An autopilot for energy models – automatic generation of renewable supply curves, hourly capacity factors and hourly synthetic electricity demand for arbitrary world regions

Supplementary material

Niclas Mattsson, Vilhelm Verendel, Fredrik Hedenus and Lina Reichenberg

Part 1: GIS procedure

Here we describe the GIS calculations that take place in our package "GlobalEnergyGIS.jl" in more detail. The package is written in Julia and is available at ("Github (GlobalEnergyGIS)," n.d.).

ERA5 datasets

Given a target year (2018 by default) we download datasets for four ERA5 reanalysis variables from Copernicus (2017), the ECMWF online data service:

- wind: 100m u- and v-components of wind speed
- solar: "surface solar radiation downwards" (SSRD) and "total sky direct solar radiation at surface" (FDIR)

We request these variables at a resolution of 0.28125 degrees (about 31 km at the equator, the maximum resolution for ERA5) producing a global raster of 1280x640 pixels with a time dimension of 8760 hours. This raw data comprises 28 GB per variable per year, but disk space usage can be significantly reduced once initial preprocessing has taken place.

Preprocessing steps

Wind preprocessing

For wind power, we calculate hourly absolute wind speeds from the u- and v-components, discarding the directional information. To further reduce disk usage, we set wind speeds at far-offshore locations to zero, which allows for efficient data compression. Since annual wind speeds are only available for near-coast locations in the Global Wind Atlas (GWA) (DTU, 2019), we use the GWA to make

this determination. This last step is optional, so users who wish to consider floating offshore wind power can retain wind speeds over oceans at the cost of disk space. The resulting global wind speed dataset is stored in a HDF5-file with compression, taking up 6.5 GB of disk space for each year of data.

Solar preprocessing

The preprocessing step for solar variables is more complex. This is because ERA5 solar variables refer to insolation on a horizontal plane. These must be transformed appropriately for different solar technologies, taking into account the angle of incidence of the insolation. This in turn requires calculating solar position in a horizontal (local) coordinate system for each raster pixel and hour.

Solar positions are calculated using the fastest algorithm (number 1) of Grena (2012). First declination and hour angle of the sun, i.e. solar position in an equatorial (absolute) coordinate system, are calculated for each hour (in universal time UTC) and longitude. Then solar zenith and azimuth angles are calculated for each latitude.

In addition to the two ERA5 variables for diffuse and direct insolation, we also need top-of-atmosphere solar insolation (TOA) variations over the year. This variable is also available in the ERA5 database, but instead of downloading additional data we choose to calculate it using this simple relation ("Wikipedia," 2019):

$$\mathrm{TOA} = I_0(1 + 0.034 \cos \frac{2\pi n}{365.25})$$

where I_0 is the solar constant (1361 W/m^2) and n is the ordinal of the day of the year.

Solar PV

The total insolation striking a tilted solar PV panel is the Global Tilted Irradiance (GTI):

$$\mathrm{GTI} = I^{\mathrm{sun}}_{\mathrm{direct}} + I^{\mathrm{sky}}_{\mathrm{diffuse}} + I^{\mathrm{ground}}_{\mathrm{diffuse}}$$

where $I^{\mathrm{sun}}_{\mathrm{direct}}$ is direct beam radiation from the sun, $I^{\mathrm{sky}}_{\mathrm{diffuse}}$ is diffuse radiation from the sky, and $I^{\mathrm{ground}}_{\mathrm{diffuse}}$ is diffuse reflected radiation from the ground. $I^{\mathrm{sun}}_{\mathrm{direct}}$ can be directly calculated from the ERA5 FDIR variable using the solar position. $I^{\mathrm{ground}}_{\mathrm{diffuse}}$ is also straightforward assuming a constant uniform ground albedo. To estimate $I^{\mathrm{sky}}_{\mathrm{diffuse}}$ there are several potentially useful models; see Loutzenhiser et al. (2007) for an overview. We use the Hay-Davies model which includes an isotropic component and circumsolar diffuse radiation to take into account that the sky is brighter nearer to the sun. The resulting equations are:

$$I^{\text{sun}}_{\text{direct}} = \text{FDIR} \cdot R_b = \text{FDIR} \cdot \frac{\cos \text{AOI}}{\cos z} = \text{DNI} \cdot \cos \text{AOI}$$
$$I^{\text{sky}}_{\text{diffuse}} = \text{DHI} \cdot \text{AI} \cdot R_b + \text{DHI} \cdot (1 - \text{AI}) \cdot \frac{1 + \cos \beta}{2}$$
$$I^{\text{ground}}_{\text{diffuse}} = \text{GHI} \cdot \rho \cdot \frac{1 - \cos \beta}{2}$$

where FDIR is direct solar insolation on a horizontal surface (ERA5 variable), R_b is the ratio of tilted and horizontal solar beam irradiance, AOI is the angle of incidence of the sun on the PV panel, z is the solar zenith angle, DNI is direct normal irradiance, DHI is diffuse horizontal irradiance, AI is the anisotropic index (a measure of nonuniformity of sky brightness), β is the tilt angle of the PV panel and ρ is ground albedo, which is assumed to be 0.2 everywhere. The variables are further related by:

$$\text{DHI} = \text{SSRD}, \quad \text{DNI} = \frac{\text{FDIR}}{\cos z}, \quad \text{AI} = \frac{\text{DNI}}{\text{TOA}}, \quad R_b = \frac{\cos \text{AOI}}{\cos z}$$

$$\cos \text{AOI} = \cos z \cos \beta + \sin z \sin \beta \cos(\alpha_{\text{sun}} - \alpha_{\text{pv}})$$

Here SSRD is the ERA5 variable "surface solar radiation downwards", α_{sun} is the azimuth angle of the sun and α_{pv} is the azimuth angle of the PV panel (assumed zero), with azimuth measured with zero due south and positive direction toward west. ERA5 radiation variables are documented in Hogan (2015).

In clear-sky weather, the optimal tilt angle of a PV module for a given location is the latitude of the panel. However, if conditions are often cloudy, more diffuse sky radiation can be captured if the tilt angle is smaller than its latitude. Therefore the optimal tilt angle is location specific. For simplicity though, we use the fitted third degree polynomials from Jacobson & Jadhav (2018) to get near-optimal tilt as a function of latitude.

