**Solar and battery adoption model**

Home of the most recent model version:

[\\batgirl-nc.nexus.csiro.au\shared-data\Paul\Adoption model AEMO 2022b](file:///\\batgirl-nc.nexus.csiro.au\shared-data\Paul\Adoption%20model%20AEMO%202022b)

**Sets used**

**t** time, the time set uses multiple subsets mainly to avoid division by zero issues and move between historical and projected processes and reporting lengths

**Scen** scenarios, the code uses numbers and these are assigned to scenario names in the reporting code in later steps. Sub-sets are to isolate scenarios as some clients will ask for different numbers of scenarios.

**m** months, reporting for AEMO is monthly and so is input data although most modelling is done annually and monthly detail added later. It is necessary to have monthly subsets in the reporting phase

**s** states, state subsets isolate different state policies or other differences. Reporting files change this set structure to STATE and REGION which specifically relates to NEM and WEM reporting requirements

**name1**, list for retail price subcomponents

**cus** customer segment based on the type of CER owned and retail contract

**custtype** customer segmentation based on residential, commercial and industrial as well as residential size

**batparams** list of some battery parameters

**captype** lists the two types of ways that capacity may be measured – installations and power

**techname** list of solar and battery cost components

**sdf** solar and battery degradation factors by technology, impact on output, price impacts

**hhpy** list of half hourly segments in a year, subsets are created that identify periods ideal to charge, discharge or avoid these activities

**loadelements** list of battery, pv and load operational data types

**priceelecments** retail price data types

**driver** list of demographic and macroeconomic types that are subject to a growth rate over time

**demoname** list of demographic properties

**demoname2** list of vehicle types that are drawn from the demography file

**SECTOR** segments of solar PV customers based on size ranges of capacity installed, several subsets of these are used because they each have different growth rate and other factors applied to their projection

**cr** customer segments based on residential and commercial

**ct** customer segments based on residential and business

**units** ways for measuring capacity installed

**hr** list of half hourly segments in a year

**dnsp** list of each distribution network service provider, subsets are created that match each dnsp to their to state (there are multiple dnsps in some states)

**polygon** an old set for Australian spatial zones, not used

**resclusters** list of half hourly residential consumption profiles used in the model, subsets identify different sized customers and which DNSP they belong to

**comindclusters** list of half hourly commercial/industry consumption profiles used in the model, a comcluster subset selects four that are most representative of commercial customers

**hhday** list of half hourly segments in a day

**nce\_e** list of generation plant types, not used

**ndg\_a** list of embedded electricity generation plant types, not used

**zones** list of distribution zone substations, not used, subsets are created which match the substations to their DNSP and substations to state. Mainly used as “z” which is a smaller set of interest

**lga**, list of local government areas, not used, subsets are created to match lga to lga number and to state and to polygon

**ws** list of demographic weighting scenarios to be tested, no testing has been conducted recently

**wc** list of categories of demographic data that are weighted

**comp** set used to collect two data sources for comparison or validation purposes (e.g. historical and projected)

**age** list of possible vehicle ages, not used, includes some subsets for new vehicles and other

**SA2** list of ABS Statistical Area 2 spatial zones, not used, subsets are created that match these to states, zone substations, exclude SA2 outside of NEM and WEM

**postcode**, list of postcodes, subsets are used to match postcodes to SA@, to states, to zone substations and to exclude SA2 outside of NEM and WEM

**rs** regression set, list of historical known data to be used in the regression required for short term forecasting, these are matched via tables to the year and month

ps prediction set, list of future dates to be used in the fitted short term forecasting equation calculated by the regression, these are matched via another table to the year and month

**Step through the code**

The short term projection and long term projections are carried out in separate code or files. For the long term projection model the short term projections are part of the historical data input file.

After the long term projection model is completed there are several separate reporting files. Some areis for developing data necessary for Excel charting to check the outputs. Others are to create csv files for reporting the outputs to AEMO to match the templates they provide which specify that data should be placed vertically and labelled in a particular way.

The short term projection models could stay in GAMS and be provided as data if it the long term projection model is the focus. Similarly the reporting files could also stay in GAMS if only required for CSIRO use. This walk through covers everything up to but not including the final vertical reporting files.

