GenX Database Documentation

1 Model setup parameters

Model settings parameters are specified in a `genx_settings.yml` file which should be located in the current working directory (or to specify an alternative location, edit the `settings_path` variable in your `Run.jl` file). Settings include those related to model structure, solution strategy and outputs, policy constraints, and others. Model structure related settings parameters affect the formulation of the model constraints and objective function. Computational performance related parameters affect the accuracy of the solution. Policy related parameters specify the policy type and policy goal. Network related parameters specify settings related to transmission network expansion and losses. Note that all settings parameters are case sensitive.

Table 1a: Summary of the Model settings parameters

Settings Parameter	Description
Model structure related	
TimeDomainReduction	1 = Use time domain reduced inputs available in the folder with the name defined by settings parameter `TimeDomainReductionFolder`. If such a folder does not exist or it is empty, time domain reduction will reduce the input data and save the results there.
	0 = Use the data in the main case folder; do not perform clustering.
TimeDomainReductionFolder	Name of the folder where time domain reduced input data is stored.
UCommit	Select technical resolution of of modeling thermal generators.
	0 = no unit commitment.
	1 = unit commitment with integer clustering.
	2 = unit commitment with linearized clustering.
NetworkExpansion	Flag for activating or deactivating inter-regional transmission expansion.
	1 = active

	0 = modeling single zone or for multi-zone problems in which inter regional transmission expansion is not allowed.
Trans_Loss_Segments	Number of segments to use in piece-wise linear approximation of losses.
	1: linear
	>=2: piece-wise quadratic
Reserves	Flag for modeling operating reserves .
	0 = No operating reserves considered.
	1 = Consider regulation (primary) and spinning (secondary) reserves.
StorageLosses	Flag to account for storage related losses.
	0 = VRE and CO2 constraints DO NOT account for energy lost.
	1 = constraints account for energy lost.
Policy related	
EnergyShareRequirement	Flag for specifying regional renewable portfolio standard (RPS) and clean energy standard policy (CES) related constraints.
	Default = 0 (No RPS or CES constraints).
	1 = activate energy share requirement related constraints.
CO2Cap	Flag for specifying the type of CO2 emission limit constraint.
	0 = no CO2 emission limit
	1 = mass-based emission limit constraint
	2 = load + rate-based emission limit constraint
	3 = generation + rate-based emission limit constraint
CapacityReserveMargin	Flag for Capacity Reserve Margin constraints.
	Default = 0 (No Capacity Reserve Margin constraints)
	1 = activate Capacity Reserve Margin related constraints

MinCapReq	Minimum technology carve out requirement constraints.
	1 = if one or more minimum technology capacity constraints are specified
	0 = otherwise
MaxCapReq	Maximum system-wide technology capacity limit constraints.
	1 = if one or more maximum technology capacity constraints are specified
	0 = otherwise
Solution strategy and outputs	
Solver	Specifies the solver name (This is not case sensitive i.e. CPLEX/cplex, Gurobi/gurobi, Clp/clp indicate the same solvers, respectively).
ParameterScale	Flag to turn on parameter scaling wherein load, capacity and power variables defined in GW rather than MW. This flag aides in improving the computational performance of the model.
	1 = Scaling is activated.
	0 = Scaling is not activated.
ModelingToGenerateAlternatives	Modeling to Generate Alternative Algorithm. For details, see here
	1 = Use the algorithm.
	0 = Do not use the algorithm.
ModelingtoGenerateAlternativeSlack	value used to define the maximum deviation from the least-cost solution as a part of Modeling to Generate Alternative Algorithm. Can take any real value between 0 and 1.
WriteShadowPrices	Get the optimal values of dual variables of various model related constraints, including to estimate electricity prices, stored value of energy and the marginal CO2 prices.
MultiStage	Model multiple planning stages
	1 = Model multiple planning stages as specified in `multi_stage_settings.yml`
	0 = Model single planning stage
MethodofMorris	Method of Morris algorithm

	1 = Use the algorithm.
	0 = Do not use the algorithm.
Miscellaneous	
PrintModel	Flag for printing the model equations as .lp file.
	1 = including the model equation as an output
	0 = the model equation won't be included as an output

Additionally, Solver related settings parameters are specified in the appropriate .yml file (e.g. `gurobi_settings.yml` or `cplex_settings.yml`), which should be located in the current working directory. Note that GenX supplies default settings for most solver settings in the various solver-specific functions found in the `src/configure_solver/` directory. To overwrite default settings, you can specify the below Solver specific settings. Settings are specific to each solver.

Table 1b: Summary of the Solver settings parameters

Settings Parameter	Description
Solver settings	
Method	Algorithm used to solve continuous models or the root node of a MIP model. Generally, barrier method provides the fastest run times for real-world problem set.
	CPLEX: CPX_PARAM_LPMETHOD - Default = 0; See link for more specifications.
	Gurobi: Method - Default = -1; See link for more specifications.
	clp: SolveType - Default = 5; See link for more specifications.
	HiGHS: Method - Default = "choose"; See link
BarConvTol	Convergence tolerance for barrier algorithm.
	CPLEX: CPX_PARAM_BAREPCOMP - Default = 1e-8; See link for more specifications.
	Gurobi: BarConvTol - Default = 1e-8; See linklink for more specifications.
Feasib_Tol	All constraints must be satisfied as per this tolerance. Note that this tolerance is absolute.