CSP

For CSP, the insolation on a 2-axis solar tower collector is the Direct Normal Irradiance (DNI), as calculated above. The insolation on a north-south oriented parabolic trough is given directly by ERA5 data as FDIR. (In the current version of our energy system model, we assume all CSP is of type solar tower.)

Saved solar variables

To save disk space, insolation over oceans is discarded (set to zero) using the land cover dataset. The transformed solar variables representing insolation of solar PV and CSP of type solar tower, i.e. GTI

and DNI (and optionally FDIR for parabolic troughs), are stored in HDF-files with compression. The resulting file size is approximately 7.5 GB per solar variable per year.

Solar and wind synchronization

ERA5 insolation and wind speed variables are consistent with each other since they are calculated from the same reanalysis model. However, wind speeds in ERA5 are instantaneous, but insolation variables measure radiation accumulated in the previous hour. Therefore the solar dataset is on average 30 minutes out of sync with the wind data. This bias could be corrected by interpolating the datasets, but since this would smooth the data somewhat (i.e. reduce temporal variations) we have chosen not to do so.

Preprocessing of other datasets

Several auxiliary datasets are used in the GIS calculations in addition to the reanalysis data. These are: administrative borders ("Eurostat NUTS," n.d.; GADM, n.d.), gridded population (Gao, 2017) and GDP (Murakami & Yamagata, 2019) in SSP scenarios (Riahi et al., 2017), land cover (Friedl et al., 2010), topography (Amante & Eakins, 2009) and protected areas (IUCN, 2019). Vector datasets (GADM, NUTS and WDPA) were rasterized and all datasets were rescaled to a common resolution of 0.01 degrees (approximately 1 km at the equator).

We use two datasets of gridded population in SSP scenarios. Gao (2017) has a high spatial resolution (0.01 degrees); Murakami & Yamagata (2019) has a much lower resolution (0.5 degrees) but includes gridded GDP (PPP) in addition to population. We combined these datasets to create an upscaled dataset for gridded GDP at 0.01 degrees by multiplying GDP per capita from the low resolution dataset with the population from the high resolution data. US consumer price indexes were used to convert to USD (2010). The baseline SSP scenario and year used in our model is SSP2, year 2050.

Unfortunately we lack a global dataset representing access to electricity. To create a rough proxy dataset, we assume that 1 km pixels with a total pixel-GDP over a threshold value of 100,000 USD (2010) are connected to the electricity grid. This approach has several limitations. For example, it incorrectly assumes that all countries and subregions expand their grids at the same rate relative to their GDP and it disregards remote areas with attractive electricity supply resources such as wind- or hydropower. However, it also has the advantage that local access to electricity scales naturally with scenarios for future GDP growth.

GIS calculations based on user-defined parameters

The variable transformations described above in the Preprocessing section are performed immediately after the ERA5 datasets have been successfully downloaded, and determine the datasets that are stored on disk. In this section, we describe GIS calculations that combine the stored wind and solar datasets with user assumptions to produce potential installed capacity for each region and resource class, and capacity factors for each region, resource class and hour.

Region definitions

To define land areas for geographical regions, input can be given as names or codes of countries or subregions in the GADM or NUTS databases. For example, "United Kingdom" or "North Yorkshire" are valid GADM names, and "UK" or "UKE22" are valid NUTS codes. Offshore areas are allocated to the region with the closest land pixel, regardless of distance to shore. Major lakes (larger than 1000 km^2) are also identified based on the land cover dataset and are available for offshore wind power.

Based on the user-defined region names, we create two region indicator rasters (1280x640 matrices of integer region codes) with 1 km resolution for onshore and offshore regions.

Transmission GIS assumptions

In our energy system model, inter-regional transmission costs are a function of distance and whether the connection is land-based or subsea. Connections can only be made between neighboring regions.

Transmission costs between regions are estimated using the region indicator rasters. The length of a transmission line is assumed to be the great circle distance between population-weighted geographical region centers. If two regions are adjacent in the onshore raster, the line is assumed to be land-based. If not, and the regions are adjacent in the offshore raster, then the line is assumed to be sub-sea.

Algorithm for calculating potentials and capacity factors

The general algorithm used to calculate potentials and capacity factors is as follows. Each step requires a number of user-specified parameters. These are listed along with their default values in a section further below.

- Use auxiliary datasets to create masks (boolean rasters) that indicate in which pixels wind or solar installations are possible.
- Allocate each pixel to a resource class based on average annual capacity factor (for solar) or wind speed.
- Iterate over each pixel:
  - if the pixel is not allowed (masked out) then skip to the next pixel
  - look up the region and resource class of that pixel
  - calculate
    `pixel potential = pixel area * power density * remaining area factor`
  - add `pixel potential` to the running total potential for that region and resource class
  - Iterate over all hours of the year:
    - calculate the hourly capacity factor of an installation in that pixel
    - add the resulting capacity factor to the running total capacity factor for that region, resource class and hour
    - increment a count variable for that region, resource class and hour

- Calculate the average capacity factor for each region, resource class and hour by dividing the running total capacity factor by the count

Dual resolutions, wind rescaling and turbine curves

The situation is actually somewhat some complex than the algorithm above indicates. The ERA5 datasets have a significantly lower spatial resolution (31 km) than the auxiliary datasets (1 km). The pixel iteration is therefore performed in two steps: first we iterate over the low resolution pixels, then we iterate over all high resolution pixels contained within each low resolution pixel. This increases computational efficiency by moving low resolution calculations and lookups outside the "hot inner loop" of the high resolution iteration.

There are also differences in the algorithms for solar and wind. For solar PV and CSP, resource class allocation is based on annual average capacity factors, i.e. simple annual means of the low resolution GTI and DNI variables created in the preprocessing step.

