TECHNICAL DOCUMENTATION

UNIVERSITY OF CALIFORNIA
Santa Barbara

Co-Locating a Power Couple: Retrofitting Existing Wind Projects with Solar PV in the U.S.

A Capstone Project submitted in partial satisfaction of the requirements for the degree of
Master of Environmental Data Science
for the
Bren School of Environmental Science & Management

by
Michelle Lam
Colleen McCamy
Alessandra Vidal Meza

Advisor: Grace Wu, UC Santa Barbara, Environmental Studies Program

Committee in Charge:
Grace Wu, UC Santa Barbara, Environmental Studies Program
Naomi Tague, Bren School of Environmental Science & Management
Ruth Oliver, Bren School of Environmental Science & Management

June 2023

Co-Locating a Power Couple: Retrofitting Existing Wind Projects with Solar PV in the U.S.

As developers of this Capstone Project documentation, we archive this documentation on the Bren School's website such that the results of our research are available for all to read. Our signatures on the document signify our joint responsibility to fulfill the archiving standards set by the Bren School of Environmental Science & Management.

Michelle Lam Date

Colleen McCamy Date

Alessandra Vidal Meza Date

The Bren School of Environmental Science & Management produces professionals with unrivaled training in environmental science and management who will devote their unique skills to the diagnosis, assessment, mitigation, prevention, and remedy of the environmental problems of today and the future. A guiding principle of the School is that the analysis of environmental problems requires quantitative training in more than one discipline and an awareness of the physical, biological, social, political, and economic consequences that arise from scientific or technological decisions.

The Capstone Project is required of all students in the Master of Environmental Data Science (MEDS) Program. The project is a six-month-long activity in which small groups of students contribute to data science practices, products or analyses that address a challenge or need related to a specific environmental issue. This MEDS Capstone Project Technical Documentation is authored by MEDS students and has been reviewed and approved by:

Dr. Grace Wu Date

Dr. Ruth Oliver Date

Acknowledgements
This study could not have been completed without the support, mentorship, and collaboration of several people at UC Santa Barbara and beyond. A special thank you to advisors and committee members: Dr. Grace Wu, Dr. Naomi Tague, and Dr. Ruth Oliver, as well as Dr. Ranjit Deshmukh, Jamie Montgomery, Sam Csik, Kat Le, and Brad Hill. This study builds on findings from the *Power of Place* study by The Nature Conservancy. A special thank you to *Power of Place* collaborators: Joe Fargione, Nels Johnson, Christel Hiltibran, Chris Hise, and Liz Kalies, as well as consultants Ryan Jones from Evolved Energy Research and Emily Leslie form Montara Mountain Energy.

Table of Contents

1.0 Abstract

The demand for renewable energy such as wind and solar power in the United States is rapidly accelerating to help meet the ambitious goal of a net-zero power grid by 2035. However, a barrier to adding new wind and solar photovoltaic (PV) development lies in delays of constructing new transmission infrastructure. Co-location of renewable energy sites by adding solar PV to existing wind projects can reduce the need for new transmission infrastructure development while increasing profit from the project. The study used an nonlinear optimization model to find the optimal system size of solar PV to be added at each of the identified over 1,200 existing wind projects in the contiguous U.S. Using environmental sensitivity scores and estimated annual levelized-costs, the analysis then ranked the co-location projects to highlight sites that may benefit most from co-location. This study provides initial evidence towards co-location as a viable and efficient solution to reaching U.S. carbon-free energy goals. The nonlinear optimization model proposed can be generalized to quantify and examine co-location in other scenarios; like the implementation of co-location at sites that have not yet been developed with wind or solar.

2.0 Executive Summary

Carbon emissions are a leading contributor to global warming and climate change. In the U.S., the utilities sector alone produced 1,551 million metric tons of carbon dioxide (CO_2) emissions in 2021 (EIA). Electricity generation through renewable solar and wind can significantly reduce CO_2 emissions, yet land use and transmission infrastructure — both characterized as siting challenges — are an obstacle to the rapid acceleration and deployment of these technologies.

Solar and wind technologies can be co-located at a single location so that one transmission line can accommodate both technologies. How much energy is delivered from energy projects to energy demand centers depends on whether the technologies have complementary generation profiles and their relative sizing (ratio of solar to wind). The ratio of solar to wind is defined as the installed solar capacity to installed wind capacity; we report installed capacity in Megawatts (MW).

In terms of land use, solar panels can be strategically placed between turbines to minimize shading losses (Ludwig, et al., 2020). Using existing transmission infrastructure while increasing revenue, existing wind projects can retrofit solar PV technology and avoid the cost of new transmission development.

However, a limited knowledge around the quantity of solar PV and wind technology that can be cost-effectively co-located in different areas in the U.S. has stalled the consideration of co-location. This study aims to provide a better understanding of the optimal solar PV to wind ratio at existing wind projects that maximizes annual profit. In addition, the study conducts techno-economic analysis to rank co-location in existing wind projects within the context of co-location based on social, environmental, and economic factors. This can help advance co-location as a clean energy strategy that provides a better output per unit of land and efficient use of existing transmission infrastructure.

