde M, Rowland J. Mapping irrigated areas in Afghanistan over the past decade using MODIS NDVI. Remote Sens Environ 2014;149:155–65. https://doi.org/10.1016/J.RSE.2014.04.008.
[269] Huete AR, Didan K, Van Leeuwen W. Modis Vegetation Index. Veg Index Phenol Lab 1999;3(1):129.
[270] Phiri D, Simwanda M, Salekin S, Ryirenda VR, Murayama Y, Ranagalage M, et al. Remote Sensing Sentinel-2 Data for Land Cover / Use Mapping: A Review. Remote Sens 2019;42:14.
[271] Lecun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44. https://doi.org/10.1038/nature14539.
[272] Zhong L, Hu L, Zhou H, Tao X. Deep learning based winter wheat mapping using statistical data as ground references in Kansas and northern Texas, US. Remote Sens Environ

2019;233:111411. https://doi.org/10.1016/J.RSE.2019.111411.
[273] Hu W, Huang Y, Wei L, Zhang F, Li H. Deep convolutional neural networks for hyperspectral image classification. J Sensors 2015;2015. https://doi.org/10.1155/2015/258619.
[274] Yue J, Zhao W, Mao S, Liu H. Spectral-spatial classification of hyperspectral images using deep convolutional neural networks. Remote Sens Lett 2015;6:468–77. https://doi.org/10.1080/2150704X.2015.1047045.
[275] Mou L, Member S, Bruzzone L, Zhu XX. Learning Spectral-Spatial-Temporal Features via a Recurrent Convolutional Neural Network for Change Detection in Multispectral Imagery 2018:1–12.
[276] Rustowicz R, Cheong R, Wang L, Ermon S, Burke M, Lobell D. Semantic Segmentation of Crop Type in Africa: A Novel Dataset and Analysis of Deep Learning Methods. CVPR Work 2019;1:75–82.
[277] Rußwurm M, Korner M. Temporal Vegetation Modelling Using Long Short-Term Memory Networks for Crop Identification from Medium-Resolution Multi-spectral Satellite Images. IEEE Comput Soc Conf Comput Vis Pattern Recognit Work 2017;2017-July:1496–504. https://doi.org/10.1109/CVPRW.2017.193.
[278] Zhou Y, Luo J, Feng L, Yang Y, Chen Y, Wu W. Long-short-term-memory-based crop classification using high-resolution optical images and multi-temporal SAR data. GIScience Remote Sens 2019;56:1170–91. https://doi.org/10.1080/15481603.2019.1628412.
[279] Xu J, Zhu Y, Zhong R, Lin Z, Xu J, Jiang H, et al. DeepCropMapping: A multi-temporal deep learning approach with improved spatial generalizability for dynamic corn and soybean mapping. Remote Sens Environ 2020;247:111946. https://doi.org/10.1016/j.rse.2020.111946.
[280] Xu J, Yang J, Xiong X, Li H, Huang J, Ting KC, et al. Towards interpreting multi-temporal deep learning models in crop mapping. Remote Sens Environ 2021;264:112599. https://doi.org/10.1016/j.rse.2021.112599.
[281] Rußwurm M, Körner M. Self-attention for raw optical Satellite Time Series Classification. ISPRS J Photogramm Remote Sens 2020;169:421–35. https://doi.org/10.1016/j.isprsjprs.2020.06.006.
[282] Bazi Y, Bashmal L, Al Rahhal MM, Dayil R Al, Ajlan N Al. Vision transformers for remote sensing image classification. Remote Sens 2021;13:1–20. https://doi.org/10.3390/rs13030516.
[283] Li Y, Zhang H, Xue X, Jiang Y, Shen Q. Deep learning for remote sensing image classification: A survey. Wiley Interdiscip Rev Data Min Knowl Discov 2018;8:1–17. https://doi.org/10.1002/widm.1264.
[284] Jia Deng, Wei Dong, Socher R, Li-Jia Li, Kai Li, Li Fei-Fei. ImageNet: A large-scale hierarchical image database 2009:248–55. https://doi.org/10.1109/cvprw.2009.5206848.
[285] Tao C, Qi J, Lu W, Wang H, Li H. Self-Supervised Paradigm Under Limited Labeled Samples 2020;19:1–5.
[286] Banerjee B, Bovolo F, Bhattacharya A, Bruzzone L, Chaudhuri S, Mohan BK. A new self-training-based unsupervised satellite image classification technique using cluster ensemble strategy. IEEE Geosci Remote Sens Lett 2015;12:741–5.

https://doi.org/10.1109/LGRS.2014.2360833.
[287] Saha S, Solano-Correa YT, Bovolo F, Bruzzone L. Unsupervised deep learning based change detection in Sentinel-2 images. 2019 10th Int Work Anal Multitemporal Remote Sens Images, MultiTemp 2019 2019:0–3. https://doi.org/10.1109/Multi-Temp.2019.8866899.