However, for wind power, resource class allocation is based on high resolution annual average wind speeds from the Global Wind Atlas (DTU, 2019). We also rescale low resolution ERA5 hourly wind speeds to match high resolution annual average wind speeds from the GWA. In this way we capture geographical variations in wind power output caused by local differences in topography and land cover. Then a turbine curve is applied (see below) to further transform rescaled hourly wind speeds into instantaneous capacity factors. The increased computational complexity of the wind algorithm caused by the higher resolution and turbine curve calculations leads to run times an order of magnitude slower than for solar (minutes instead of seconds).

Subtechnologies and masks

The algorithm above is executed separately for wind and solar. For each of these, we calculate masks (and ultimately capacity factors and potentials) for various subtechnologies. For wind, we distinguish between onshore wind class A (within x kilometers of our grid proxy), onshore class B (further than x km from the proxy) and offshore wind. Solar is similarly divided into utility-scale PV class A and class B, rooftop PV, and CSP class A and B. Again, each of these are further divided into a number of resource classes, five by default. Default parameters are listed below.

Accounting for PV and CSP area overlap

Both utility-scale PV and CSP plants use the same parameter assumptions for masks. In other words, the land areas available for both technologies overlap perfectly, and the total solar capacity potential in the downstream energy model is roughly twice as high as it should be. To correct this error, our GIS code for solar potentials not only calculates potential capacity per region and resource class for PV and CSP individually, but also the total area available for solar installations by region, PV class and CSP class. This information on the combined (two-dimensional) regional potential of PV and CSP is passed on to the energy model, which can then correctly account for the land overlap (see

equations in our energy model description).

We do not currently account for overlap between solar and onshore wind, since both technologies can coexist on the same site with only marginal losses of land area due to the need for maintenance roads for wind turbines. If required, the code can be extended to account for this limited overlap as well.

Note on PV efficiency and rooftop assumptions

Our approach does not explicitly take into account PV module efficiency or inverter losses. Instead, these are implicitly included in parameters that are expressed in terms of peak output of the PV system, e.g. investment cost (€/kW_p) and maximum power density of a PV plant (W_p/m^2).

Similarly, for rooftop PV we do not explicitly consider factors such as building density or rooftop area, orientation and inclination. These are implicitly included in the parameters for power density and remaining area available for installations. Rooftop installations are assumed to be oriented towards south (or towards north for the southern hemisphere). One disadvantage of this approach is that we do not currently capture temporal smoothing of aggregated PV output from buildings with roofs with "suboptimal" orientation. In real power systems, the power output profile from rooftop PV installations that face nearly east or west can be shifted several hours compared to south-oriented systems. The smoothing effect that arises from rooftops with a continuous distribution of directions from east to south to west may indeed be advantageous from a system perspective.

Hydro GIS assumptions

We also estimate potentials and hourly production for hydropower based on the work of Gernaat, Bogaart, Vuuren, Biemans, & Niessink (2017). Gernaat et al. used global hydrology and topography maps combined with a hydropower investment model to identify 60,000 suitable sites for hydropower. For each site, its longitude and latitude coordinates, location-specific levelized cost of electricity (LCOE), potential reservoir size, annual electricity generation and monthly water inflow were estimated. Gernaat also kindly supplied his model calculations for about 7000 existing hydropower plants based on the Global Reservoir and Dam Database (Lehner et al., 2011) with coordinates, annual generation and monthly inflow.

The database for existing hydropower is incomplete. To account for missing capacity, we scale individual sites to match total hydropower capacity by country based on World Energy Council (2016). Similarly, the databases only extend to 60 degrees northern latitude, which impacts Russia and Canada and almost entirely neglects hydropower in the Nordic countries. Existing hydropower capacity is corrected using World Energy Council (2016) as above, but for this reason there is no future hydro potential in these regions above 60 degrees north. This is not a major limitation for the Nordic countries due a broad political agreement to not exploit remaining hydro potential in northern Scandinavia [REF].

We disaggregate hydropower potential by region, resource class and reservoir capacity. Default

parameters are listed below. Hourly water inflow is obtained by assuming that the inflow remains constant during each month (i.e. we divide monthly inflow by 720 hours). Finally, our energy system model tracks reservoir levels for one aggregated reservoir for each region and resource class.

Default GIS parameters

Parameters for onshore and offshore wind power

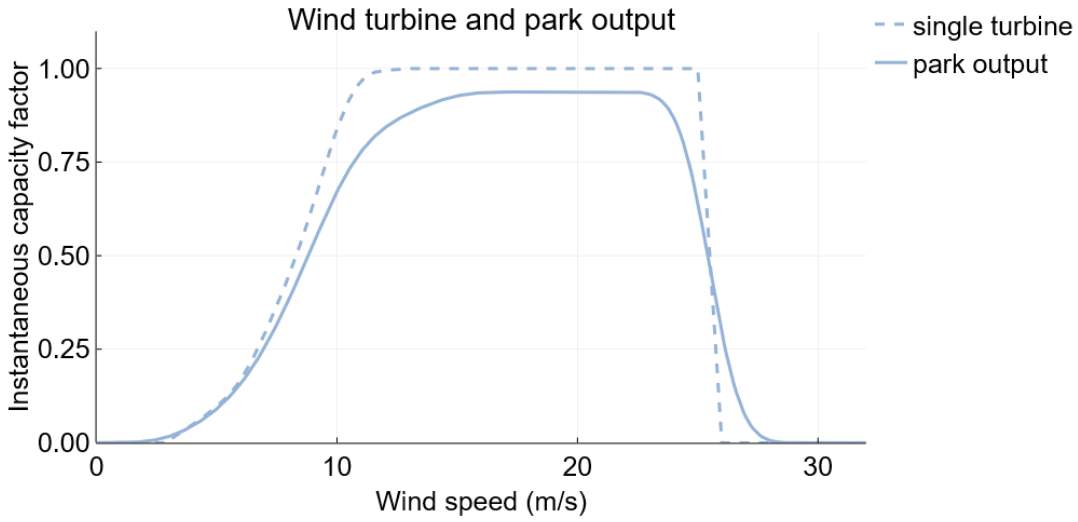
Below we list our default parameter assumptions for the GIS wind module. These parameters are all configurable by the user.