Here, the study presents an analysis that quantifies the optimal ratio of solar PV to wind when retrofitting existing wind projects in the contiguous U.S. using estimates of hourly wind and solar generation profiles. Using hourly meteorological wind speed and solar radiation data from National Renewable Energy Laboratory (NREL), the study models hourly energy generation and divides energy generation in each hour by the total system size to calculate hourly capacity factors for both wind and solar through the NREL System Advisor Model. These estimated capacity factors were generated for all existing wind project sites. The estimated wind and solar capacity factors for each existing wind project over three years was solved by an nonlinear optimization model with the objective to maximize annual profit subject to potential generation, actual generation, annual costs, and transmission line capacity at a given hour. The objective function solves for an optimal solar PV to wind ratio for each existing wind project.

This study identifies that, on average, for every 1.00 unit of wind, 1.01 units of solar may be added to the project to maximize annual profit for all viable existing wind projects, ceteris paribus. Overall this may lead to 115 Gigawatts (GW) of additional solar PV production within the U.S. utilizing existing transmission infrastructure and an improved output of energy per unit of land. In addition, around 12 percent of projects are within counties with IRA-eligible census tracts and less than 3 percent of projects were sited within counties with census tracts of Disadvantage Community status. These findings are illustrated further in an interactive map application.

With these results, this study hopes to spark interest in the potential for the co-location of solar PV and wind technology within the U.S. This is a preliminary analysis evaluating the complementarity of solar PV and wind technology in the case of retrofitting existing wind projects; additional research is necessary to understand project-specific challenges and incentives.

3.0 Problem Statement

Co-location projects of solar PV and wind technology have been successfully developed in India and China; however, co-location projects have yet to emerge at a similar scale in the U.S. (Ludwig, et al., 2020). The limited knowledge around the quantity of solar PV and wind technology that may be cost-effectively co-located across different meteorological profiles is likely a barrier to considering co-location as a decarbonization solution in the U.S. Better understanding of the optimal ratio of solar PV to wind at existing wind projects can advance co-location as a clean energy strategy with better energy generation output per unit of land and efficient use of existing transmission infrastructure.

4.0 Specific Objectives

This study filled this knowledge gap through the objectives below:

1) Determine Optimal Co-Location Ratio: Build a nonlinear optimization model to solve for the optimal solar PV to wind technology ratio at each existing wind project that maximizes the annual profit subject to the estimated generation, annual costs, and transmission line capacity of each viable project.

2) Conduct Multi-Criteria Techno-Economic Analysis: Rank existing wind projects on environmental sensitivity and estimated annual levelized-costs with co-location. Reported summary statistics for co-location in wind projects that are sited within Department of Energy disadvantaged communities and Inflation Reduction Act energy communities.

3) Create Interactive Map Application: Present model results and techno-economic analysis in an interactive map application for energy industry leaders and professionals to learn more about retrofitting for co-location.

5.0 Summary of Solution Design

The study's modeling and analysis can be summarized in four parts: existing wind projects data preprocessing, meteorological data acquisition, modeled energy capacity, nonlinear optimization computation, and multi-criteria techno-economic analysis.

The detailed abbreviations and definitions used in this study are listed in Table 1.

Table 1: Abbreviations and Definitions

Abbreviation	Definition
USWTDB	U.S. Wind Turbine Database
NREL	National Renewable Energy Laboratory
NSRDB	National Solar Radiance Database

WIND Toolkit	Wind Integration National Dataset Toolkit
SAM	System Advisor Model
HIFLD	Homeland Infrastructure Foundation-Level Data
GEA	Cambium Generation and Emission Assessment
IRA	Inflation Reduction Act
DOE	Department of Energy
DAC	Disadvantaged Community
RCI	Rural Capacity Index

5.1 Existing Wind Projects Data Preprocessing

This study focuses on existing wind projects as locations where co-location through retrofitting with solar PV may provide additional energy generation and profit without new transmission development. The USWTDB is a publicly available and spatially-validated database for the point coordinates and attributes of the individual utility-scale wind turbines sited across the United States (Rand et al., 2020). This study acquired a version of this data product from the Lawrence Berkeley National Laboratory (Berkeley Lab) with their added permitted project identifiers for each individual wind turbine. Permitted project identifiers were used to aggregate individual wind turbines to create convex hull polygons representing existing wind projects. The point centroids were then determined for each of the existing wind project polygons and these centroid point locations were used for the study's optimization and analysis. Existing wind project parameters included the number of wind turbines and the project capacity. Nearest neighbor analysis was also conducted to find the distance from each existing wind project to the nearest electric substation on the Homeland Infrastructure Foundation-Level Data.

This study notes a data validation issue in the aggregation of individual wind turbines by the permitted project identifiers. The Berkeley Lab assigns identifiers by the unique permits issued for wind turbine deployment and not by unique wind projects, resulting in several identifiers for a wind project that deployed wind turbines in several phases.