[288] Yu X, Wu X, Luo C, Ren P. Deep learning in remote sensing scene classification: a data augmentation enhanced convolutional neural network framework. GIScience Remote Sens 2017;54:741–58. https://doi.org/10.1080/15481603.2017.1323377.
[289] Stivaktakis R, Tsagkatakis G, Tsakalides P. Deep Learning for Multilabel Land Cover Scene Categorization Using Data Augmentation. IEEE Geosci Remote Sens Lett 2019;16:1031–5. https://doi.org/10.1109/LGRS.2019.2893306.
[290] Bazzi H, Baghdadi N, Fayad I, Zribi M, Belhouchette H, Demarez V. Near real-time irrigation detection at plot scale using sentinel-1 data. Remote Sens 2020;12. https://doi.org/10.3390/RS12091456.
[291] de Lima RP, Marfurt K. Convolutional neural network for remote-sensing scene classification: Transfer learning analysis. Remote Sens 2020;12. https://doi.org/10.3390/rs12010086.
[292] Main-Knorn M, Pflug B, Louis J, Debaecker V, Müller-Wilm U, Gascon F. Sen2Cor for Sentinel-2. In: Bruzzone L, editor. Image Signal Process. Remote Sens. XXIII, vol. 10427, SPIE; 2017, p. 37–48. https://doi.org/10.1117/12.2278218.
[293] Hagen A, Strube J, Haide I, Kahn J, Jackson S, Hainje C. A Proposed High Dimensional Kolmogorov-Smirnov Distance. Mach Learn Phys Sci Work 34th Conf Neural Inf Process Syst 2020:1–6.
[294] Breiman L. Random Forests. Mach Learn 2001:1–28. https://doi.org/10.1201/9780429469275-8.
[295] Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features 2019.
[296] Wang Z, Yan W, Oates T. Time series classification from scratch with deep neural networks: A strong baseline. Proc Int Jt Conf Neural Networks 2017;2017-May:1578–85. https://doi.org/10.1109/IJCNN.2017.7966039.
[297] King G, Zeng L. Logistic regression in rare events data. Polit Anal 2001;9:137–63. https://doi.org/10.18637/jss.v008.i02.
[298] Kingma DP, Ba JL. Adam: A method for stochastic optimization. 3rd Int Conf Learn Represent ICLR 2015 - Conf Track Proc 2015:1–15.
[299] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. Int J Comput Vis 2020;128:336–59. https://doi.org/10.1007/s11263-019-01228-7.
[300] Small C. Grand Challenges in Remote Sensing Image Analysis and Classification. Front Remote Sens 2021;1:1–4. https://doi.org/10.3389/frsen.2020.605220.
[301] Zanaga D, Van De Kerchove R, De Keersmaecker W, Souverijns N, Brockmann C, Quast R, et al. ESA WorldCover 10 m 2020 v100 2021. https://doi.org/10.5281/ZENODO.5571936.
[302] Karra K, Kontgis C, Statman-Weil Z, Mazzariello JC, Mathis M, Steven P, et al. Global land use / land cover with Sentinel 2 and deep learning 2021.
[303] OpenStreetMaps Contributors. OpenStreetMaps 2021.

https://planet.openstreetmap.org/.
[304] Corbane C, Sabo F, Politis P, Syrris V. GHS-BUILT-S2 R2020A - builtup grid derived from Sentinel-2 global image composite for reference year 2018 using Convolutional Neural (GHS-S2Net). Eur Comm Jt Res Cent 2020. https://doi.org/doi:10.2905/016D1A34- B184-42DC-B586-E10B915DD863.
[305] Corbane C, Syrris V, Sabo F, Politis P, Melchiorri M, Pesaresi M, et al. Convolutional neural networks for global human settlements mapping from Sentinel-2 satellite imagery. Neural Comput Appl 2021;33:6697–720. https://doi.org/10.1007/s00521-020-05449-7.
[306] Pekel JF, Cottam A, Gorelick N, Belward AS. High-resolution mapping of global surface water and its long-term changes. Nature 2016;540:418–22. https://doi.org/10.1038/nature20584.
[307] Nussbaumer P, Bazilian M, Modi V. Measuring energy poverty: Focusing on what matters. Renew Sustain Energy Rev 2012;16:231–43. https://doi.org/10.1016/j.rser.2011.07.150.
[308] Modi V, McDade S, Lallement D, Saghir J. Energy Services for the Millennium Development Goals. UN Millenium Dev Proj 2005:116.
[309] OECD/IEA. IEA Statistics 2020.
[310] Intergovernmental Panel on Climate Change. Drivers, Trends and Mitigation. 2015. https://doi.org/10.1017/cbo9781107415416.011.
[311] DeCarolis J, Daly H, Dodds P, Keppo I, Li F, McDowall W, et al. Formalizing best practice for energy system optimization modelling. Appl Energy 2017;194:184–98. https://doi.org/10.1016/j.apenergy.2017.03.001.