- General assumptions
  - SSP scenario and year: SSP2, 2050
  - year of ERA5 solar and wind data: 2018
- Area & capacity assumptions
  - onshore turbine density: 5 W/m^2
  - offshore turbine density: 8 W/m^2
  - onshore share of remaining area available for wind farms: 8 %
  - offshore share of remaining area available for wind farms: 33 %
- Mask assumptions
  - maximum distance to electricity grid (for class B and offshore): 300 km
  - maximum population density: 150 persons/km^2
  - maximum water depth for offshore wind: 40 m
  - minimum distance to shore for offshore wind: 5 km
  - excluded land cover types for onshore wind: water, wetlands, urban
  - excluded protected area categories (IUCN codes from the World Database of Protected Areas): "Strict Nature Reserve", "Wilderness Area", "National Park", "Natural Monument", "Habitat/Species Management", "Not Reported"
- Resource class assumptions
  - wind resource classes are determined by annual average wind speed in each high resolution pixel (i.e. 0.01 degrees), according to the following table:

	Onshore	Offshore
class 1	2 - 5 m/s	3 - 6 m/s
class 2	5 - 6 m/s	6 - 7 m/s
class 3	6 - 7 m/s	7 - 8 m/s
class 4	7 - 8 m/s	8 - 9 m/s
class 5	8+ m/s	9+ m/s

- Wind turbine and park output curves
  - To convert instantaneous wind speeds to capacity factors, we use an output profile based on the Vestas V112 3 MW wind turbine. Local wind speed variations are considered

by smoothing the turbine power curve by applying a convolution with a Gaussian probability density function with a standard deviation of 1 m/s. Availability and electrical losses are assumed to be 6% and average wind farm wake losses at 11.5%. Single turbine and resulting wind farm output are shown in the figure below.



Parameters for solar PV and CSP

Below we list our default parameter assumptions for the GIS solar module. These parameters are all configurable by the user.

- General assumptions
  - SSP scenario and year: SSP2, 2050
  - year of ERA5 solar and wind data: 2018
- Area & capacity assumptions
  - PV plant power density: 45 W_p/m^2
  - CSP plant power density: 35 W/m^2
  - share of remaining area available for PV or CSP plants: 5 %
  - share of remaining area available for PV rooftop installations: 5 %
- Mask assumptions
  - maximum distance to electricity grid (for class B and offshore): 300 km
  - maximum population density for PV or CSP plants: 150 persons/km^2
  - minimum population density for PV rooftop installations: 200 persons/km^2
  - excluded land cover types for PV or CSP plants: water, forest, croplands and “woody savannas” (a category that includes much of the forests of Scandinavia)
  - excluded protected area categories (IUCN codes from the World Database of Protected Areas): “Strict Nature Reserve”, “Wilderness Area”, “National Park”, “Natural Monument”, “Habitat/Species Management”, “Not Reported”
- Resource class assumptions
  - PV and CSP resource classes are determined by annual average capacity factor of the

solar insolation in each high resolution pixel (i.e. 0.01 degrees), according to the following table:

	solar PV	CSP
class 1	8 - 14 %	10 - 18 %
class 2	14 - 18 %	18 - 24 %
class 3	18 - 22 %	24 - 28 %
class 4	22 - 26 %	28 - 32 %
class 5	26+ %	32+ %

Parameters for hydropower

- Four classes of reservoir sizes, measured in time required to completely discharge a full reservoir when running the turbines at maximum capacity.

	discharge time [weeks]
class A	0 - 0
class B	0 - 6
class C	6 - 26
class D	26+

- Four levelized cost classes:

	LCOE [$/MWh]
class 1	0 - 5
class 2	5 - 10
class 3	10 - 20
class 4	20 - 50

- Storage classes A-D and cost classes 1-4 are combined to create a matrix of 16 classes representing future hydro potential (A1, A2, A3, A4, B1, B2, …). Then we add one class X0 for existing hydro capacity.

Part 2: Energy system model (Supergrid.jl)

The following is a mathematical representation of our capacity expansion model. All equations are linear; the resulting optimization model is therefore a pure linear programming model. By convention, model variables use upper-case letters and parameters use lower-case letters. The full model is written in Julia using the JuMP optimization package and is available at ("Github (Supergrid)," n.d.).

Sets

$r \in R$	model region (aliases: $r1$, $r2$)
$k \in K$	electricity generation technologies
$c \in C$	resource classes (aliases: $c1$, $c2$)
$h \in H$	hour of the year
$s \in S$	electricity storage technologies
$f \in F$	fuels for electricity generation

Variables

$E_{r,k,c,h}$	electricity generation in hour h of region r, technology k and resource class c [GWh_{elec}/period]
$C_{r,k,c}$	installed capacity in region r of technology k and resource class c [GW_{elec}]
$C^{pv}_{r,c1,c2}$	installed solar PV capacity in region r in areas (GIS pixels) with PV class $c1$ and CSP class $c2$ [GW_{elec}]
$C^{csp}_{r,c1,c2}$	installed CSP capacity in region r in areas (GIS pixels) with PV class $c1$ and CSP class $c2$ [GW_{elec}]
$T_{r1,r2,h}$	electricity transmission in hour h from region $r1$ to region $r2$ [GWh_{elec}/period]
$C^{T}_{r1,r2}$	installed transmission capacity between region $r1$ and $r2$ [GW_{elec}]
$E^{charge}_{r,s,c,h}$	electricity used for charging in hour h of region r, storage technology s and resource class c [GWh_{elec}/period]
$S_{r,s,c,h}$	stored electricity in hour h of region r, storage technology s and resource class c [GWh_{elec}]
$F_{r,f}$	annual fuel use of fuel f in region r [GWh_{fuel}/year]
M_r	CO_2 emissions in region r [kton CO_2/year]

Parameters

$p^{potential}_{r,k,c}$	maximum potential capacity in region r of technology k and resource class c [GW]
t	length of a model time period [h/period] (one hour by default)
$d_{r,h}$	electricity demand in region r and hour h [GW]
$f_{r,k,c,h}$	capacity factor in hour h of region r, technology k and resource class c