**Short term projection model: small scale solar PV(<100kW)**

File: Regression\_2022.gms

This file produces monthly solar PV data for <100kW systems by residential and commercial, by number of installations and power capacity (kW) from 2014 until the end of the next year ahead which was December 2023 in this case. These dates need to be changed for future implementations but generally aiming for about an 18 month forecast (in this case it was July 2022 to December 2023)

The current code proceeds as follows:

* Some sets and subsets from the master list above are defined at the start of this file
* Historical solar PV capacity data is loaded from DataTest.xlsx. This file is provided by AEMO and only modified a little to make it easier to load including relabelling column headings to match the sets used in the model
* Because some historical data is revised the most up to date data for the last two years is loaded from CER\_PV\_POSTCODE\_RES\_BUS\_ACTUAL\_JAN\_2021-JULY\_2022.xlsx. this data is provided by AEMO
* The Scendat.xlsx file loads the differences from the linear regression that will be assigned to each scenario to provide differences between them and these differences also allow for non-linearity in the projections
* When historical data is not available from AEMO, data is loaded directly form CER files (CERPostcodeData2021.xlsx and CERPostcodeData2022.xlsx). This data is lacking in residential and business breakdown. This situation might occur if we are doing a projection without AEMO as client and we therefore only have public CER data available.
* To cope with the lack of breakdown residential and business in public CER data, the next set of code finds the historical share of residential and business in the last known year (2020) and applies that to 2021 and 2022.
* When AEMO has provided 2021 and 2022 residential and business breakdown the code simply writes over this previous procedure for all but the Northern Territory. This is because AEMO has tended to neglect to provide data for NT with the system there being outside their jurisdiction
* The historical data is converted into the right format for the regression by replacing the time and month with a regression set
* A linear least squares regression model is defined where the unknown variables are the objective value, the deviation from predicted and actual, the slope on the month, the slope on the year and the intercept.
* The equations are defined next. The first equation is the objective function to be minimised which is the sum of squared deviations of the observed from predicted in each historical month. The second equation calculates the deviation by subtracting the predicted from the observed
* The model is defined by these two equations
* A loop is set up so that this minimisation problem can be calculated for each postcode, capacity type and customer type by replacing the monthly data with one instance of these three dimensions each loop.
* The model is non-linear because of the squaring of unknown variable deviation. Therefore the code calls on a nonlinear solver (conopt) to solve the model through minimisation. A solver option file is assigned which sets a tolerance limit.
* Using the results of the regression for the intercept and slopes, the linear equation is solved for future explanatory data values (which is just months and years) and stored as predicted data
* Some postcodes with bad data might result in a declining slope. The code disallows this prediction and replaces it with a flat trend
* The predicted data is converted back to month and year format
* The predicted data is added via extension to the historical data set by creating a new parameter called AllData
* All data is saved to a gdx file for use in the next process

File: PVSize\_2022.gms

The data in the previous process produces historical capacity and installations. Average PV size per postcode in each month is defined as the monthly change in capacity in the postcode divided by the monthly change in installations. This can be calculated for historical data. In the full projection model we need a prediction of PV size out to 2060 based on historical trends with future uncertainty explored via scenario assumptions. This file creates that projection.

The current code proceeds as follows:

* Some sets and subsets from the master list above are defined at the start of this file
* The scenario file is loaded and this contains some assumptions that define the preferred growth rates of PV size by scenario. These are currently designed to asymptote at different levels and timing.
* The first set of code continues to deal with the predicted solar PV capacity and overwrites the scenario factors that were design the differentiate the predicted values by scenario. The code increases those differences and adds a higher than linear slope by virtue of multiple all by values higher than 1.
* Business is treated differently to residential. The quantity of installations of business gets the same treatment as residential. However, business capacity in 2022 gets only 20% of residential growth and in 2023 only 70% of residential growth. This would have been judged based on looking at the plausibility of business and residential trends and concluding that business capacity was not growing as fast.
* From the postcode data, some state, NEM and SWIS parameters are created by summing up the relevant postcodes.
* State level data is brought in in two tables on the average size in 2021 and on the historical growth rate in each state
* Some default levels for PV size are defined for residential and business
* PV size by postcode and scenario is calculated from the historical data (including the short term projection) via a loop starting with the first month. A loop is used because the code checks each year that the size is not falling. If it is falling it gets the default value. If it’s the same, it stays the same. If it’s growing it can keep the growth as the PV size.
* After the first month is calculated the other months are calculated in the same way. This can’t be done in one step because the previous month occurs in another year for month 1 but in the same year for months 2 to 12.
* Any residential greater than 15kW are assigned 15kW. Any commercial greater than 100kw are assigned 100kW. These should not exist because these maximums define the data but bad data can occur.
* The average monthly PV size growth rate in 2021 is calculated. If less than 1 then assigned 1 (i.e. no growth). This data is currently not used. Instead, the table of average size and growth is used in the next step. The state detail is needed for a better result.
* The 2021 average state PV size is assigned to 2022 and 2023 with assumed growth of 5.8% p.a.
* For business the approach is modified because it has less growth in PV size. The five year average postcode PV size is calculated and assigned to 2023
* This concludes the historical period (including the short term forecast). The long term PV size forecast follows.
* The default annual historical growth rate (hgr) is assigned 1 by default. It is then calculated based on 2021. Any greater than 2 (300%) are assigned 2.
* The historical PV size is added to the projected data for 2022 and 2023
* Using 2023 as the starting point, the hgr is applied to 2024 and beyond postcode data via a loop. The hgr is adjusted via a scenario factor to give difference answer for each scenario.
* The average state PV size in 2014 is calculated by weighting by postcode in each state.
* For the remainder of the period to 2023 the average state size is calculated as the difference from June to June each year in the capacity divided by the installations
* For the projection the state hgr (hgrs) adjusted by a scenario factor is used to project the state average PV size for the period 2024 to 2060
* SWIS and NEM average forecast PV sizes are calculated
* The historical solar PV capacity and installations are sent to one gdx file and written to Excel. The PV size data is sent to another gdx file and written to Excel