	CPLEX: CPX_PARAM_EPRHS - Default = 1e-6; See link for more specifications.
	Gurobi: FeasibilityTol - Default = 1e-6; See link for more specifications.
	clp: PrimalTolerance - Default = 1e-7; See link for more specifications.
	clp: DualTolerance - Default = 1e-7; See link for more specifications.
Optimal_Tol	Reduced costs must all be smaller than Optimal_Tol in the improving direction in order for a model to be declared optimal.
	CPLEX: CPX_PARAM_EPOPT - Default = 1e-6; See link for more specifications.
	Gurobi: OptimalityTol - Default = 1e-6; See link for more specifications.
Pre_Solve	Controls the presolve level.
	Gurobi: Presolve - Default = -1; See link for more specifications.
	clp: PresolveType - Default = 5; See link for more specifications.
Crossover	Determines the crossover strategy used to transform the interior solution produced by barrier algorithm into a basic solution.
	CPLEX: CPX_PARAM_SOLUTIONTYPE - Default = 2; See link for more specifications.
	Gurobi: Crossover - Default = 0; See link for more specifications.
NumericFocus	Controls the degree to which the code attempts to detect and manage numerical issues.
	CPLEX: CPX_PARAM_NUMERICALEMPHASIS - Default = 0; See link for more specifications.
	Gurobi: NumericFocus - Default = 0; See link for more specifications.
TimeLimit	Time limit to terminate the solution algorithm, model could also terminate if it reaches MIPGap before this time.
	CPLEX: CPX_PARAM_TILIM- Default = 1e+75; See link for more specifications.
	Gurobi: TimeLimit - Default = infinity; See link for more specifications.
	clp: MaximumSeconds - Default = -1; See link for more specifications.
MIPGap	Optimality gap in case of mixed-integer program.
	CPLEX: CPX_PARAM_EPGAP- Default = 1e-4; See link for more specifications.

	Gurobi: MIPGap - Default = 1e-4; See link for more specifications.
DualObjectiveLimit	When using dual simplex (where the objective is monotonically changing), terminate when the objective exceeds this limit.
	clp: DualObjectiveLimit - Default = 1e308; See link for more specifications.
MaximumIterations	Terminate after performing this number of simplex iterations.
	clp: MaximumIterations - Default = 2147483647; See link for more specifications.
LogLevel	Set to 1, 2, 3, or 4 for increasing output. Set to 0 to disable output.
	clp: logLevel - Default = 1; See link for more specifications.
	cbc: logLevel - Default = 1; See link for more specifications.
InfeasibleReturn	Set to 1 to return as soon as the problem is found to be infeasible (by default, an infeasibility proof is computed as well).
	clp: InfeasibleReturn - Default = 0; See link for more specifications.
Scaling	Sets or unsets scaling; 0 -off, 1 equilibrium, 2 geometric, 3 auto, 4 dynamic(later).
	clp: Scaling - Default = 3; See link for more specifications.
Perturbation	Perturbs problem; Switch on perturbation (50), automatic (100), don't try perturbing (102).
	clp: Perturbation - Default = 3; See link for more specifications.
maxSolutions	Terminate after this many feasible solutions have been found.
	cbc: maxSolutions - Default = -1; See link for more specifications.
maxNodes	Terminate after this many branch-and-bound nodes have been evaluated
	cbc: maxNodes - Default = -1; See link for more specifications.
allowableGap	Terminate after optimality gap is less than this value (on an absolute scale)
	cbc: allowableGap - Default = -1; See link for more specifications.
ratioGap	Terminate after optimality gap is smaller than this relative fraction.
	cbc: ratioGap - Default = Inf; See link for more specifications.

threads	Set the number of threads to use for parallel branch & bound.
	cbc: threads - Default = 1; See link for more specifications.

2 Inputs

All input files are in CSV format. Running the GenX model requires a minimum of five input files. Additionally, the user may need to specify five more input files based on model configuration and type of scenarios of interest. Description and column details of all potential input files are included in the `Input_data_explained` folder in the `Example_Systems` folder. Names of the input files and their functionality is also given below. Note that names of the input files are case sensitive.

Table 2: Summary of the input files

Column Name	Description
Mandatory Files	
Fuels_data.csv	Specify fuel type, CO2 emissions intensity, and time-series of fuel prices.
Network.csv	Specify network topology, transmission fixed costs, capacity and loss parameters.
Load_data.csv	Specify time-series of load profiles for each model zone, weights for each time step, load shedding costs, and optional time domain reduction parameters.
Generators_variability.csv	Specify time-series of capacity factor/availability for each resource.
Generators_data.csv	Specify cost and performance data for generation, storage and demand flexibility resources.
Settings-specific Files	
Reserves.csv	Specify operational reserve requirements as a function of load and renewables generation and penalty for not meeting these requirements.
Energy_share_requirement.csv	Specify regional renewable portfolio standard and clean energy standard style policies requiring minimum energy generation from qualifying resources.
CO2_cap.csv	Specify regional CO2 emission limits.
Capacity_reserve_margin.csv	Specify regional capacity reserve margin requirements.