$f^{\text{inflow}}_{r,s,c,h}$	capacity factor of inflow in hour h of region r, storage technology s and resource class c
$l^{T}_{r1,r2}$	electricity transmission losses between region $r1$ and $r2$
c^{F}_{f}	cost of fuel f [€/MWh$_{\text{fuel}}$]
c^{CO2}	carbon tax [€/ton CO_2]
c^{var}_{k}	variable maintenance cost of technology k [€/MWh$_{\text{elec}}$]
$c^{\text{hydro}}_{r,c}$	levelized cost of hydropower of resource class c in region r [€/MWh$_{\text{elec}}$]
$c^{\text{invest}}_{k,c}$	investment cost of technology k and resource class c [€/kW]
c^{fixed}_{k}	fixed maintenance cost of technology k [€/kW/year]
$c^{T,\text{invest}}_{r1,r2}$	transmission investment cost between region $r1$ and $r2$ [€/kW]
$c^{T,\text{fixed}}_{r1,r2}$	transmission fixed maintenance cost between region $r1$ and $r2$ [€/kW/year]
p^{CRF}_{k}	capital recovery factor for technology k [1/year]
$p^{T,\text{CRF}}$	capital recovery factor for transmission [1/year]
$t^{\text{discharge}}_{r,s,c}$	discharge time of storage technology s [h]
p^{inflow}	minimum share of hydro inflow for electricity generation
f_k	fuel f used for electricity generation technology k
m_f	CO_2 emissions of fuel f [kg CO_2/kWh$_{\text{fuel}}$]
m^{cap}	global cap on CO_2 emissions per electricity demand [kg CO_2/kWh$_{\text{elec}}$]
$a^{\text{solar}}_{r,c1,c2}$	total area available for solar PV or CSP in region r of areas (GIS pixels) with PV class $c1$ and CSP class $c2$ [thousand km^2]
$p^{\text{density,pv}}$	area density of solar PV plants [W/m^2 = MW/km^2]
$p^{\text{density,csp}}$	area density of CSP plants [W/m^2 = MW/km^2]
p^{bio}	maximum share of bioelectricity in annual electricity generation
p^{ramp}	maximum hourly ramping rate (as share of installed capacity)

Variable bounds

All variables except CO_2 emissions M_r have a lower bound of zero (we allow negative emissions using bioenergy with CCS).

$$E_{r,k,c,h}, C_{r,k,c}, C^{\text{pv}}_{r,c1,c2}, C^{\text{csp}}_{r,c1,c2}, T_{r1,r2,h}, C^{T}_{r1,r2}, E^{\text{charge}}_{r,s,c,h}, S_{r,s,c,h}, F_{r,f} \geq 0 \quad \forall r, k, c, c1, c2, h, r1, r2, s, f$$

Renewable capacities are limited by potentials determined by our GIS package.

$$C_{r,k,c} \leq p^{\text{potential}}_{r,k,c} \quad \forall r, k, c : k \text{ is renewable}$$

No transmission between non-neighboring countries (indicated by zero investment cost):

$$C^{T}_{r1,r2} = 0 \quad \forall r1, r2 : c^{T,\text{invest}}_{r1,r2} = 0$$

Constraints

Electricity generation is limited by capacity:

$$E_{r,k,c,h} \leq f_{r,k,c,h} C_{r,k,c}\, t \quad \forall r, k, c, h$$

Hourly electricity balance - generation plus imports minus exports minus stored electricity must exceed demand:

$$\sum_{k,c} E_{r,k,c,h} + \sum_{r2} (1 - l^{T}_{r2,r}) T_{r2,r,h} - \sum_{r2} T_{r,r2,h} - \sum_{s,c} E^{\text{charge}}_{r,s,c,h} \geq d_{r,h}\, t \quad \forall r, h$$

Stored energy balance: current storage level minus previous storage level must not exceed charging electricity (batteries) plus renewable inflow (hydro or CSP) minus discharge. Hour indices wrap around ($S_{r,s,c,0} = S_{r,s,c,8760}$). We do not currently consider self-discharge. Storage variables are also indexed by class since hydropower has resource classes.

$$S_{r,s,c,h} - S_{r,s,c,h-1} \leq E^{\text{charge}}_{r,s,c,h} + f^{\text{inflow}}_{r,s,c,h} C_{r,s,c}\, t - \frac{1}{\eta_s} E_{r,s,c,h} \quad \forall r, s, c, h$$

Storage capacity limit: current storage level must not exceed its installed capacity (in GW) multiplied by its discharge time (hours).

$$S_{r,s,c,h} \leq C_{r,s,c} t^{\text{discharge}}_{r,s,c} \quad \forall r, s, c, h$$

Minimum water flow for hydropower: hydropower must use at least a fraction p^{inflow} of its hourly inflow for electricity generation.

$$E_{r,\text{hydro},c,h} \geq p^{\text{inflow}} f^{\text{inflow}}_{r,\text{hydro},c,h} C_{r,\text{hydro},c}\, t \quad \forall r, c, h$$

No charging of hydro and CSP: electricity cannot be stored in these technologies. (We do not currently implement pumped hydro storage in our model. Hydropower and CSP are classified as storage technologies so their water and thermal storage levels can be tracked using the constraints above, but they cannot be charged using surplus electricity.)

$$E^{\text{charge}}_{r,s,c,h} = 0 \quad \forall r, s, c, h : s \in \{\text{hydro}, \text{CSP}\}$$

Electricity transmission between regions is limited by transmission capacity:

$$T_{r1,r2,h} \leq C^{T}_{r1,r2}\, t \quad \forall r1, r2, h$$

Symmetry of transmission capacity:

$$C^T_{r1,r2} = C^T_{r2,r1} \quad \forall r1, r2$$

Calculate annual fuel use:

$$F_{r,f} = \sum_{k,c,h:\ f_k=f} \frac{E_{r,k,c,h}}{\eta_k} \quad \forall r, f$$

Calculate annual CO_2 emissions:

$$M_r = \sum_f m_f F_{r,f} \quad \forall r$$

Global cap on CO_2 emissions:

$$\sum_r M_r \leq m^{\text{cap}} \sum_{r,h} d_{r,h}\, t$$

Limit on bioelectricity share of annual electricity demand:

$$\sum_{k,c,h:\ f_k=\text{bio}} E_{r,k,c,h} \leq p^{\text{bio}} \sum_h d_{r,h}\, t \quad \forall r$$