[312] Häfele W, Rogner H-H. A Technical Appraisal of the IIASA Energy Scenarios? A Rebuttal. Policy Sci 1984;17:341–65.
[313] Alfaro JF, Miller S, Johnson JX, Riolo RR. Improving rural electricity system planning: An agent-based model for stakeholder engagement and decision making. Energy Policy 2017;101:317–31. https://doi.org/10.1016/j.enpol.2016.10.020.
[314] Poncelet K, Delarue E, Six D, Duerinck J, D'haeseleer W. Impact of the level of temporal and operational detail in energy-system planning models. Appl Energy 2016;162:631–43. https://doi.org/10.1016/j.apenergy.2015.10.100.
[315] Rockström J, Sachs J, Schmidt-traub G, Öhman M. Sustainable development and planetary boundaries. Sustain Dev Solut Netw 2013:1–46.
[316] Kahrl F, Williams J, Jianhua D, Junfeng H. Challenges to China's transition to a low carbon electricity system. Energy Policy 2011;39:4032–41. https://doi.org/10.1016/j.enpol.2011.01.031.
[317] Zhang X, Shahidehpour M, Alabdulwahab A, Abusorrah A. Optimal Expansion Planning of Energy Hub with Multiple Energy Infrastructures. IEEE Trans Smart Grid 2015;6:1–10. https://doi.org/10.1109/TSG.2015.2390640.
[318] Castellano A, Kendall A, Nikomarv M, Swemmer T. Powering Africa. 2015. https://doi.org/10.4102/sajs.v108i1/2.1062.
[319] Avila N, Carvallo J, Shaw B, Kammen DM. The energy challenge in sub-Saharan Africa: Generating energy for sustainable and equitable development. Oxfam Res Backgrounder Ser 2017:100.

[320] Energy Information Administration (EIA). Energy Outlook 2020 2019;58:7250–7. https://doi.org/10.1128/AAC.03728-14.
[321] EIA. Electric Power Annual 2018. 2019.
[322] QSEL. Shared Solar n.d. https://qsel.columbia.edu/shared-solar/ (accessed April 3, 2020).
[323] Soto D, Basinger M, Rodriguez-Sanchez S, Adkins E, Menon R, Owczarek N, et al. A prepaid architecture for solar electricity delivery in rural areas. ACM Int Conf Proceeding Ser 2012:130–8. https://doi.org/10.1145/2160673.2160691.
[324] Waite M, Modi V. Electricity Load Implications of Space Heating Decarbonization Pathways. Joule 2019;4:376–94. https://doi.org/10.1016/j.joule.2019.11.011.
[325] Wolfram C. Good Morning Vietnam, Energy-wise. 2018.
[326] Kapadia K. Productive uses of renewable energy. A Review of Four Bank-GEF Projects. 2004.
[327] Mainuddin M, Das Gupta A, Onta PR. Optimal crop planning model for an existing groundwater irrigation project in Thailand. Agric Water Manag 1997;33:43–62. https://doi.org/10.1016/S0378-3774(96)01278-4.
[328] Bose RK, Shukla M. Elasticities of electricity demand in India. Energy Policy 1999;27:137–46. https://doi.org/10.1016/S0301-4215(99)00011-7.
[329] Fobi S, Kocaman AS, Taneja J, Modi V. A scalable framework to measure the impact of spatial heterogeneity on electrification. Energy Sustain Dev 2021;60:67–81. https://doi.org/10.1016/j.esd.2020.12.005.
[330] European Commission. Photovoltaic Geographical Information System 2021.
[331] Modi V. Sample residential consumption timeseries for 45 low-consuming households. n.d.
[332] US Energy Information Administration (EIA). Annual household site end-use consumption by fuel in the U.S. 2018:8. https://www.eia.gov/consumption/residential/data/2015/c&e/pdf/ce4.1.pdf.
[333] US Energy Information Administration (EIA). Commercial Building Energy Consumption Survey (CBECS) 2018. https://www.eia.gov/consumption/commercial/data/2018/.
[334] Hydro-Quebec. Hydro Quebec Annual Report 2019 2020. https://www.hydroquebec.com/data/documents-donnees/pdf/annual-report.pdf.
[335] New York State Climate Action Council. Technical Advisory Group, Integration Analysis -- Key Drivers and Outputs (updated December 30, 2021) 2021. https://climate.ny.gov/-/media/Project/Climate/Files/IA-Tech-Supplement-Annex-2-Key-Drivers-Outputs.ashx.
[336] Wakjira MT, Peleg N, Anghileri D, Molnar D, Alamirew T, Six J, et al. Rainfall seasonality and timing: implications for cereal crop production in Ethiopia. Agric For Meteorol 2021;310:108633. https://doi.org/10.1016/j.agrformet.2021.108633.