5.2 Hourly Meteorological Data Acquisition

This study used data from the NREL NSRDB and WIND Toolkit as inputs for energy generation modeling and to generate the energy capacity factors by technology. This study acquired solar and wind hourly meteorological profiles for the years 2012 to 2014. These profiles were downloaded through an API from NREL NSRDB and NREL WIND Toolkit for each coordinate corresponding to the centroid of the polygons for existing wind projects in the contiguous U.S. from the USWTDB (See Section 5.1).

5.3 Modeled Energy Generation

The hourly meteorological data described above served as the inputs to energy generation simulation runs using NREL's System Advisory Model (SAM) Python-API, resulting in an hourly time series of energy output in kilowatt hours (kWh) for wind and solar PV technology at the existing wind projects. Additional information about SAM is outlined in Section 8.1.

5.4 Nonlinear Optimization Computation

A nonlinear optimization model was constructed to find the optimal solar PV to wind ratio by satisfying the objective equation at each existing wind project that maximizes the annual profit of co-location, subject to potential generation, actual generation, annual costs, and transmission line capacity at a given hour. The open-source Python library, Pyomo, was used to define and initialize the model and CPLEX was used to solve the optimization model. Additional information about the Pyomo module and CPLEX solver is outlined in Section 8.1.

This model solves by maximizing total annual profit using price estimates of the potential revenue from electricity generation and the cost of transmission capacity of each existing wind project. Wind and solar project costs are from the NREL Annual Technology Baseline and hourly wholesale electricity prices are from the NREL Cambium 2022 Project datasets. The NREL Annual Technology Baseline uses robust modeling assumptions to identify technology-specific cost and performance parameters. The NREL Cambium 2022 Project provides hourly emission, cost, and operational estimates for several forward-looking U.S. electricity scenarios for the contiguous U.S. through 2050. This study assumed a high electrification scenario in the year 2030 and accounted for potential cross-site differences in energy markets by matching each wind project to the GEA region it belongs to. GEA regions cover the contiguous U.S. based on the eGRID regions of the U.S. EPA but not identical due to the geographic extent and resolution of the Cambium workflow (Gagnon et al., 2021).

5.5 Multi-Criteria Techno-Economic Analysis

This study also disaggregated and evaluated the ratios with key criteria for the social and environmental implications of the energy transition. This multi-criteria techno-economic analysis considered three main criteria: **(1)** environmental impact, **(2)** economic development, and **(3)** equitable investment. This study acquired open-source data products for each criteria as chosen by the Client.

This study used environmental impact scores from the *Power of Place* study by The Nature Conservancy to evaluate the environmental sensitivity of solar PV deployment within existing wind projects. The scores are high-level estimates of spatially-specific environmental positives and negatives, where larger raster values indicate higher negative impacts, lower raster values indicate lower negative impacts. The environmental impact index consists of data falling into the following categories: wetland, managed areas, threatened and endangered species habitat, intact habitat, and local bird and bat habitats.

This study evaluated economic development to understand the potential of co-location within areas that may be eligible for a 10 percent increase in federal tax benefits for renewable energy development using Resources for the Future data for the census tract-level categorization of IRA energy communities. This study coupled this analysis with Headwaters Economics' county-level RCI data. The RCI evaluates for the staffing and resources needed to submit funding applications, satisfy rigorous reporting requirements, and plan, build, and sustain projects over the long term.

This study also evaluated equitable investment in renewable energy deployment using DOE census tract-level data on disadvantaged indices and DAC status. The DOE uses fossil dependence, energy burden, environmental and climate hazard, and sociodemographic data at the census tract-level to assess for DAC status. Spatial aggregation of the data products was conducted to contextualize the energy injustice landscape of existing wind projects. All data products were adjusted to have a consistent extent, coordinate reference system, resolution, and projection. Metadata that outlines different attributes and assumptions of these data are available in the metadata.xml file.

6.0 Products & Deliverables

The study delivered four main products in addition to study findings: **(1)** a version-controlled repository, **(2)** data and metadata, **(3)** a methods summary, all for scientific reproducibility of the modeling and analysis approach; and **(4)** an interactive map application for industry communication of the model results. These four products fall within the two goals of the study's deliverables to maintain scientific reproducibility and provide a tool for industry communication.

6.1 <u>Scientific Reproducibility</u>

The version-controlled repository houses the Python and R scripts for extracting the meteorological data, generating time series technology capacity factors, building and solving an nonlinear optimization model, and conducting the multi-criteria techno-economic analysis. To execute the script and solve the optimization model, data and metadata for the model inputs are provided; likewise, data and metadata for the model outputs are provided, for reproducibility and transparency. A methods summary was written to document the optimization model building and selection for reproducibility in similar feasibility, implementation, and co-location siting research. The methods summary may be included in a publication by the Client for their renewable energy deployment research.

6.2 <u>Industry Communication</u>

The interactive map application presents the optimization model results and multi-criteria techno-economic analysis for industry leaders to explore and learn more about solar PV and wind technology co-location as a strategy to advance U.S. decarbonization goals. This application also presents map layers and summary statistics of the equitable investment landscape of existing wind projects and project eligibility for a 10 percent increase in federal tax benefits for renewable energy development as part of the IRA. This application is a part of the Client's public and industry outreach materials for their continued research on renewable energy deployment.