Actions required before proceeding to long erm projection model:

* Check output files PVRegResults.xlsx and PVSizeResults.xlsx to ensure the spread of one year projections suits the scenarios
* While the regression focusses on absolute or total numbers the results files also provides insight into the implied new installation rate and implied change in PV systems sizes. Even though the projected total installations and capacity may look fine it is necessary to check the installation rate and solar PV size changes for plausibility.
* Make adjustments to scenario factors or other growth rates as needed.

**Short term projection model: *Solar-large scale (>100kW but <30MW)***

Large-scale solar projection commences with updating the excel file LargeSolarPowerStations\_v11.xlsx. That file does the following:

* Loads the latest postcode level data on large-scale
* Sorts the data into the required size categories (defined by AEMO)
* Determines an average build rate

This average build rate is applied in a later uploaded file to the longer term projection model called LS\_soalr\_v2.gms. The inclusion file calculates the revenue ratio associated with each scenario. Builds proceed at the build rate if the revenue ratio is positive. Similar to the other short term forecasting models, the build rate can be faster than linear in some scenarios. This scenario build rate is modified through a parameter called “Mod”.

**Short term projection model: *Batteries***

Battery sales and capacity are projected at the state/territory level in a file called BatteryInstallations\_2022.xlsx. The projections are based on a simple regression of annual state level data. The historical data comes from a combination of SunWiz data and AEMO DERR data. The outputs of this file are number of sales and capacity added.

**Long term projection model: all solar PV and batteries**

To update the longer term projections for all solar PV sizes (small and large) and batteries this occurs in two files consecutively, firstly AdoptionMod\_Res\_v6.gms which activates AdoptionMod\_Com\_v4.gms (this second file then calls on some reporting files).

The current code proceeds as follows:

* Some sets and subsets from the master list above are defined at the start of this file
* The current structure of residential retail electricity price components by state are loaded via an in-code table. This comes from the AEMC’s retail electricity prices annual publication. These components make up the cost stack that is the retail electricity price. These are in c/kWh
* The generation component is converted to $/MWh
* NT and WA prices are adjusted to take account of state control of retail prices there
* National averages are calculated
* Historical and projected retail and generation prices based on input from AEMO are loaded from AEMO\_ELEC\_P\_working.xlsx
* Solar export prices are loaded from Yearly\_SolarPrices2022.xlsx (AEMO provided)
* Export prices after 2050 are assumed to equal 2050 export prices. Null data before 2020 is assigned 2020 prices
* The Generation price that was available from the AEMC table is replaced with the AEMO data
* scen2019 is not used and can be ignored
* ‘Speedrate’ data is defined. This is used to control the maximum rate of adoption over time should the adoption curve lead to something implausible. This was more of a problem when using unfiltered costs and prices that had sudden jumps – this could lead to rapid change in the payback and subsequently large change in adoption. Costs and price inputs have since been smoothed.
* The smart tariff share is assigned and this controls how much the share of customer bills that are based on smart tariffs contribute to the average payback versus flat tariffs
* State solar PV capacity factors are added in code in a table. AEMO provides half hourly data but these averages are also useful in the code and are calculated from the half hourly data
* The state solar PV capacity factors are assigned to postcodes and DNSPs
* A number of scalar parameters are defined in code
* AEMO\_solar\_data is loaded but can be ignored. It is no longer used.
* The file GALLM202021costs\_v3.xlsx loads the capital cost data for batteries (disaggregated by battery pack and balance of plant) small scale solar and large scale solar. Battery capital subsidies are also included in this file.
* The subsidies available for large-scale and small-scale solar in the renewable energy target to 2030 are loaded from the STC\_v3.xlsx file. Beyond 2030 these subsidies are extended by virtue of the safeguards mechanism and other offset schemes.
* The degradation.xlsx file loads the degradation rates for the physical capacity of solar PV and batteries as well as the degradation in export revenue received due to market congestion. This file also deals with estimated retirement due to age of assets through survival rate data and the net increase in capacity of installations owing to replacement of older installations
* The Monthly\_working\_v2.xlsx file loads the monthly installation trends by state and customer/size segment
* The DERR\_data\_2022.xlsx file loads the postcode location of battery installations in the AEMO DERR data
* The BatteryInstallations\_2022.xlsx loads the historical yearly, state battery installation data by scenario and customer type including a two year ahead forecast to 2024.
* The gdx files for the historical solar PV installations and capacity and PV size are loaded next. The solar PV data includes a 18month ahead monthly forecast at the postcode level by scenario. The PV size data includes PV size to 2060 at the state and postcode level by scenario.
* This solar PV historical data is assigned to a new parameter called RSCapacitym and the June data (end of financial year) aggregated up from postcode level to annual state parameter RSCapacityState. This allows some calculations to be done at the state and annual level.
* Some state and annual level PV sizes are calculated from the imported data
* The parameter BusModelVal is the saturation level in the logistic curve. Values are assigned here. The values have been calculated in a file called ‘Market Segment\_v6.xlsx’ which s not stored here but rather tin the project file. This content is developed from scenario assumptions specific to each project. The BusModelVal is slightly adjusted in some cases due to the difficulty of hitting the national level target using individual postcode level logistic curves (noting this problem has since been solved in the new implementation of the EV model code). The name of the parameter comes from the fact that the main constraint on reaching higher levels of adoption of solar and batteries is whether sufficiently attractive business models can be found that make it viable to put these technologies in rental and apartment buildings.
* Standard battery sizes (kWh) and peak power (kW) for small and large customer are assigned
* Some additional battery performance characteristics are loaded in batterydata2.xlsx. It appears that some battery cost data is loaded here as well but only used for some historical years where current and future year data comes from the GALLM file only.
* Residential and commercial half hourly consumption before battery or solar profiles are loaded from CustomerProfiles45\_wHVAC\_v5\_working\_fy.xlsx. tYhey are financial year beginning.
* The hourssetmap.xlsx file data is loaded that includes some large sets covering half hourly time match to hourly time, days, common half hourly daily segments, winter months and summer months.
* State half hourly solar production files are loaded. This is AEMO data for 2016-17. There is more recent data available but any year is probably fine for the broad purposes of calculating representative bills. Ideally we should probably use 2012-13 as I think that is the weather year for the consumption profiles.
* The factorGrowth.xlsx file loads annual growth rates for state GDP, customers, separate dwellings and home ownership by state.
* Missing customer profiles for essential energy, ergon energy and northern territory distribution zones are created by copying from other distribution zones with similar climate.
* State solar PV subsidies are downscaled to distribution zones and residential profiles by distribution zone.
* Some battery cost data is assigned to historical years
* Annualised solar, battery and electricity meter costs are calculated along with the levelised cost of solar
* State retail, generation and export prices are converted to distribution zones
* These distribution zone parameters are assigned to customers in those distribution zones

At this point LS\_soalr\_v2.gms is loaded which deals with solar PV installations greater than 100kW but less than or equal to 30MW. We call these large-scale solar in the explanations below, but they are only large relative to rooftop solar. We are not talking about 200-1000MW solar farms here.

The code proceeds as follows:

* LargeSolarPowerStations\_v11.xlsx is loaded and this is where short term projections of historical capacity and installations in the different size categories were made. Data imported in this file includes average build rates and time between builds, current capacity and installations by postcode, monthly installations and capacity by state from 2014 to 2024 by size segment. Finaly there are factors included which used to turn this historical and linear projected data into 5 different scenarios
* The scenario factor is assigned to all sizes.
* Some redundant code is texted out.
* Some parameters are defined.
* Different build rates are assigned to different scenarios through a parameter called Mod(scen).
* The three different paths of prices received for renewable generation under the large-scale renewable energy target (LRET) are assigned to different scenarios.
* The levelised cost of large-scale solar PV is calculated increasing the capacity factor to recognise that large-scale solar has a higher capacity factor.
* The ratio of revenue to cost of generation is calculated
* AEMO provides a postcode file that ensure only the right postcodes in the NEM and WEM are associated with each state. This code removes postcodes that don’t fit that data.
* The scenario factor (UncFac) is applied to differentiate values in 2024 for the capacity and installation of large-scale solar by scenario. 2023 is also calculated so that the rate of new installation between 2024 and 2023 can be calculated at a later step
* Installations per kW in 2024 are calculated.
* The rate of new installation between 2024 and 2023 is calculated.
* A loop is created for the smaller sized large-scale solar which adds new installations and new capacity to years where the ratio of revenue to cost of generation is greater than 1.1. Capacity is added at the historical build rate modified by a scenario factor Mod(scen). Installations are then determined by the previously calculated installations per kW
* A further loop follows that is set up such that the larger sized large-scale also deploy new capacity but only if the revenue ratio is sustained for 5 years. The installations are determined such that the installations must fit within the size category. That is in the 5MW to 10MW category, each installation is on average 7.5MW
* A further loop that follows requires states that have not previously deployed the larger large-scale solar to sustain a revenue ratio of 1.1 for seven years before they can deploy the larger large-sale solar categories
* Scenario 1 is assigned 80% of the installations and capacity of scenario 2 if the projected values are lower than that.
* Historically, some solar sizes are not deployed each year. These loops zero out year in which they cannot be deployed to match the averaged observed space between years ranging from none to up to 4 years. Half the states are assigned to deploy them one year after the other half so that the national trend is smoother and not all installation occur in the same years
* Parameters which capture cumulative installations and capacity (no degradation) are assigned ‘historical’ installations and capacity with some forecast years modified by the scenario factor (UncFac) to create differentiation between scenarios.
* The remainder forecast installations and capacity are added via loop which adds the new installations and capacity over time.
* A degradation rate of 1% is added to projected capacity
* Some older code that sort to smooth the joining of historical and projected data has been texted out
* The generation from this capacity is calculated but not used elsewhere

Back to AdoptionMod\_Res\_v6.gms

The code continues as follows:

* An energy efficiency parameter is assigned but not used (by virtue of assigning a value of 1)
* Some distribution one level hourly solar production data is loaded from a gdx file and converted to half hourly. However, this data is not used. The previously imported state SolarProd data is used and assigned to each distribution customer segment.
* kW consumption data is converted to kWh
* Half hourly grid consumption per residential customer by scenario and year is calculated as consumption minus solar production which is a function of the solar profile per kw multiplied by the solar system size in kWs. If negative it is assigned zero.
* Half hourly exports per residential customer by scenario and year is calculated as solar production minus consumption
* Customer peak demand before or with no solar is calculated as the maximum of annual half hourly consumption times 2 (converting kwh to kW)
* Customer peak demand with solar is calculated as the maximum of annual half hourly grid consumption times 2 (converting kwh to kW)
* Battery operation is simulated under a flat tariff with a simple solar shift algorithm as follows:
  + The battery starts half full
  + If solar production is greater than consumption and battery is not full
    - And if the available solar is greater than maximum charge, assign maximum charge to that half hour and add that to the state of charge
    - Else if what’s available from solar is less than a full charge then do the maximum charge available and what’s available to the state of charge
  + Else if solar production is less than consumption
    - And if the gap is greater than maximum discharge, give maximum discharge and take that from state of charge
    - Else if the gap is less than maximum discharge, give the smaller amount and take that from state of charge
  + Else if neither applies then don’t discharge or discharge
  + Update starting battery level for next loop from current state of charge
* Texted code is used to gather one week of data for checking the results of the above algorithm
* Residential grid consumption after storage is installed is calculated as grid consumption before storage minus battery discharge
* Residential exports after storage is installed is calculated as exports before storage minus battery charging
* The residential bill for a household with no solar and no storage is calculated as the retail daily charge times days in a year plus consumption times the consumption charge (divided by 100 to convert cents per kwh to $ per kWh)
* A retail price index is created relative to the year 2015
* An alternative type of retail bill includes a capacity charge whereby there is a $/kW charge based on peak demand. The formula for the capacity charge is constructed such that it results in the same electricity bill as the energy based bill just calculated. The capacity charges are assigned to each distribution zone customer segment. However, this capacity charge is no longer used in the code that follows. It was decided to put greater emphasis on energy based bills but with a rebate provided for customers who avoid peak demand (potentially through a VPP arrangement or TOU tariff)
* Summer and winter average daily consumption profiles are calculated for reporting purposes.
* Various parameters needed to calibrate the adoption curve as a function of payback are introduced
* Payback over time for household installing only solar is calculated as the cost of solar after subsidies plus the cost of the smart meter divided by the bill before solar minus the bill after solar. The bill after solar is the fixed daily charge times days per year plus grid consumption times the consumption charge minus export revenue. A factor called APF remains a part of this formula but has fallen out of use since it has identical values across all scenarios. It was originally designed to capture periods of heightened changes in the affordability of electricity.
* The payback for a household with both solar and storage is calculated in the same way except that the with storage values for grid consumption and exports are used
* There is a provision for scenario 1 to be subject to a different calculation procedure however this is not currently utilised.
* The battery operation algorithm is run again but this time for a household with an incentive to avoid peak demand charging and provide peak demand discharging. The algorithm is similar to the previous one except that charging can only occur at certain charge times to avoid the peak. Furthermore, discharging can only occur at preferred discharging times which align with peak demand.
* Texted code is used to gather one week of data for checking the results of the above algorithm
* Grid consumption and exports with storage is recalculated with the new charge and discharge results.
* A second payback over time is calculated for with solar and storage, this time with the re-calculated grid consumption and exports as well as a rebate payment (std – smart tariff discount).
* New parameters are introduced to assigned payback data into different spatial configurations. Several set matched conversions are gone through to arrive at a postcode level payback parameter.
* Income, separate dwellings and home ownership by postcode are loaded from ABS\_income\_pc.xlsx and ABS\_HouseStats\_pc.xlsx
* Scalar data of the ideal and minimum values for the above data are defined
* Define some other scalars to be calculated
* Average capacity per installation is calculated and used to calculate capacity form installations and then generation but none of this is used or reported and can be ignored.
* A minimum of 10 houses is assigned to postcodes with zero houses to avoid division by zero errors.
* The share of installations in each postcode is calculated.
* The maximum and minimum solar share by postcode are calculated but these are just for interest and not used anywhere
* If the solar share is greater than 75% then the historical capacity in the postcode is adjusted to be 75%. A minimum of 0.001 is also assigned. It would be better if the historical installations were not modified but rather just overwrite solar share.
* Some additional parameters are introduced associated with calculating the postcode score by which the targeted maximum adoption level is adjusted to reflect the demographics in each postcode
* State growth rates for different demographic drivers are assigned to postcodes in that state
* The weights used to assign more or less positive effects of demographic characteristics in calculating the total score are imported
* The score is calculated over time as the sum of different demographic factors multiplied by their weights and normalised somewhat by dividing by the maximum range of values.
* The maximum and minimum scores for 2023 are calculated for interest’s sake. This not used elsewhere.
* The postcode score is forced to fall between -0.25 and 0.25 via applying a formula. The values for 2023 and 2050 are captured for interest’s sake but not used anywhere
* The formula for calculating a two point calibration of a logistic curve function is implemented in a number of steps that are broken down into various “intermediateVar” parameters.
* The maximum saturation point of the logistic curve we call f1 but calculate as intermediateVar5 is calculated as the addition of BusModValue (the national target) and the postcode score that was forced between +/-0.25. This means that each postcode meets the national target but varied by the postcode score. Postcode values greater than 0.999 or less than 0.001 are ruled out.
* Other parameters in the logistics curve, alpha and beta, are calculated using formulas which include the current installation share, the maximum saturation point, current payback. The future level of payback at which maximum saturation occurs is also a feature of these formulas but is often assigned as 1 year. The value of 1 falls out of the formula in some cases.
* The payback over time (PaybackRS) is entered into the formula and the adoption rate for all future years, postcode and scenarios is stored in the parameter ‘Adoption’.
* There are some texted out controls on the Adoption rate. These have been used less frequently as the calibration of historical data has improved.
* A fall in adoption is ruled out. This can only happen if the payback period increase. This is possible as the government subsidy for solar PV is designed to decline over time.
* A maximum speed of adoption is applied in two different ways.
* The largely texted out code around projected sales and sales comparison dates back from when the model was being tested for how well is calibrated against historical deployment and is not used in the adoption projection.
* The code eliminates parts of Australia that AEMO is not interested in such as northern WA and Mount Isa in Queensland that is not part of the NEM. For clients who want the whole of Australia modelled these restrictions need to be texted out.
* The projected solar PV installations are calculated by multiplying adoption by the number of households and the assumed customer growth over time.
* The projection is recalibrated so that June 2023 exactly matches history in all postcodes.
* The projected solar PV capacity in kW is calculated in a loop by staring with historical data then adding the projected annual change in installations multiplied by the assumed annual PV size
* The number of households over time is captured for reporting purposes
* Production of electricity from solar PV is calculated but not used anywhere
* A degradation factor by year is calculated from the known installation rates from the beginning of deployment in the early 2010s. This is applied to historical installations in the reporting file
* Extra installations are added on the basis that there is a 1.75% replacement rate due to storm damage etc
* The battery projection model now starts and commences with the BusModVals or maximum saturation rates that will be used for batteries.
* Various parameters are defined that will be used in the battery projection steps.
* Cumulative state historical battery installations are calculated from historical sales.
* Battery installations in postcodes not related to the NEM or WEM are eliminated
* The available postcode data on battery installations is used the calculate a postcode share of each state.
* This postcode share of state is then used to create historical postcode sales data by applying that share to the historical state sales data.
* The installation share is then calculated as the share of sales divided by households. This should be changed. We have the installation share from the cumulative sales and that would be more consistent with the method.
* Installation shares less than 0.001 and more than 0.75 are ruled out.
* The logistic curve formula is applied in the same way that it was for solar PV except that we use the battery installation share and the payback period for solar PV with a battery.
* The Adoption rate by postcode is calculated from the logistic curve formula applied to future payback rates and modified in that it can’t be lower than a pervious year and can’t increase from year to year greater than certain values.
* Some largely texted out code calculates sales for model testing but these are not used.
* Battery installations in postcodes not related to the NEM or WEM are eliminated
* Historical and future battery sizes are assigned.
* The projected battery installations are calculated by multiplying adoption by the number of households and the assumed customer growth over time.
* The projection is recalibrated so that June 2024 exactly matches history in all postcodes.
* Texted out code applies state subsidy policies, modifying the adoption so that it meets those policies. These are texted out because most of those policies have been removed over time.
* The capacity in giga watt hours is calculated from the installations and battery size
* State level data is calculated from the aggregated postcodes in the state.

