**Documentation for GCAMUSAJobs: An R package for employment projections based on GCAM-USA power sector outcomes**

Di Sheng, Brian O'Neill, Stephanie Waldhoff, and Matthew Binsted

Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD, USA

Table of Contents

[1. Package Overview 3](#_Toc178345327)

[2. Inputs of the package 3](#_Toc178345328)

[2.1 GCAM-USA output 3](#_Toc178345329)

[2.2 GCAM-USA assumptions 4](#_Toc178345330)

[2.2.1 Capacity factor 4](#_Toc178345331)

[2.2.2 Retirement 4](#_Toc178345332)

[2.2.3 O&M cost 5](#_Toc178345333)

[2.3 JEDI employment factors 5](#_Toc178345334)

[2.3.1 Prerequisite 5](#_Toc178345335)

[2.3.1 Biomass and biomass-CCS 6](#_Toc178345336)

[2.3.2 Coal and coal-CCS 6](#_Toc178345337)

[2.3.3 Natural gas and natural gas-CCS 6](#_Toc178345338)

[2.3.4 Refined liquids and Refined liquids-CCS 7](#_Toc178345339)

[2.3.5 Solar and rooftop PV 7](#_Toc178345340)

[2.3.6 Onshore wind and offshore wind 7](#_Toc178345341)

[2.3.7 Nuclear 7](#_Toc178345342)

[2.3.8 Hydro 7](#_Toc178345343)

[2.3.9 Geothermal 7](#_Toc178345344)

[3. Processes 8](#_Toc178345345)

[3.1 Power production calculation 8](#_Toc178345346)

[3.2 Capacity activity calculation 9](#_Toc178345347)

[3.2.1 Operational capacity and installed capacity 9](#_Toc178345348)

[3.2.2 Capacity retirement 10](#_Toc178345349)

[3.2.3 Simultaneous capacity addition and pre-mature retirement 10](#_Toc178345350)

[3.3 Job calculation 10](#_Toc178345351)

[3.3.1 Adjustment for O&M jobs 11](#_Toc178345352)

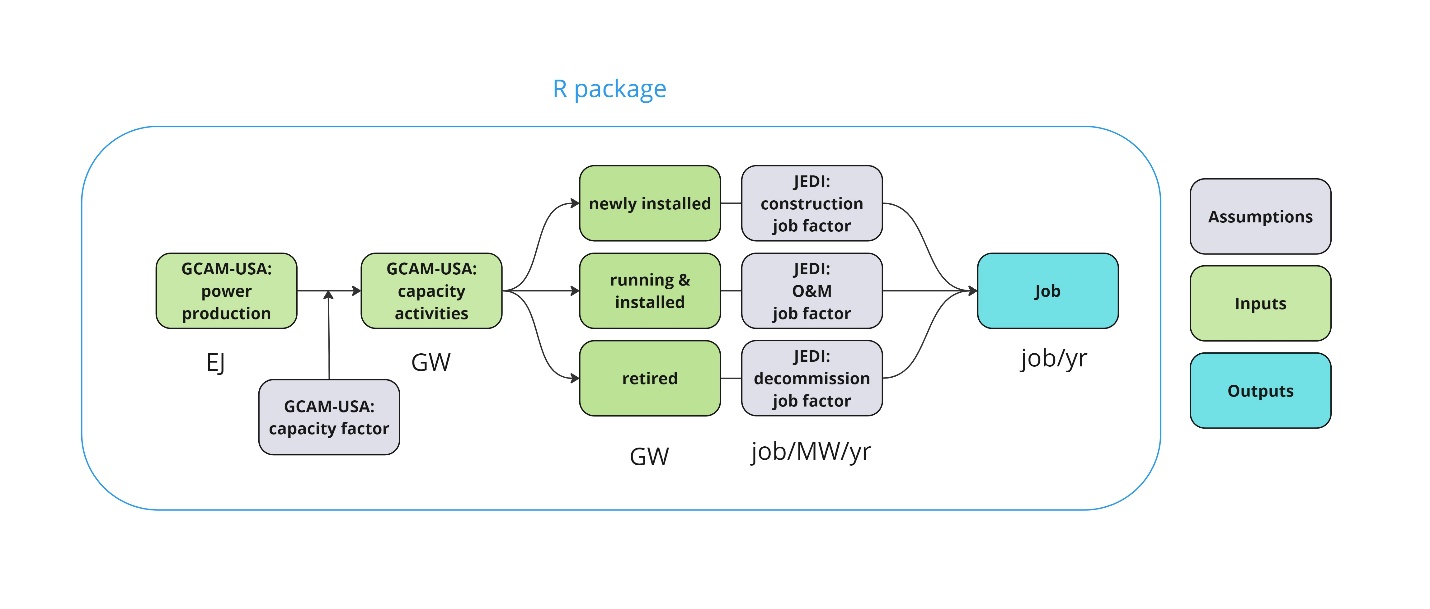
[3.3.2 Annual construction job EF 11](#_Toc178345353)