6.3 Study Findings

This study reports four important findings of retrofitting existing wind projects with solar PV in the contiguous U.S.: **(1)** additional potential solar energy, **(2)** combined annual revenue and additional profit, **(3)** the optimal ratio of solar PV to wind, and **(4)** multi-criteria techno-economic analysis. This study reports findings for the subsample of 1,288 existing wind projects viable to be retrofitted with solar PV and defines viability as existing wind projects that have an estimated solar capacity greater than zero.

This study finds an estimated 115 GW of potential solar energy across all viable existing wind projects. This addition of potential solar energy represents the opportunity to nearly double capacity at existing wind projects while avoiding the development of new transmission infrastructure. Retrofitting existing wind projects with solar PV is also estimated to deliver over 11 billion USD in additional profit from solar energy production and 30 billion USD in annual combined revenue with wind energy production.

In addition, this study finds an average optimal ratio of solar PV to wind of 1.01 for all viable existing wind projects. The optimal ratio suggests that, on average, for every 1.00 unit of wind, 1.01 units of solar may be added to the project to maximize profit, ceteris paribus. While the distribution of optimal ratios ranges from zero to 1.50, most existing wind projects have a ratio of solar PV to wind greater than 0.70. We thus conclude that co-location of solar PV and wind at viable existing wind projects may be a 'low-hanging fruit' pathway towards reaching clean electricity goals in the U.S.

This study couples these nonlinear optimization model findings with a multi-criteria techno-economic analysis for environmental impact of solar PV deployment, economic development, and equitable investment.

This study finds an average environmental impact score of 7.80 across all viable existing wind projects and a weighted average environmental impact score of 6.61 by potential solar capacity. These estimates suggest that retrofitting viable existing wind projects with solar PV has minimal direct environmental impact and may make use of environmental and physical sunk costs already incurred by projects, like fencing, auxiliary, roads infrastructure, and transmission infrastructure.

This study also finds that around 12 percent of viable existing wind projects are sited within counties with IRA energy communities and account for over 13,900 MW of potential solar capacity. This suggests that around 12 percent of viable existing wind projects may receive a 10 percent increase in federal tax benefits. However, the local governments of IRA-eligible viable existing wind projects also operate, on average, at near 76 percent capacity. The local governments of where these projects are sited may not have the staffing and resources to access these federal tax benefits. This study concludes that there is uncertainty about how and where retrofitting viable projects with solar PV may provide opportunities for economic development in the energy transition.

Similarly, this study finds that less than 3 percent of viable existing wind projects are within counties with census tracts of DAC status. This percentage of existing wind projects is low likely due to systemic biases and rural-urban and racial disparities that influenced the siting of existing wind projects (Ross et al., 2022).

Largely, retrofitting existing wind projects with solar PV has the potential to help place the U.S. closer to reaching its ambitious goal of a net-zero U.S. power grid by 2035. Deploying the potential of 115 GW of solar PV at viable existing wind projects can also increase energy production per square kilometer of land and utilize existing infrastructure in effect reducing the need for new transmission development.

7.0 Summary of Testing

7.1 Testing of Code Script

Quality assurance and quality control were conducted as data were acquired to validate that data were in their correct spatial and temporal resolution and contained all attributes of interest. A temporal resolution of three years was chosen to address the variability of the meteorological data and sensitivity considerations. Data types were tested using error messages throughout the script with the *py.test*; these error messages helped ensure reproducibility and confirmed that future users use the correct data types. Team members also implemented random unit testing through the different stages to ensure the data were joined correctly into a consolidated dataset.

7.2 Testing Data Acquisition & Optimization

A subset of 10 sites from the identified 1,332 existing wind projects in the contiguous U.S. was used to estimate the render time for data acquisition of hourly meteorological data over three years and troubleshoot API errors. Similarly, the subset was used to estimate the computational time of solving the optimization model with CPLEX.

7.3 Testing Interactive Map Application

In addition to one-on-one user tests with energy industry stakeholders, the interactive map application was 'monkey tested' to test if the application crashed and ensure that it had consistent libraries, features, and dependencies across operating systems. See Appendix 1 for Final Evaluation Testing Plan.

8.0 User Documentation

8.1 Software & Tools

This study uses Visual Studio Code for all scripts in Python and RStudio for all scripts in R. Generally, work was performed within local GitHub repositories from a Bren PC via a Azure

Virtual Machine (VM) and the servers Taylor and Tsosie were accessed for data storage via CyberDuck. A .yml within the solar-wind-ratios GitHub repository also outlines the required Python packages, libraries, and dependencies to set up an environment.