Consider overlap of PV and CSP land area by introducing variables for combined solar capacity for each combination of PV and CSP resource classes. The total combined capacities are then linked to the standard capacity variables.

$$\frac{1}{p^{\text{density,pv}}} C^{\text{pv}}_{r,c1,c2} + \frac{1}{p^{\text{density,csp}}} C^{\text{csp}}_{r,c1,c2} \leq a^{\text{solar}}_{r,c1,c2} \quad \forall r, c1, c2$$

$$C_{r,\text{pv},c} = \sum_{c2} C^{\text{pv}}_{r,c,c2} \quad \forall r, c$$

$$C_{r,\text{csp},c} = \sum_{c1} C^{\text{csp}}_{r,c1,c} \quad \forall r, c$$

Optional ramping limits for thermal technologies. Hour indices wrap around ($E_{r,k,c,0} = E_{r,k,c,8760}$).

$$E_{r,k,c,h} - E_{r,k,c,h-1} \leq p^{\text{ramp}} f_{r,k,c,h} C_{r,k,c}\, t \quad \forall r, k, c, h : k \text{ is thermal}$$
$$E_{r,k,c,h} - E_{r,k,c,h-1} \geq -p^{\text{ramp}} f_{r,k,c,h} C_{r,k,c}\, t \quad \forall r, k, c, h : k \text{ is thermal}$$

Objective function

Minimize the sum of all cost components [M€/year]: fuel, carbon tax, variable maintenance, levelized cost of hydropower (from Gernaat, Bogaart, Vuuren, Biemans, & Niessink (2017)), discounted invest-

ments plus fixed maintenance, and discounted transmission investments plus fixed maintenance.

$$\frac{1}{1000}(\sum_{r,f} c_f^F F_{r,f} \;+\; \sum_r c^{\mathrm{CO2}} M_r \;+\; \sum_{r,k,c,h} c_k^{\mathrm{var}} E_{r,k,c,h} \;+\; \sum_{r,c,h} c_{r,c}^{\mathrm{hydro}} E_{r,\mathrm{hydro},c,h}) \\ +\; \sum_{r,k,c} (c_{k,c}^{\mathrm{invest}} p_k^{\mathrm{CRF}} + c_k^{\mathrm{fixed}}) C_{r,k,c} \;+\; \frac{1}{2} \sum_{r1,r2} (c_{r1,r2}^{T,\mathrm{invest}} p^{T,\mathrm{CRF}} + c_{r1,r2}^{T,\mathrm{fixed}}) C_{r1,r2}^T$$

CSP investment costs

In our energy model, CSP is assumed to be of type solar tower. Solar tower investment costs vary with the size of the solar collector field (parameterized by the "solar multiple", a measure of the collector field output relative to the generator capacity) and its thermal storage capacity (measured in full load hours of potential power output). Our default plant has a solar multiple of 3 and 12 hours of thermal storage and has a current investment cost of about 9200 USD(2010)/kW (IRENA, 2012). After conversion to EUR and assuming a 25% cost reduction to 2050, the resulting default investment cost in our model is 6000 €/kW.

Approximately 35% of the total investment cost can be attributed to the solar collector field and about 10% to the thermal storage system for a plant with our default parameters (IRENA, 2012). For solar towers with other parameters, we assume investment costs vary linearly with solar multiple (SM) and thermal storage capacity (TSC). The investment cost (IC) relation is therefore:

$$\mathrm{IC} = \mathrm{IC}_{\mathrm{default}}(0.55 + 0.35\frac{\mathrm{SM}}{3} + 0.10\frac{\mathrm{TSC}}{12})$$

Part 3: Additional model results

Here we expand on figure 5 in the main paper by providing the electricity supply mix in Europe disaggregated by model region. A main driver of model results is the relative lack of renewable resources in Germany. In the case with unlimited transmission, Germany becomes a net importer and relies on neighboring regions with more abundant renewables. When no interregional transmission is allowed, Germany must exhaust virtually all its resources to fulfil the annual demand, including the high-cost low-grade resources. This scenario is only barely feasible.