[4. Output of the package 12](#_Toc178345354)

[5. Future development 12](#_Toc178345355)

# 1. Package Overview

This package aims to provide employment projections in the power sector based on GCAM-USA output.



This package processes GCAM-USA output and integrates it with capacity factor and employment factor assumptions to estimate direct employment related to power generation. Employment factors (annual jobs per capacity) for various fuel and job types are derived from the Jobs & Economic Development Impacts (JEDI) Model by NREL (<https://www.nrel.gov/analysis/jedi/models.html>).

Section 2 outlines the inputs and assumptions, while Section 3 details the calculation methodology. Section 4 presents the outcome of this package, and Section 5 discusses future development.

# 2. Inputs of the package

This package uses three main inputs: GCAM-USA output, GCAM input data (supplemented by other sources), and employment factors from the JEDI model (supplemented by other sources). Below is an overview of each input.

## 2.1 GCAM-USA output

GCAM-USA provides annual power output projections (in exajoules, EJ) by technology and vintage for each model year. The model is calibrated to a base year (2015) and projects output every five years until 2100. For more information, refer to the [GCAM-USA model documentation](https://jgcri.github.io/gcam-doc/gcam-usa.html). This package is compatible with both GCAM-USA output in the original GCAM-USA output database (baseX), as well as queried project data (.dat). The package utilizes the rgcam package (available at [rgcam GitHub](https://github.com/JGCRI/rgcam)) to query outcomes from GCAM-USA output, and a query file containing all necessary queries is included in the package (*GCAMUSAJobs/inst/extdata/my\_batch.xml*). For guidance on querying relevant outcomes, please refer to the package vignette (<https://jgcri.github.io/GCAMUSAJobs/articles/package_vignette.html>).

Since the JEDI employment factors (discussed in ***Section 2.3***) are based on capacity, we need data on GCAM-USA electricity outcomes in capacity terms, including operational capacity, newly installed capacity, and retired capacity. However, GCAM-USA does not directly track capacity levels. Instead, we derive capacity information from annual power generation outcomes.

Power generation outcomes are obtained through two queries below, and detailed query information can be found in *GCAMUSAJobs/inst/extdata/my\_batch.xml*.

*1. 'elec energy input by elec gen tech by vintage'*

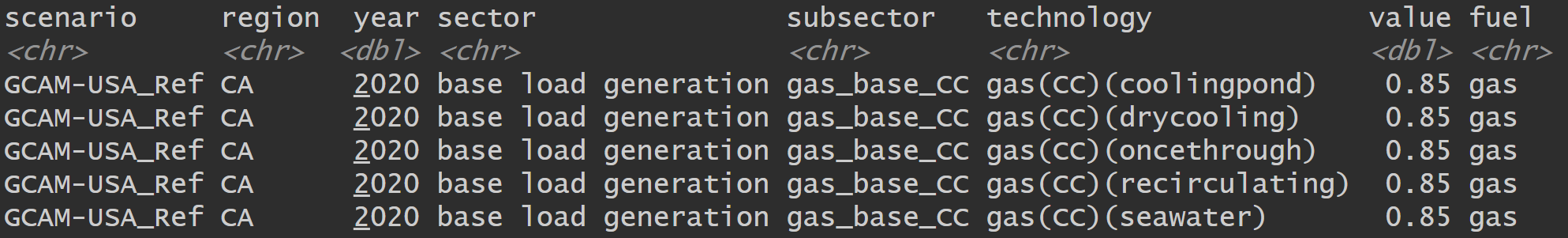
*2. 'elec gen by gen tech and cooling tech (incl cogen) by vintage'*