8.1.1 *Modeled Energy Generation*

The SAM application was installed and ran through a VM with the pySCC API. SAM uses tools to automate simulations for batch processing of parametric and sensitivity analysis and weather variability to model many different types of renewable energy systems and financial models for residential, commercial and grid-scale projects. Since the launch of the public tool in 2007, there have been over 35,000 downloads of the tool (NREL). In addition to research, the tool is used by: manufacturers to evaluate efficiency improvements and cost reductions; energy project developers to examine system configurations for maximizing energy output; and policy makers to identify impacts of incentive structures (NREL). SAM was used to model energy capacity for solar PV and wind technology. The default variables were conserved.

8.1.1.1 **Instructions to Access the NREL API**

- Go to https://developer.nrel.gov/ and click on *Get an API Key* under *API Key*
- Enter API signup information like name, email, and institution or organization
- Upon entering the information, you will receive an email with the API key
- Save and secure this key for future use within the repository

8.1.1.2 **Instructions to Download SAM**

- SAM is a free application that can be downloaded online
- Follow the instructions to download to your local computer: https://sam.nrel.gov/download.html

8.1.2 *Nonlinear Optimization Computation*

Pyomo (version 6.4.4) is used to define and initialize the model and CPLEX solver is used to solve an nonlinear optimization model. Pyomo is a Python-based, open-source optimization modeling language for linear, quadratic, nonlinear, mixed-integer, and stochastic mathematical programming. CPLEX is a high-performance mathematical programming solver for linear, mixed-integer, and quadratic programming by IBM. A free version of the CPLEX solver through an academic registration account was used. Pyomo library's latest release (version 6.4.4) was installed through an Anaconda environment on the VM to ensure consistency and reproducibility. Pyomo and CPLEX streamlines the iterative process of computing to find the optimal solar PV to wind ratio at each existing wind project.

8.1.3 *Industry Communication*

The interactive map application was built using Shiny in R. Shiny is an open-source and free R package by Posit for developing and deploying interactive web applications. This application is maintained and hosted on the Shiny server under the Client's Shiny account.

8.2 Repository Access

The wattmaps GitHub organization houses the public repositories that hold all the code scripts required to reproduce this study. The solar-wind-ratios repository stores code scripts used in the acquisition of hourly meteorological data, SAM modeled energy generation, and used to build and solve the optimization model. A separate repository wind-retrofit-app stores the code scripts to deploy and host the Shiny application.

8.3 Repository Structure

The solar-wind-ratios repository contains folders for each technology (i.e., solar and wind). These folders store code scripts for the iterative acquisition of hourly meteorological data and modeled energy generation for each technology at each existing wind project. A separate folder for code scripts associated with building and solving the optimization model exists within the repository. Also housed within the repository is a folder that holds all of the scripts for data cleaning and preprocessing required for the modeling and optimization. The repository also contains a README.md file and .gitignore file. The .gitignore folder specified data file types to avoid pushing data to the repository. See Appendix 2 for a complete file structure.

8.4 Analysis Execution

Included below provides helpful information needed to run the scripts used in this analysis.

8.4.1 *Wind Speed & Solar Radiance Meteorological Data Acquisition*

- Use the NSRDBapi_SAMsim_solarPV.py file to acquire solar radiance data and the WTKapi_SAMsim_wind_hourly.py file to acquire the wind speed data. Both files contain the API code script with blank user inputs.
- Input user API key and email, year of interest, and resource-specific parameters (like such as wind turbine hub height) and determine what the Universal Time Coordinate (UTC) should be used for the meteorological data.
- Read in a file with latitudes and longitudes of points of interest for the wind and solar meteorological data acquisition.
- The NREL WIND Toolkit server produced random errors when downloading the data. When this error occurs, the time series data will not populate within the csv file and the first row of the csv will be populated with the API query inputs. For this study, we manually fixed this error by deleting the file that the data didn't store properly and ran the script starting from the project ID that caused the error. Future runs of this model energy generation could automate this manual fix by including an additional if statement within the code.

8.4.2 *Model Energy Generation*

- The step uses the same files used for downloading the data as well as an additional SAM assumption file.
- To get specific SAM assumptions, open the SAM application on your local computer

- On the SAM application, navigate to *File* and upload a .sam file. The .sam file contains all of the user-defined assumptions for wind. Alternatively, you can set your settings within the SAM tool.
- In the main panel, click the down arrow and select the option *Generate Code*
- Select Python for the language and input this script into the *SAM_assumptions_hourly_wind.py for wind and SAMassumptions_hourly_solar.py file for solar* where the code script notes "### SAM GENERATED CODE ###"
- Now when you run the *API scripts* it will generate power profiles for the given latitudes and longitudes based on your assumptions. It takes about 3-4 seconds to run this for each coordinate pair.
- Refer to Appendix 3 and Appendix 4 for solar PV- and wind-specific assumptions.