At this point the commercial solar PV and battery projection model is loaded. The code proceeds as follows:

* Parameters are defined which are similar to the residential model but with ‘C’ added to denote commercial
* Historical commercial postcode installation and capacity data is assigned from ‘AllData’ to a new parameter name.
* A state level version of this data is assigned by summing up the relevant postcodes. If equal to zero, at least 1 installation and 6kW is assigned to historical years.
* Average commercial solar PV size is calculated from the installation and capacity data. A non-state or time based value uses the 2022 NSW value.
* The costs of solar PV and battery systems are annualised and the LCOE for commercial solar PV calculated.
* More parameters are defined.
* Commercial consumption profiles are assigned to a new parameter name
* Solar profiles from distribution customer zones are assigned to distribution zones
* Commercial solar PV export prices are taken from those used in Ausgrid distribution zone for the residential model
* Commercial retail prices are taken from residential retail prices (for small businesses they are often not very different but a factor of 0.79 is applied later when used in bill calculations)
* Energy efficiency changes are not applied
* The commercial consumption profile data is scaled up to be consistent with observed solar system sizes
* To save repetition the calculations for grid consumption before and after batteries, exports before and after batteries, battery charge and discharge with and without peak avoidance, commercial bills with and without solar PV and batteries and payback period commercial bills with and without solar PV and batteries proceed in the same way as explained in the residential model
* The payback period by customer in each distribution zone is then converted through various steps to geta to postcode level.
* The historical share of installations in each postcode is calculated based on historical installations and, unfortunately, households even those this is the commercial model. This decision was likely made on the basis that businesses per postcode data were not available. This approach could be changed if the required number of business premises data can be found.
* BusModVal data is defined which is the maximum saturation rate.
* The same formula for postcode score and the logistic curve are applied as was in the residential model. The residential postcode score is used due to lack of business data. The normalisation of that score is changed to be in the range of +/- 0.15 reflecting that longer term range is assumed to be narrower. Commercial installation shares, maximum saturation rates and payback rates are applied to the formulas.
* Future Adoption is calculated using the resulting logistic curve parameters and some maximum changes in the adoption rate are imposed on that projection.
* Some additional parameters are defined.
* Postcodes that are not in the NEM or WEM are eliminated
* Installations are calculated as the projected adoption ate multiplied by households adjusted by growth in customers
* The result is recalibrated so that is exactly reproduces June 2023.
* The capacity in kW is calculated from the change installations and assumed commercial solar PV sizes over time.
* The degradation factor over time is calculated from historical and projected installations. This is applied to historical installations in the reporting file.
* The replacement rate is calculated at 1.75% for losses due to damage and added to the projected installations
* Electricity production from solar PV is calculated but not used anywhere.
* Replacement sales and total installations are assigned to the state level by summing up the relevant postcodes.
* The commercia battery model commences.
* By this point there are no new code elements compared to the residential battery model and so will save on repetition
* At the end of the commercial battery model the share of batteries in solar installations is calculated for both residential and commercial installations.
* Average battery charging and discharging behaviour is captured for winter or summer with the other texted out.
* The postcode installation and capacity levels are sent to a gdx file. These are used later in postcode level reporting.
* The 2023 and 2050 market saturation levels are captured and sent to a gdx and Excel file for checking and reporting purposes.
* Projected adoption, payback and other reporting variable are sent to a dgx and Excel file.
* A larger reporting file is called.