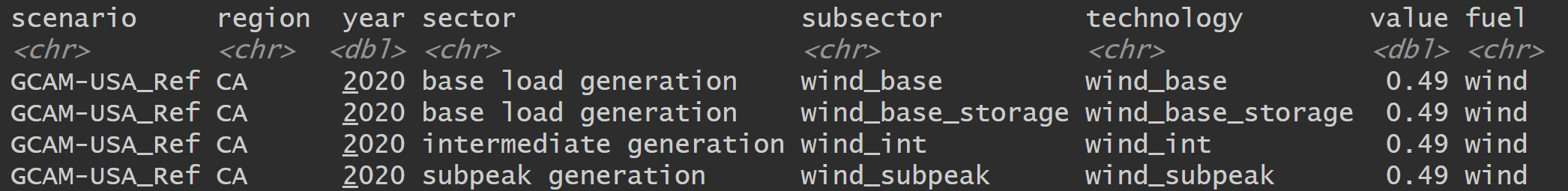
## 2.2 GCAM-USA assumptions

GCAM-USA assumptions used in this package consist of (1) *capacity factor by technology (supplemented by other sources)*, (2) *retirement assumptions*, and (3) *variable and fixed operation and maintenance (OM) cost*. Capacity factors and retirement assumptions help derive capacity levels for operational, newly installed, and retired capacity. The variable and fixed O&M costs are used to differentiate between variable and fixed O&M employment factors from JEDI (see ***3.4.1 Adjustment for O&M job***).

### 2.2.1 Capacity factor

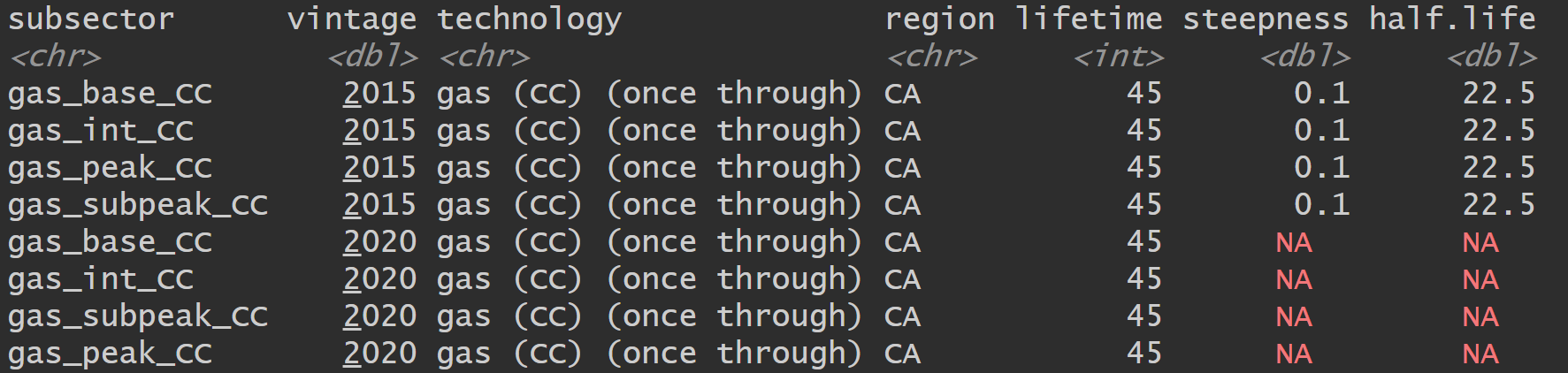
Query '*GCAM\_USA elec cap-fac by cooling tech*' provides state-level capacity factor by fuel technology and segment. For example,





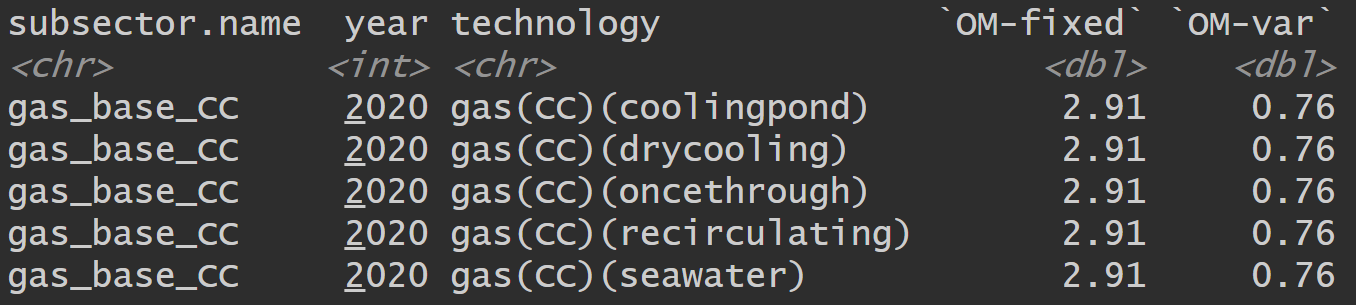
Note that GCAM-USA does not include capacity factors for hydro and geothermal power, and we use alternative data sources. For hydropower, we use monthly state-level capacity factors from Hall et al. (2003)1, with the annual state-level average used in this package. For geothermal power, due to limited data, we use a capacity factor of 0.76, based on estimates from the EIA2.