8.4.3 Solve for Optimal Ratios

- The colocation_opt_loop.py script pulls in the following data stored on your local computer:

Table 2: Objects and Descriptions

Object Name	**Type**	**Description**
substation_pids.csv	csv	Contains the distance to the nearest substation for each PID and is used to define transmission and substation costs
US_pids_GEAs.csv	csv	Contains the GEA that each PID corresponds to and is used to match up each PID with a GEA
cambiumHourlyPrice2030	Folder	Contains projected hourly price data for each GEA (each file in the folder represents pricing for a specific GEA) and is used to define hourly wholesale electricity prices
potentialInstalledCapacity	Folder	Contains a file for potential solar installed capacity and a file for wind installed capacity and is used to define potential capacity
solar_capacityFactor_filePerPID	folder	Contains hourly capacity factors for solar at each site over three years and is used to define solar hourly energy generation, where each file should contain three years of data for a PID
wind_capacityFactor_filePerPID	folder	Contains hourly capacity factors for wind at each site over three years and is used to define wind hourly energy generation, where each file should contain three years of data for a PID

- See the data repository within Zenodo for an example of the data structure and the necessary data for the constraints and parameters.
- Install the CPLEX solver from IMB's download portal and direct the CPLEX environment path to your local computer's path to the CPLEX solver files.
- Scalar values and parameters may be changed by updating values within the script and reading in new data for certain parameters.
- This file will loop through each of the different project locations and determine an optimal ratio that maximizes annual profit. The final output is a dataframe with the ratio results and annual profit for each existing wind project location.

Note: Depending on your data inputs, additional preprocessing and wrangling may be required to execute these scripts. See the dataCleaningPreprocessing folder within the repository to account for missing data from the energy generation modeling and other data preprocessing and wrangling steps to execute this analysis.

9.0 Archive Access

9.1 Research Data

Research data for this study came from publicly available and open-source datasets. Table 3 outlines the datasets used in the study.

Table 3: Data Sources

Data	File Type	Source	Description of data and use
United States Wind Turbine Database (USWTDB)	csv	Provided by Client from Lawrence Berkeley National Laboratory	Point data of existing wind turbines in the contiguous U.S; data as of 2021. Used to identify and assign existing wind project IDs (PID).
National Solar Radiation Database (NSRDB)	csv	National Renewable Energy Laboratory; NSRDB	Hourly values of the three most common measurements of solar radiation (global horizontal, direct normal, and diffuse horizontal irradiance) and meteorological data; data from 2012-2014. Used to generate solar meteorological profiles for each PID.
Wind Integration National Dataset (WIND) Toolkit	csv	National Renewable Energy Laboratory; WIND Toolkit	Meteorological data at multiple hub heights for more than 2,488,136 sites in the continental United States for the years 2007–2014; data from 2012-2014. Used to generate wind meteorological profiles for each PID.
Electric Substations	shapefile	Provided by Client, Homeland Infrastructure Foundation-Level Data (HIFLD)	Locations of electric power substations primarily associated with electric power transmission; data as of 2022. Used to find the nearest electric substation and substation characteristics (like voltage) to each PID.

Cambium 20222 Project Data	csv	National Renewable Energy Laboratory	Modeled hourly emission, cost, and operational data for a range of electrification scenarios of the U.S. through 2050; data as of 2022. Used to find the projected price scenarios of each PID at the region-level.
Environmental Sensitivity Index	raster	Power of Place	Environmental impact score dataset for solar PV from Power of Place - National study; data as of 2022. Used to assign an environmental impact score to each PID.
Inflation Reduction Act (IRA) Energy Communities Identifier	shapefile	Resources for the Future (REF)	Locations of IRA energy communities as defined by the interpretation of eligibility from RFF; data as of 2022. Used to classify eligibility for tax incentives by the IRA of each PID at the census tract-level.
Disadvantaged Communities Identifier	shapefile	Department of Energy	Identifies disadvantaged communities based on 36 burden indicators like fossil dependence, energy burden, environmental and climate risk, and socio-economic vulnerabilities at the census tract-level; data as of 2022. Used as part of equitable investment analysis of each PID at the county-level.
Rural Capacity Index	shapefile	Headwaters Economics	Identifies communities with limited capacity needed to apply for federal programs on a scale of 0 (low capacity) to 100 (high capacity); data as of 2022. Used as part of IRA analysis of each PID at the county-level.

9.2 Data Sharing & Access

As required by the Creative Commons licenses provided by the data sources, the publicly accessible data is cited under references in the final Technical Documentation Plan.

9.3 Data Archival & Preservation

The servers Taylor and Tsosie housed within the Bren School of Environmental Science & Management, were used to store and archive data. In addition, all code scripts were version controlled and archived using git via GitHub repositories. The study's data, metadata, and scripts were provided to the Client for their archives and continued research.