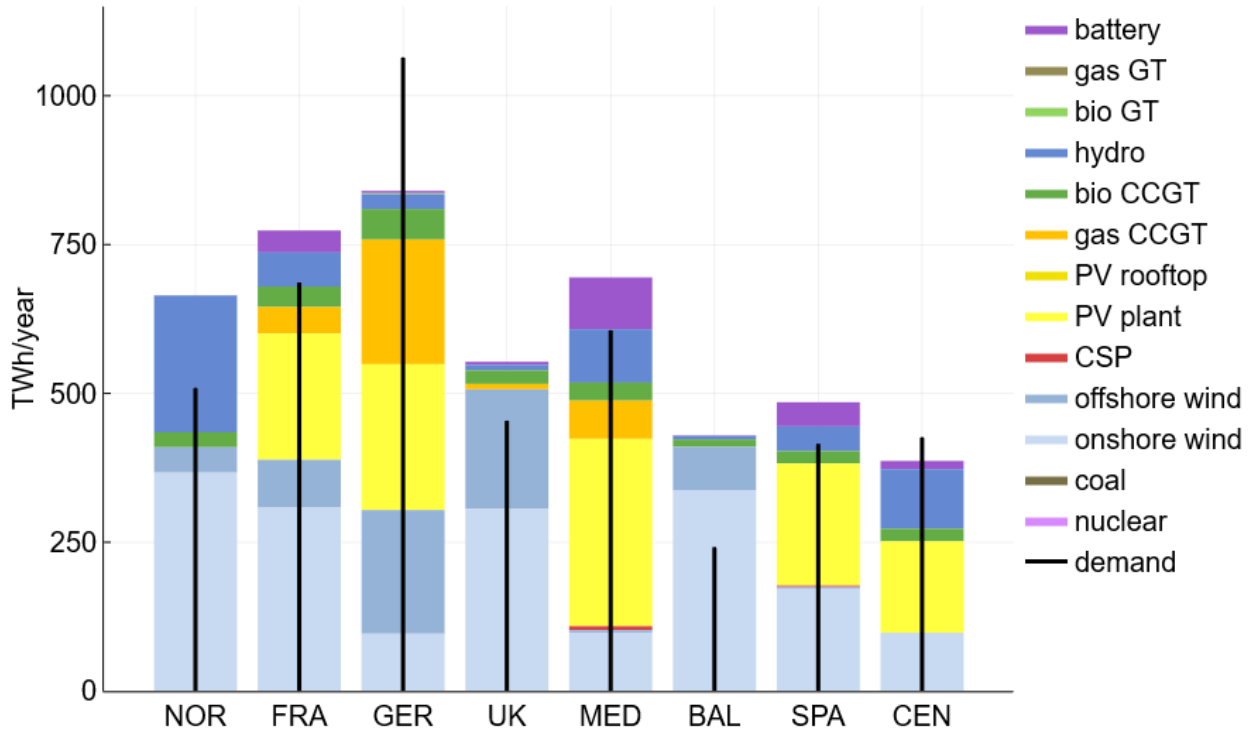


Figure S1. Annual electricity generation in European regions (TWh/year) for the default case (with unlimited transmission). Black vertical lines are annual electricity demand. Scenario: carbon cap 25 g CO_2/kWh, no nuclear, existing hydro only.

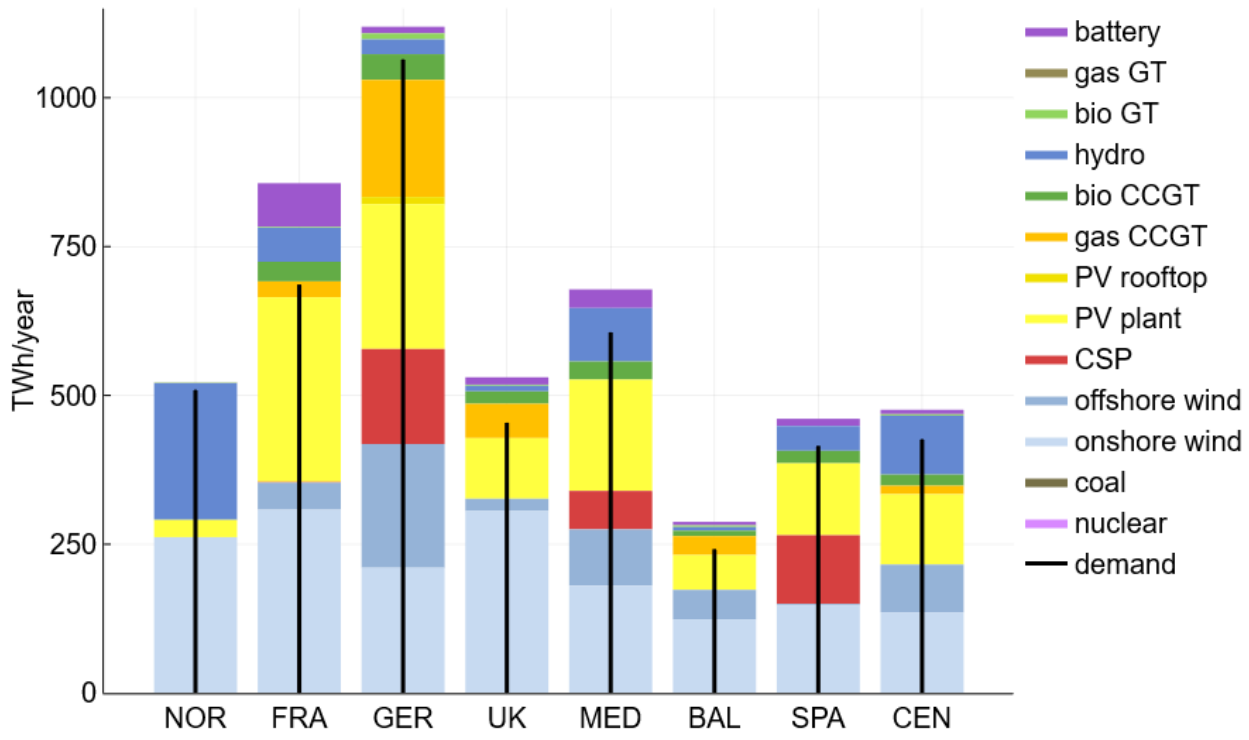


Figure S2. Annual electricity generation in European regions (TWh/year) for the alternative case with no transmission. Black vertical lines are annual electricity demand. Scenario: carbon cap 25 g CO_2/kWh, no nuclear, existing hydro only.