### 2.2.2 Retirement

The ***SCurve***, which provides assumed retirement assumptions by technology and vintage, is generated based on GCAM-USA v7.1 assumptions (GCAMUSAJobs/inst/extdata/GCAM/Assumption.R) and is included as package data. Users may need to update these retirement assumptions when using this package with different versions of GCAM-USA. For convenience, scripts and data to generate the ***SCurve*** for GCAM-USA v6 is also provided (GCAMUSAJobs/inst/extdata/GCAM/GCAM6). 

### 2.2.3 O&M cost

The ***OM\_cost***, providing fixed O&M cost (1975$/kW/year) and variable O&M cost (1975$/MWh) by fuel technology over time, is generated based on GCAM-USA v7.1 assumptions (GCAMUSAJobs/inst/extdata/GCAM/Assumption.R) and included as package data. Users should update the ***OM\_cost*** if GCAM-USA outputs are produced with different O&M cost assumptions.



## 2.3 JEDI employment factors

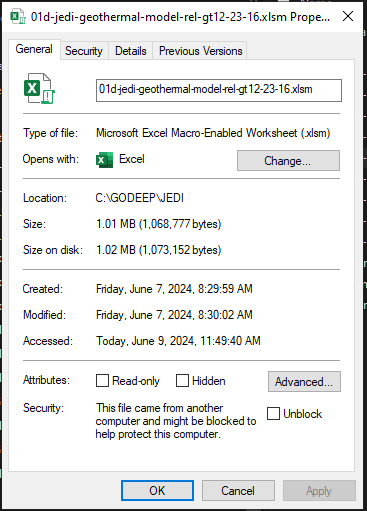
[NREL JEDI Models](https://www.nrel.gov/analysis/jedi/models.html), which we accessed on August 29, 2024, are used to calculate the annual employment factor (EF) input for this package, modified in some cases with other sources. The JEDI model provides job outcomes for a specified type of project, and we calculate the EF by dividing the jobs by project capacity. Additional adjustments for EF are summarized in **3.4.1** and **3.4.2**.

To extract JEDI outcomes across fuel types and states, we developed Python scripts. This section outlines the workflow for obtaining the JEDI employment factor, starting with the prerequisites for running the JEDI model via Python, followed by detailed steps for running individual JEDI models by fuel type.

### 2.3.1 Prerequisite

The JEDI model is Excel-based, and its macros are disabled by default for security reasons. To run the JEDI model through Python, macros must be enabled for each JEDI Excel file by the following process:

Excel file 🡪 Properties 🡪 General 🡪 check the “Unblock” for ***Security***



While the specific Python scripts for running each JEDI model may vary (see **2.3.1** to **2.3.9**), they follow a common structure:

Import library 🡪 specify JEDI model file 🡪 update “project data sheet” 🡪 update “local share” 🡪

specify macro for job results 🡪 data collection 🡪 save the data

where the “project data sheet” specifies the characteristics of the power plant, and “local share” indicates the proportion of economic activities related to power production occurring within the region, ranging from 0% to 100%.

### 2.3.1 Biomass and biomass-CCS

For **biomass** fuel type in GCAM, we use the JEDI’s **Biopower** model (***01d-jedi-biopower-model-rel-b12-23-16.xlsm***), and the associated Python script is **jedi\_biopower*.ipynb***. Output is saved in ***EF/JEDI\_bioenergy.csv***.

We modify the JEDI’s Biopower model following Tables S15, S16, and S17 from Xie et al. (2023)3 for the **biomass-CCS**fuel type in GCAM (***01d-jedi-biopower-CCS-Xie.xlsm***), and run it with **jedi\_biopower\_CCS*.ipynb***. Note that in Table S17, per kW cost modifications in JEDI are implemented by adjusting the total cost/project size, as the per kW cells in JEDI are locked and cannot be modified directly. Output is ***EF/JEDI\_bioenergy\_CCS.csv***.

### 2.3.2 Coal and coal-CCS

For **coal** fuel type in GCAM, we run JEDI’s **Coal** model (***01d-jedi-coal-model-rel-c12-23-16.xlsm***), via Python script **jedi\_coal*.ipynb***, and produce outputs ***EF/JEDI\_coal.csv***.