References

About — System Advisor Model. (n.d.). National Renewable Energy Laboratory. Retrieved June 1, 2023, from https://sam.nrel.gov/about-sam.html

Frequently Asked Questions (FAQs) – Renewable. (n.d.). U.S. Energy Information Administration. Retrieved June 1, 2023, from https://www.eia.gov/tools/faqs/

Gagnon, P., Frazier, W., Cole, W., & Hale, E. (2021). *Cambium Documentation: Version 2021*. National Renewable Energy Laboratory. https://doi.org/10.2172/1835075

Ludwig, D., Breyer, C., Solomon, A. A., Seguin, R., Ludwig, D., Breyer, C., Solomon, A. A., & Seguin, R. (2020). Evaluation of an onsite integrated hybrid PV-Wind power plant. *AIMS Energy*, *8*(5), Article energy-08-05-988. https://doi.org/10.3934/energy.2020.5.988

Rand, J. T., Kramer, L. A., Garrity, C. P., Hoen, B. D., Diffendorfer, J. E., Hunt, H. E., & Spears, M. (2020). A continuously updated, geospatially rectified database of utility-scale wind turbines in the United States. *Scientific Data*, *7*(1), Article 1. https://doi.org/10.1038/s41597-020-0353-6

Ross, E., Day, M., Ivanova, C., McLeod, A., & Lockshin, J. (2022). Intersections of disadvantaged communities and renewable energy potential: Data set and analysis to inform equitable investment prioritization in the United States. *Renewable Energy Focus*, *41*, 1–14. https://doi.org/10.1016/j.ref.2022.02.002

Vasiliki Klonari, Daniel Fraile, Raffaele Rossi, & Michael Schmela. (2019, May 23). *Exploring the Viability of Hybrid Wind-Solar Power Plants*. 4th International Hybrid Power Systems Workshop, Crete, Greece.

Appendix 1: Interactive Map Application Testing Script
Think-Aloud Tasks for Representative Users
Please 'think-aloud' as you perform these tasks by just saying everything that comes to mind as you are making choices and navigating through the interface. I will be happy to repeat the task at any time if you forget any part of it.

Industry User Tasks

- Imagine you have been sent this dashboard by a coworker who researches the latest technologies within the renewable energy sector. Your focus is looking at how federal and regional incentive structures can aid your organization in developing new renewable energy projects. While thinking-aloud, **find the suitable colocation sites that fall within the federal IRA incentive areas as defined as "energy communities"**, within the map dashboard.
- Consider that you attended a Lunch & Learn about co-location of wind and solar resources and the presenter mentioned this dashboard that identified suitable areas for co-location within the U.S. You focus on the Western Interconnected Grid within your work. While 'thinking aloud', **filter the map to present co-location sites specific to the Western Interconnected Grid**. Secondly, **download the raw data** to your local computer for your own analysis.
- Suppose you are meeting with the finance team about reducing the cost of transmission infrastructure and curtailment losses. With your research, you stumble upon the dashboard highlighting the benefits of co-location of wind and solar. While 'thinking aloud', **navigate to where you can find the cost estimates of co-location in the Eastern Interconnection Grid**.

Research User Tasks

- Suppose you follow a link from a policy brief to a dashboard that explores co-location of wind and solar in the state of Texas in the U.S. You have just spoken to a regional advocacy group looking for new strategies to deploy renewable energy in Texas. You are interested in learning about the number of suitable sites and the optimal ratio of wind to solar at each suitable site in Texas. While 'thinking aloud', **filter the data to represent co-location sites specific to Texas.**
- Imagine you are preparing for a presentation to the general public on key strategies to meet our decarbonization goals in the electricity sector in the West and stumble upon this dashboard exploring co-location. While 'thinking aloud', **find the wind-only or solar-only estimates for a site.**
- Suppose you are researching wind and solar complementarity and come across this dashboard exploring co-location. You are interested in how the ratio of wind to solar technology at each site was calculated. While `thinking aloud`, **navigate to where you can find links to the project's GitHub repository and data citations**.
- Imagine you are interested in adding facts and figures for several co-location sites from the dashboard. Before you add this information to your own paper, you need to

state the assumptions. While ‘thinking aloud’, **navigate to where you can find the assumptions made for the optimal ratio of wind to solar technology.**

