

Datasets in repository

Use limitations and disclaimer:

Candidate project areas (CPAs), their spur lines (gen-ties), and their estimated costs were generated using published studies (see paper) and expert opinion on renewable energy siting and transmission routing criteria. Potential interzonal transmission lines were modeled using assumptions based on expert opinion among those on the author team regarding which lines are most likely candidates for reconductoring and colocating and their potential capacity, as well as the most likely interconnection points for new lines. Modeling assumptions have not been vetted or endorsed by any planning authorities. These data should only be treated as schematic representations of possible areas for renewable infrastructure development and transmission availability and only for scenario analysis if directly used. Use of these datasets for planning infrastructure projects without review and modifications through a planning process is NOT recommended.

Spatial data (supply curve):

- Environmental exclusions (“PoPWest_EnvExclusions_Rasters.gdb.zip”):
 - File geodatabase contains the three environmental exclusion categories (binary rasters) that were aggregated to create the environmental Siting Levels
- Candidate Project Areas (CPAs) for offshore wind, onshore wind and utility-scale solar PV (“CPA.zip”): one shapefile per technology and siting level
 - Area_km2: Area in km2
 - CPA_ID: unique identifier for CPAs. Use as foreign key for joining spur lines using the same column name.
 - Cap_MW: capacity in MW
 - length_m: spur line length in km (to the nearest 230 kV substation)
 - txTot_usd: total transmission costs for the spur line in USD 2018 including substation costs
 - CPA_zone: zone (state or region) for RIO modeling
 - annGen_MWh: annual average generation in MWh
 - rf_sc: random forest prediction score indicating the likelihood that a wind or solar power plant would be sited in that located. Higher values indicate higher likelihood whereas lower values indicate lower likelihood (used in downscaling).
 - Area (offshore wind only): Area in km2
 - depth (offshore wind only): ocean sea floor greater or less than -50m below sea level
 - CF_mean: area weighted average capacity factor
 - type (offshore wind only): fixed or floating turbines
 - length_km (offshore wind only): spur line length in km (to the nearest 230 kV substation)

- txUsdPerkW (offshore wind only): total transmission costs in USD 2018 per kW, including substation costs
- BOEM_Area (offshore wind only): whether or not the CPA is located in an BOEM designated lease area
- Power lines - Spur lines corresponding to the CPA ("PowerLines_spurs.zip"): one shapefile per technology and siting level, except for solar PV which is divided into north (N) and south (S).
 - CPA_ID: unique identifier for CPAs. Used for associating the CPAs with the spurs.
 - PathCost: cost of the spur line alone in usd
 - SpurlID: ignore
 - Length: length of the spur line in meters
- Power lines - interzonal transmission lines ("PowerLines_transmission.zip"):
 - Type: colocate, new, reconductor (and voltages)
 - Corridor: state → state
 - Usd_mile: usd 2018 per mile (including substation costs)
 - Cap_MW: maximum new capacity on the line
 - Shape_leng: length of line in meters

RIO capacity expansion modeling results:

- Excel workbook with pivot table of capacity (GW) selected per scenario (columns) per state (rows) per year (filter) per technology (filter) ("RIO_capacity.xlsx")
- Excel workbook with pivot table of generation per scenario (columns) per state (rows) per year (filter) per technology (filter) ("RIO_annual_energy.xlsx")
- Excel workbook with pivot table of interzonal transmission capacity selected per scenario (rows), per corridor (filter), per year (filter) ("RIO_total_transmission_capacity.xlsx")

Environmental and social impact assessment results:

- Excel pivot table of environmental impacts ("EnvlImpacts_allCategorical.xlsx"):
 - See spreadsheet for description and how to use pivot table
- Excel spreadsheet of social impacts ("SocImpacts_allContinuous.csv")
 - Infrastructure: selSites (candidate project areas), bulk (interzonal transmission), spurs (gen-ties that connect CPAs to nearest 230 kV substation)
 - Method: downscaling method used. Choose from: rf (random forest), logReg (logistic regression), LCOE (levelized cost of electricity). Paper figures show random forest results.
 - Tech: solar, wind, bulk (interzonal transmission)
 - SitingLevel: environmental siting level (1-3)
 - metric: metric assessed—perPoverty (percent living below poverty line), income (average individual income), perUnemp (percent unemployed), and popDen (population density in persons/km²). See Social Vulnerability Index (2018) released by the CDC/ATSDR documentation for more details on the perPoverty, income, and perUnemp metrics.

- Scenario: he or high_elect = high electrification; se = slow electrification; heRE = high electrification with 100 renewables; heO = electrification only (assuming high electrification rates); IS = in state; lb or low_bio= low biomass; heO = electricity only; refEO = reference electricity only; RC or reConst = renewables constrained; numbers following UNDERSCORES correspond to the siting level; abbreviated names are used with CPAs (selSites) and long names are used with bulk tx lines.
- region: all or individual states
- mean: mean score value (units correspond to envMetric)