The **coal-CCS** model (***01d-jedi-coal-CCS-Xie.xlsm***) is a modified version of the coal model, incorporating adjustments based on Tables S12, S13, and S14 from Xie et al. (2023)3. The associated Python script is ***jedi\_coal\_CCS.ipynb***, and output is ***EF/JEDI\_coal\_CCS.csv***.

### 2.3.3 Natural gas and natural gas-CCS

For **gas** fuel type in GCAM, we run JEDI’s **Natural Gas** model (***01d-jedi-ngas-model\_rel-ng4-17-17.xlsm***), via Python script is **jedi\_natural\_gas*.ipynb*** and produce output ***EF/JEDI\_natural\_gas.csv***.

The **gas-CCS** model (***01d-jedi-ngas-CCS-Xie.xlsm***) is a modified version of the natural gas model, incorporating adjustments from Tables S9 and S10 from Xie et al. (2023)3. Note that in Table S10, the “Variable Subtotal” is listed as 6.19 $/MWh, while when inputting costs for "Routine Turbine Maintenance," "Water," and "Catalysts & Chemicals" into the JEDI natural gas model, the internal formula calculates a subtotal of $5.89/MWh. We maintain the $5.89 value for the "Variable Subtotal" in our modified gas-CCS model. The associated Python script is ***jedi\_natural\_gas\_CCS.ipynb***, and output is ***EF/JEDI\_natural\_gas\_CCS.csv***.

### 2.3.4 Refined liquids and Refined liquids-CCS

The **refined liquids** model (***01d-jedi-refined-liquids-Xie.xlsm***) is based on the JEDI’s coal model, with a heat rate of 10767 Btu/kWh following Xie et al. (2023)3. Output is ***EF/JEDI\_refined\_liquids.csv***, obtained through ***jedi\_refined\_liquids.ipynb***. The **refined liquids-CCS** model is identical to the coal-CCS model, and the outcome is ***EF/JEDI\_refined\_liquids\_CCS.csv***, generated with ***jedi\_coal\_CCS.ipynb***.

### 2.3.5 Solar and rooftop PV

For concentrated solar power (**CSP)**, we run JEDI’s CSP model (***01d-jedi-csp-trough-model-rel-csp12-23-16.xlsm***) via**jedi\_csp*.ipynb***, and the outcome is ***EF/JEDI\_csp.csv***.

For photovoltaic (PV) and rooftop PV, we use JEDI’s PV model (***jedi-pv-model.xlsm***), which includes three system types: utility, commercial, and residential. The utility system is used for **PV**, the outcome is ***EF/JEDI\_pv\_utility.csv***, obtained with**jedi\_pv\_utility*.ipynb***. The residential system and commercial system are used for **rootop\_pv**. Outcomes are ***EF/JEDI\_pv\_residential.csv*** and ***EF/JEDI\_pv\_commercial.csv***, obtained with**jedi\_pv\_residential*.ipynb*** and **jedi\_pv\_commercial*.ipynb***, respectively.

### 2.3.6 Onshore wind and offshore wind

JEDI’s land-based wind model (**JEDI Beta Windows/ jedi-lbw-model-w2020.xlsm**) is used for **wind** model. The outcome is ***EF/JEDI\_onshore\_wind.csv***, obtained with**jedi\_onshore\_wind*.ipynb***.

JEDI’s offshore wind model (**JEDI\_installer\_Windows/ jedi-osw-model-rel.2021-3.xlsm**) is used for **offshore** **wind** model. The outcome is ***EF/JEDI\_offshore\_wind.csv***, obtained with**jedi\_offshore\_wind*.ipynb***.

### 2.3.7 Nuclear

The **nuclear** model (***01d-jedi-nuclear-model-Xie.xlsm***) is based on the JEDI’s coal model, following Tables S3, S4, and S5 in Xie et al (2023)3. Output is ***EF/JEDI\_nuclear.csv***, obtained through ***jedi\_nuclear.ipynb***.

### 2.3.8 Hydro

The **hydro** model is based on the JEDI’s conventional hydro model (***01d-jedi-chydro-model-rel-ch12-23-16.xlsm***). The Existing Plant Expansion/Upgrade system is specified in the project data sheet, as in GCAM-USA v7.1, hydropower capacity is assumed to be constant over time, with no addition and retirement throughout the model years. Output is ***EF/JEDI\_hydro.csv***, obtained through ***jedi\_hydro.ipynb***.