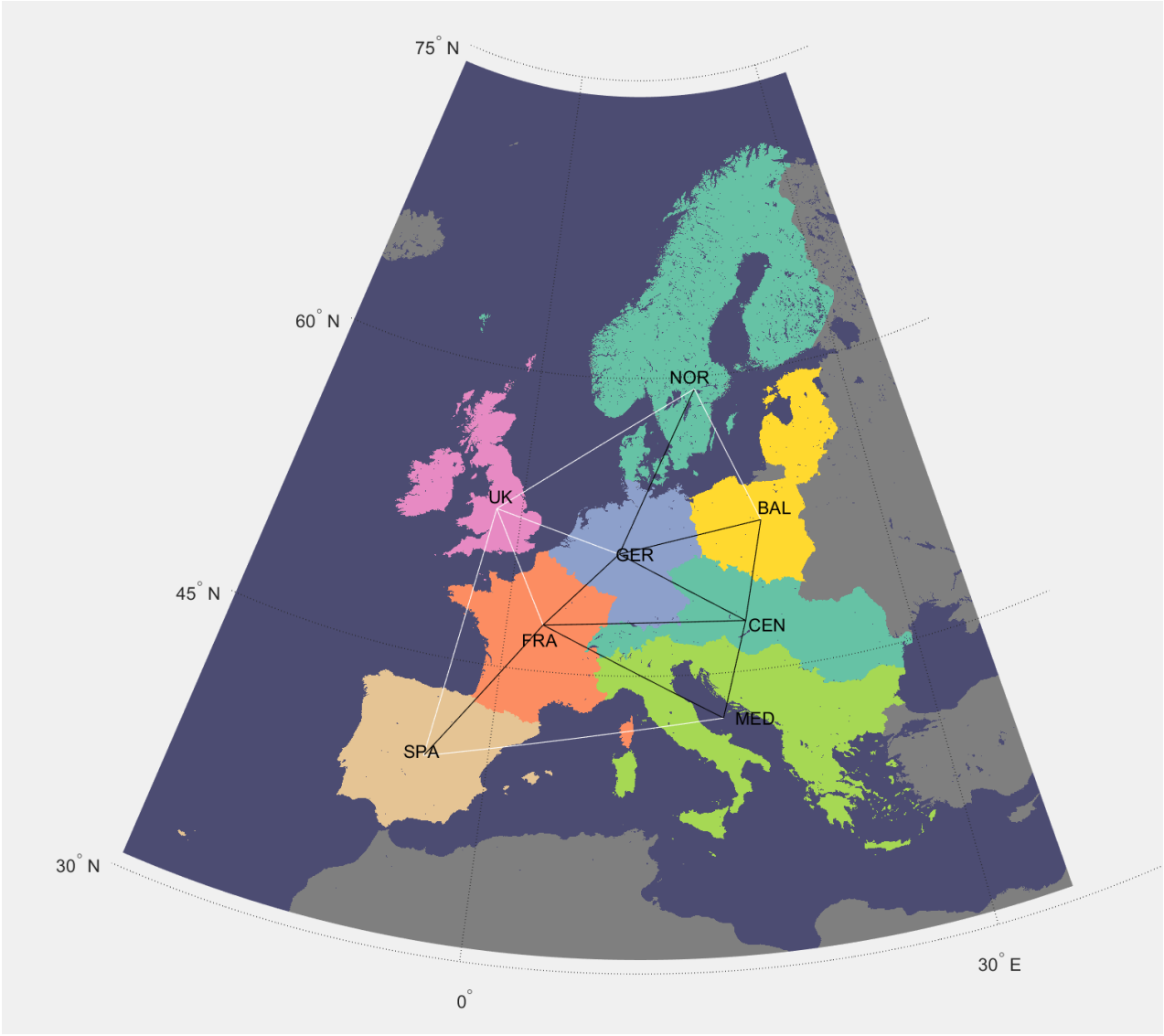


Figure S3. Model regions in Europe. Black lines are potential land-based transmission connections and white lines are potential subsea connections.

References

Amante, C., & Eakins, B. W. (2009). *ETOPO1 arc-minute global relief model: Procedures, data sources and analysis*. Retrieved from doi:10.7289/V5C8276M

Copernicus. (2017). ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Retrieved September 30, 2019, from Copernicus Climate Change Service Climate Data Store (CDS) website: https://cds.climate.copernicus.eu/cdsapp#!/home

DTU. (2019). Global Wind Atlas 2.0. Retrieved September 30, 2019, from Global Wind Atlas website: https://globalwindatlas.info

Eurostat NUTS. (n.d.). Retrieved September 30, 2019, from Nomenclature of territorial units for statistics website: https://ec.europa.eu/eurostat/web/nuts/background

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*(1), 168–182. https://doi.org/10.1016/j.rse.2009.08.016

GADM. (n.d.). Global dataset of administrative areas, version 3.6. Retrieved September 30, 2019, from https://gadm.org/

Gao, J. (2017). Downscaling global spatial population projections from 1/8-degree to 1-km grid cells. *NCAR Technical Note, NCAR/TN-537+ STR. Doi*, *10*, D60Z721H.

Gernaat, D. E. H. J., Bogaart, P. W., Vuuren, D. P. van, Biemans, H., & Niessink, R. (2017). High-resolution assessment of global technical and economic hydropower potential. *Nature Energy*, *2*(10), 821–828. https://doi.org/10.1038/s41560-017-0006-y

Github (GlobalEnergyGIS). (n.d.). Retrieved from https://github.com/niclasmattsson/GlobalEnergyGIS/

Github (Supergrid). (n.d.). Retrieved from https://github.com/niclasmattsson/Supergrid/

Grena, R. (2012). Five new algorithms for the computation of sun position from 2010 to 2110. *Solar Energy*, *86*(5), 1323–1337. https://doi.org/10.1016/j.solener.2012.01.024

Hogan, R. (2015). *Radiation Quantities in the ECMWF model and MARS*. 9. Retrieved from https://www.ecmwf.int/sites/default/files/elibrary/2015/18490-radiation-quantities-ecmwf-model-and-mars.pdf

IRENA. (2012). Renewable Energy Cost Analysis - Concentrating Solar Power. Retrieved May 14, 2019, from https://www.irena.org/publications/2012/Jun/Renewable-Energy-Cost-Analysis---Concentrating-Solar-Power

IUCN, U.-W. (2019). *The World Database on Protected Areas (WDPA)*. Retrieved from https://www.protectedplanet.net/

Jacobson, M. Z., & Jadhav, V. (2018). World estimates of PV optimal tilt angles and ratios of sunlight incident upon tilted and tracked PV panels relative to horizontal panels. *Solar Energy*, *169*, 55–66. https://doi.org/10.1016/j.solener.2018.04.030

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., … Wisser, D. (2011). High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment*, *9*(9), 494–502. https://doi.org/10.1890/100125

Loutzenhiser, P. G., Manz, H., Felsmann, C., Strachan, P. A., Frank, T., & Maxwell, G. M. (2007). Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation. *Solar Energy*, *81*(2), 254–267. https://doi.org/10.1016/j.solener.2006.03.009

Murakami, D., & Yamagata, Y. (2019). Estimation of gridded population and GDP scenarios with spatially explicit statistical downscaling. *Sustainability*, *11*(7), 2106.

Riahi, K., van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., … Tavoni, M. (2017). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, *42*, 153–168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Wikipedia: Solar irradiance. (2019). In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Solar_irradiance&oldid=918929661

World Energy Council. (2016). *World Energy Resources - Hydropower 2016*. Retrieved from https://web.archive.org/web/*/https://www.worldenergy.org/wp-content/uploads/2017/03/WEResources_Hydropower_2016.pdf