Reporting\_template\_v6.gms is now called and it is the main reporting file for charts used in the AEMO reports and also prepares some parameters which are used in the vertical csv reporting that AEMO requires. This file adds degradation of battery and solar capacity. The code proceeds as follows:

* Some desired reporting parameters are defined
* Historical data is loaded and renamed splitting out installations and kW capacity
* Postcodes not in the NEM or WEM are removed from the historical data
* Some olde code adjusting installations for an older Victorian subsidy policy is texted out.
* The historical postcode data is loaded into a state level parameter by summing the appropriate postcodes.
* It is necessary to recalibrate the projections because the historical levels have been modified by zeroing out the non-NEM and non-WEM postcodes.
* The historical data is added to the parameter which hold the projections
* The projection model was slackening off the recent high installation rate too quickly and so a method was introduced to make installations in 2024 higher and ensure these additional installations were carried through to all future years.
* A method is imposed to ensure that Victoria has at least 70,000 solar PV installations per year to reflect targets there.
* An older method for adjusting projections is texted out.
* Some new parameters are assigned residential and commercial installations.
* Residential and commercial capacity in kW are calculated from the change in installations and solar PV size data. Degradation is also applied at this stage.
* Replacement capacity is calculated from replacement sales time size
* Some texted out code calculates cumulative replacement capacity
* Cumulative replacement sales are calculated
* The installations and capacities of large-scale solar are disaggregated into individual parameters.
* Battery installations and capacities in MW and MWh are assigned to parameters from parameters which hold the historical and projected data.
* The historical data doesn’t break batteries down by size. So we need a method to assign a ratio of small and large batteries to historical and projected commercial batteries. We do that by calculating the ratio of large to small commercial solar. Then we create the large and small battery categories.
* The state level annual data is output into gdx and Excel files
* Some reporting of battery operation data is texted out
* The remaining code is about creating monthly data. We start by defining new parameter names that will hold the monthly data.
* Monthly historical data for commercial solar PV by size category is added first since it is available.
* There is some code here that works on batteries in months 1 to 5 but is is written over below and redundant.
* We assign all annual data to month 6.
* The small scale solar monthly data is assigned using a monthly share. The months 7-12 are assigned first since month 6 is now known and can be built upon.
* Now month 12 is known, months 1-5 can be calculated since month 1 need sot build on month 12 in the previous year.
* Commercial solar and commercial and residential solar is completed in the same way as residential solar.
* The monthly data is saved to a gdx and Excel file

Reporting\_template\_v6Nodeg.gms is then called and repeats all the same steps above but does not add degradation. This non-degradation data is required because AEMO requests that both degraded and undegraded capacity be provided. Note, that because the number of installations is never degraded, those steps are not repeated in this code.

The vertical reporting files are not called but are run separately. They are required by AEMO and they are at the state/territory level and at the postcode level. Due to the size of the data produced, small and large batteries and solar have their own reporting files for both spatial levels. The vertical reporting files:

* Define new sets for the specific labels that AEMO requires
* Defines the reporting parameters
* Aggregates the results in to single parameters with multiple sets as labels
* Outputs vertical reports as a single parameter in a csv file format

The files are:

State level

* Vert\_Reporting\_template\_pv2.gms
* Vert\_Reporting\_template\_batt4.gms

Postcode level

* Vert\_Reporting\_template\_pcode\_PV\_v6.gms (small scale)
* Vert\_Reporting\_template\_pcode\_PVNSG\_v6.gms (large scale)
* Vert\_Reporting\_template\_pcode\_Batt\_v4.gms

No code walk-through for these is completed yet. They could stay in GAMS potentially or be completed after the above code is completed. They only require data that is already produced from the above steps plus the addition of some new sets to create AEMO’s preferred data labelling. Then the code just aggregates all the dimensions into a single parameter. There is some rounding and controlling of decimal points.