### 2.3.9 Geothermal

The **geothermal** model is based on JEDI’s conventional hydro model (***01d-jedi-chydro-model-rel-ch12-23-16.xlsm***). The Existing Plant Expansion/Upgrade system is specified in the project data sheet, as in GCAM-USA v7.1, hydropower capacity is assumed to be constant over time, with no capacity addition and capacity retirement throughout the model years. Output is ***EF/JEDI\_geo.csv***, obtained through ***jedi\_geo.ipynb***.

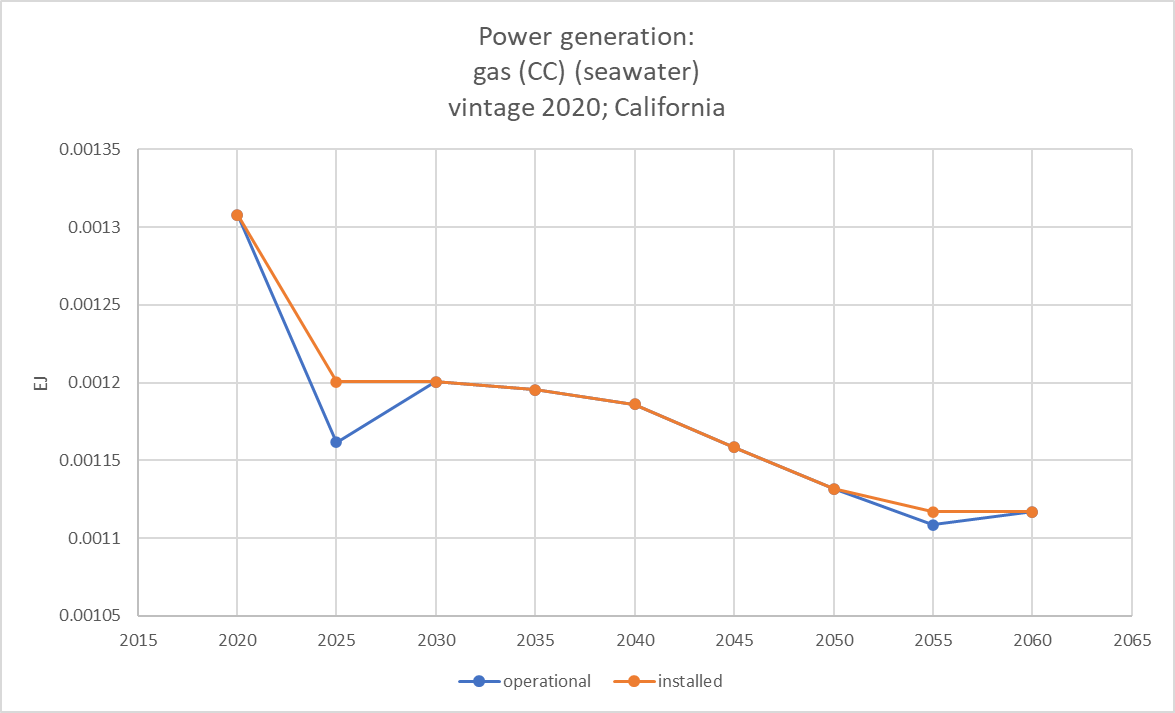
# 3. Processes

## 3.1 Power production calculation

GCAM-USA provides power generation outcome , based on which we calculate the power generation associated with different activities .

* Power generation from operational capacity:
* Maximum power generation potential from installed capacity:
* Power generation from newly installed capacity:
* Power generation loss from retired capacity:

The observed is the power generation from the operational capacity in year t, for fuel i, technology k, and vintage v. We calculate the power generation potential (i.e., generation possible with installed capacity) in period t, , as the maximum of in all future periods forwards, i.e., . For example, the figure below presents the power generation of operational capacity and installed capacity for a specific gas technology for vintage 2020 over time.



Power generation of the newly installed capacity is . The inter-period difference in the power generation of installed capacity by vintage denotes the power generation loss due to capacity retirement.

While most fuels and technologies can use the query ***”elec gen by gen tech and cooling tech (incl cogen)”*** to obtain the , wind and solar power need extra caution. Wind and solar require backup power in GCAM USA v7.1 because their energy generation is intermittent, depending on weather conditions and daylight. Backup power is meant to ensure a stable electricity supply when wind or sunlight is insufficient. As backup power is not generated by wind and solar capacities themselves, it is excluded from the capacity and job calculation for wind power and solar power.