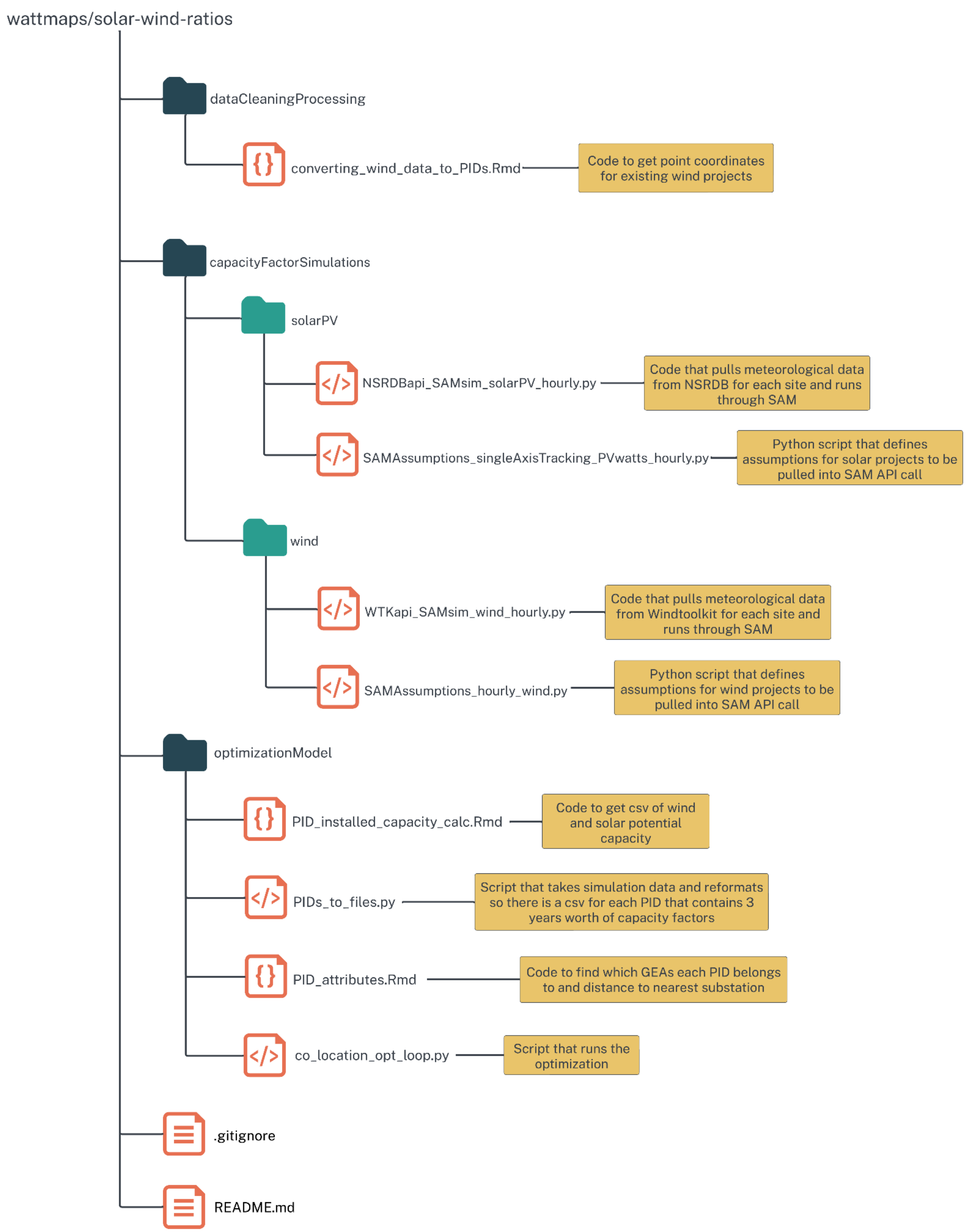
Example Questions about Tasks and Experience Completing Tasks

The section outlines questions that will elicit evaluator impressions of the tasks provided:

- How realistic were the tasks that were given to you?
- Did the tasks align with what you would typically look for in a dashboard when exploring a new renewable energy deployment strategy? **If these tasks didn’t align with what you would typically do**, what is a more realistic task you would perform when exploring a dashboard on co-location?
- What additional tasks would you like to be able to complete?
- What task was most difficult or challenging based on the design?
- Were there any moments during the tasks that felt confusing or unclear?

Appendix 2: solar-wind-ratios Repository File Structure

The below diagram illustrates the organization of code in the solar-wind-ratios repository. A stable version of the structure with suggested data object organization can be found here.



Appendix 3: SAM Assumptions for Solar PV

Table 4: SAM Assumptions for Solar PV

Parameter	Unit	Value
Array type	Factor	1-axis tracking
Tilt	Degree	0
Azimuth	Degree	180
Ground coverage	Ratio	0.3
Ground coverage (albedo)	Ratio	0.2
DC to AC	Ratio	1.34
Inverter efficiency	Percent	96
Soiling loss	Percent	2
Shading loss	Percent	3
Mismatch loss	Percent	2
Wiring loss	Percent	2
Connection loss	Percent	2
Light-induced degradation loss	Percent	1.5
Nameplate loss	Percent	1
Age loss	Percent	0
Availability loss	Percent	1

Appendix 4: SAM Assumptions for Wind

Table 5: SAM Assumptions for Wind

Parameter	Unit	Value
Wind speed variations	Factor	Weibull distribution
Wind turbine height	Meter	100
Maximum coefficient of power for wind turbine	NA	0.45
Shear coefficient	NA	0.40
A simple wake model	Factor	0
Turbulence coefficient	Ratio	0.10
System capacity	Kilowatt	90,000
Internal wake loss	Percent	0%
External wake loss	Percent	1.10%
Future wake loss	Percent	0%
Turbine loss	Percent	1.91%
Balance of plant	Percent	0.50%
Grid	Percent	1.50%
Sub-optimal performance	Percent	1.10%
Generic power curve adjustment	Percent	1.70%
Site-specific power curve adjustment	Percent	0.81%
High wind hysteresis	Percent	0.40%
Icing loss	Percent	0.21%
Environmental loss	Percent	0.40%
Degradation	Percent	1.80%
Exposure change	Percent	0%
Load curtailment	Percent	0.99%
Environmental and permit curtailment	Percent	1.00%
Grid curtailment	Percent	0.84%
Operational strategies	Percent	0%
Total uncertainty	Percent	12.085%

Grid interconnection limit	kW_AC	100,000