The query ***“elec energy input by elec gen tech”*** provides the inputs for wind and solar power, including “distributed\_solar,” “PV\_resource,” “onshore wind resource,” “offshore wind resource,” “CSP\_resource,” “csp\_backup,” and “backup electricity.” In GCAM-USA v7.1, the input-output ratio is 1:1 for wind and solar technologies, so aggregating energy input levels, excluding “backup electricity,” provides accurate wind and solar power generation levels; these are the generation levels used to calculate wind and solar capacity in this package. Note that “csp\_backup” is treated differently from “backup\_electricity.” The former is a combustion turbine within the CSP facility, so electricity generated with “csp\_backup” input is considered solar power.

Rooftop PV in GCAM-USA is not currently vintaged, requiring post-processing to establish the vintage structure based on observed power output and retirement assumptions. Refer to <https://github.com/JGCRI/GCAMUSAJobs/blob/paper/docs/tutorials/vintage.xlsx> for an example.

## 3.2 Capacity activity calculation

Since GCAM-USA does not track capacity information directly, we can derive the capacity level for different activities based on power generation outcomes by activities and the capacity factor :

* Operational capacity:
* Installed capacity:
* newly installed capacity:
* retired capacity:

### 3.2.1 Operational capacity and installed capacity

As we expect fixed O&M jobs to be associated with installed capacity, and variable O&M jobs to be associated with operational capacity, we need to distinguish installed from operational capacity. The installed capacity represents the theoretical upper limit of operational capacity . Due to the possibility of profit-driven temporary shutdown, operational capacity for a given vintage may be less than installed capacity (i.e., ). We consider temporary shutdowns as capacity underutilization instead of capacity retirements.

These two measures of generation then provide their associated capacities:

, and

The installed capacity of a period consists of newly installed capacity and capacity built in previous periods that have not retired yet. Power generation from newly installed capacity is , and we can calculate the newly installed capacity .

### 3.2.2 Capacity retirement

denotes the capacity loss due to retirement. In GCAM-USA, capacity retirement consists of

* Natural retirement : Capacity built in or before 2015 is considered vintage 2015. An S-curve assumption is applied for the natural retirement of this vintage (). For capacities installed from 2020 onwards, retirement occurs when they reach their lifetime.
* Pre-mature retirement : Occurs when the power generation of a technology is no longer profitable. When total retirement exceeds natural retirement, the excess is attributed to premature retirement.

### 3.2.3 Simultaneous capacity addition and pre-mature retirement

GCAM-USA's modeling structure allows capacity additions and premature retirements to occur simultaneously for a given technology in period t. This structure is effective for regions with multiple facilities of a given technology in a given year. However, this approach may be less practical for small regions with only one facility of the technology.

For larger states, such as Texas and California, it would be reasonable to adopt a “Total” method where capacity additions and pre-mature retirements of a given technology can happen in the same period. For smaller states, such as Rhode Island, it would be reasonable to adopt a “Net” method where capacity addition and pre-mature retirement net out each other.

Specifically, with the “Total” method, and are used for job calculation, while with the “Net” method, and are used for job calculation, where

= + max(0, )

= max(0, )

## 3.3 Job calculation

In this version of the package, we focus on direct jobs in power generation, including onsite construction jobs, construction-related jobs (i.e. design and management), operation & maintenance (O&M) jobs, and decommission jobs. Construction jobs are associated with capacity additions, while decommission jobs are tied to capacity retirements. Operational capacity is linked to variable O&M jobs, and installed capacity corresponds to fixed O&M jobs. Therefore, to calculate the job outcomes, we apply the construction employment factor (EF) to capacity additions, decommission EF to capacity retirement, variable O&M EF to operational capacity, and fixed O&M EF to installed capacity.

### 3.3.1 Adjustment for O&M jobs

JEDI O&M employment factor (EF), , is capacity-based, including both fixed and variable O&M jobs. There are a few adjustments we need to make to calculate the O&M jobs.

Firstly, GCAM allows for temporary profit-driven shutdowns, during which fixed O&M jobs remain while variable O&M jobs are reduced. Therefore, we apply variable O&M EF to the operational capacity and fixed O&M EF to the installed capacity. We use GCAM-USA's assumptions on the variable () and fixed () O&M costs (<https://jgcri.github.io/gcam-doc/inputs_supply.html#energy>) to separate JEDI’s O&M EF accordingly, assuming the cost share reflects labor share.

Secondly, GCAM-USA models multiple load segments (i.e., base load, peak load, subpeak load, etc.) for power generation, where non-base load capacities are designed to operate for only a small part of the year, reflected in lower capacity factors. We expect non-base load capacities to have a lower level of variable O&M jobs compared to the base load capacity. Unfortunately, JEDI’s EF remains unchanged across different capacity factors, so we scale down the variable O&M EF for non-base load capacities by using the ratio of non-base load to base-load capacity factors. Specifically,

for base load segment (s = base):

and for non-base load segments (s != base):

### 3.3.2 Annual construction job EF

JEDI’s construction jobs of a specified project type is measured in total full-time equivalents over the construction period, where a “full-time equivalent” is defined as one person working full-time for one year. For example, if it takes 2 years and 40 full-time-equivalents to build a project with a capacity level of one MW, it means 20 annual (full-time) jobs for 2 years. Define the construction period of a project as T, the JEDI reported construction jobs (full-time-equivalents) as N, and the project capacity as C, then the annual construction EF is simply N/(TC).

However, GCAM-USA runs every 5 years by default, so in some instances, adjustments need to be made. In the simple case of a construction period no greater than 5 years (T <= 5), then the annual construction EF is N / 5C jobs per MW. If T > 5, then the annual construction EF between t-4 and t is N / TC, and the rest of the construction jobs are assigned to the previous model period (from t-9 to t-5), with an EF of N\*(T-5) / (5TC).

For example, in the case of nuclear capacity construction, due to the absence of specific construction period data from JEDI, we assume an 8-year construction period based on NREL ATB data4. Define the annual construction EF for nuclear as γ (jobs/MW). If GCAM-USA suggests a 10 MW nuclear capacity addition in 2050, this capacity addition would create 10γ annual construction jobs from 2046 to 2050, and an additional 10 annual construction jobs from 2041 to 2045.

# 4. Output of the package

The outputs of this package consist of several data frames and plots. The function GCAM\_EJ() takes the GCAM-USA output as input and produces the average annual power generation output, disaggregated by state, fuel type, technology, and activity. The function GCAM\_GW() processes the output of GCAM\_EJ() to calculate the average annual capacity levels by state, fuel type, and technology for different activities such as operation, addition, and retirement, supporting both the “Total” and “Net” methods. The GCAM\_JOB() function then utilizes the output from GCAM\_GW() to generate the average annual job estimates, broken down by state, fuel type and job type. Users can select between the "Total" or "Net" method, with "Total" as the default.

For visualization, PLOT\_EF() produces a plot of the annual average employment factors (EF) during the project period, assuming that decommissioning projects for all fuel types span five years. The construction project period varies across fuel types, and O&M EF by default is on an annual basis during the lifetime of a facility. PLOT\_GW(), using the output from GCAM\_GW(), creates a plot of annual average capacity by activity and fuel between each GCAM model step year. The PLOT\_JOB() function visualizes the average annual direct jobs between each GCAM model step year, categorized by fuel and job types. PLOT\_JOB\_TYPE() plots the jobs by job type across all fuel types in a specified year. All of these plotting functions will produce results either for an individual state or for the US as a whole. Lastly, MAP\_JOB() offers a visual representation of state-level total power sector direct jobs for the chosen year. Further examples and detailed visualizations are available in the package vignette.

# 5. Future development

There are a few limitations of this package, and we propose potential development to improve the package's performance in the future.

First, this package currently only accounts for direct jobs related to power generation. In the future, we plan to expand the scope to include indirect jobs across the supply chain, such as employment in biomass production, fuel resource extraction, and so on.

Second, this package currently adopts a constant employment factor over time. A future version of the package can introduce exogenous assumptions for labor productivity growth over time, mainly driven by mechanization.

Third, this package currently solely adopts employment factors from the JEDI model, and a future version of the package can introduce alternative data sources for employment factors.

Reference

1. Hall, D., Hunt, R., Reeves, K. & Carroll, G. *Estimation of Economic Parameters of U.S. Hydropower Resources*. None, 1218138, 3750 https://www.osti.gov/servlets/purl/1218138/ (2003) doi:10.2172/1218138.

2. EIA. Nearly half of U.S. geothermal power capacity came online in the 1980s - U.S. Energy Information Administration (EIA). https://www.eia.gov/todayinenergy/detail.php?id=42036 (2019).

3. Xie, J. J., Martin, M., Rogelj, J. & Staffell, I. Distributional labour challenges and opportunities for decarbonizing the US power system. *Nat. Clim. Change* **13**, 1203–1212 (2023).

4. NREL. 2019 Annual Technology Baseline. https://atb.nrel.gov/electricity/2019.