Minimum_capacity_requirement.csv	Specify regional minimum technology capacity deployment requirements.

2.1 Mandatory input data

2.1.1 Fuels_data.csv

• **First row:** names of all fuels used in the model instance which should match the labels used in `Fuel` column in the `Generators_data.csv` file. For renewable resources or other resources that do not consume a fuel, the name of the fuel is `None`.

• **Second row:** The second row specifies the CO2 emissions intensity of each fuel in tons/MMBtu (million British thermal units). Note that by convention, tons correspond to metric tonnes and not short tons (although as long as the user is internally consistent in their application of units, either can be used).

• **Remaining rows:** Rest of the rows in this input file specify the time-series for prices for each fuel in /MMBtu. A constant price can be specified by entering the same value for all hours.

- ** First column:** The first column in this file denotes, Time_index, represents the index of time steps in a model instance.

2.1.2 Network.csv

This input file contains input parameters related to: 1) definition of model zones (regions between which transmission flows are explicitly modeled) and 2) definition of transmission network topology, existing capacity, losses and reinforcement costs. The following table describe each of the mandatory parameter inputs need to be specified to run an instance of the model, along with comments for the model configurations when they are needed.

Table 3: Structure of the Network.csv file

Column Name	Description
Settings-specific Columns	
Multiple zone model	
Network_Lines	Numerical index for each network line. The length of this column is counted but the actual values are not used.
z* (Network map) **OR** Origin*Zone, Destination*Zone	See below
Line_Max_Flow_MW	Existing capacity of the inter-regional transmission line.
NetworkExpansion = 1	

Line_Max_Reinforcement_MW	Maximum allowable capacity addition to the existing transmission line.
Line_Reinforcement_Cost_per_MWyr	Cost of adding new capacity to the inter-regional transmission line.
Trans_Loss_Segments = 1	
Line_Loss_Percentage	fractional transmission loss for each transmission line
Trans_Loss_Segments > 1	
Ohms	Line resistance in Ohms (used to calculate I^2R losses)
kV	Line voltage in kV (used to calculate I^2R losses)
CapacityReserveMargin > 0	
CapRes_*	Eligibility of the transmission line for adding firm capacity to the capacity reserve margin constraint. * represents the number of the capacity reserve margin constraint.
	1 = the transmission line is eligible for adding firm capacity to the region
	0 = the transmission line is not eligible for adding firm capacity to the region
DerateCapRes_*	(0,1) value represents the derating of the firm transmission capacity for the capacity reserve margin constraint.
CapResExcl_*	(-1,1,0) = -1 if the designated direction of the transmission line is inbound to locational deliverability area (LDA) modeled by the capacity reserve margin constraint. = 1 if the designated direction of the transmission line is outbound from the LDA modeled by the capacity reserve margin constraint. Zero otherwise.
MultiStage == 1	
Capital_Recovery_Period	Capital recovery period (in years) used for determining overnight capital costs from annualized investment costs for network transmission line expansion.
Line_Max_Flow_Possible_MW	Maximum possible line flow in the current model period. Overrides Line_Max_Reinforcement_MW, which is not used when performing multi-stage modeling.

There are two interfaces implemented for specifying the network topology itself: a matrix interface and a list interface. Only one choice is permitted in a given file.

The list interface consists of a column for the lines origin zone and one for the line's destination zone. Here is a snippet of the Network.csv file for a map with three zones and two lines:

```
Network_Lines, Origin_Zone, Destination_Zone,
            1,           1,                2,
            2,           1,                3,
```

The matrix interface requires N columns labeled `z1, z2, z3 ... zN`, and L rows, one for each network line (or interregional path), with a `1` in the column corresponding to the 'origin' zone and a `-1` in the column corresponding to the 'destination' zone for each line. Here is the same network map implemented as a matrix:

```
Network_Lines, z1, z2, z3,
            1,  1, -1,  0,
            2,  1,  0, -1,
```

Note that in either case, positive flows indicate flow from origin to destination zone; negative flows indicate flow from destination to origin zone.

2.1.3 Load_data.csv

This file includes parameters to characterize model temporal resolution to approximate annual grid operations, electricity demand for each time step for each zone, and cost of load shedding. Note that GenX is designed to model hourly time steps. With some care and effort, finer (e.g. 15 minute) or courser (e.g. 2 hour) time steps can be modeled so long as all time-related parameters are scaled appropriately (e.g. time period weights, heat rates, ramp rates and minimum up and down times for generators, variable costs, etc).

Table 4: Structure of the Load_data.csv file

Column Name	Description
Mandatory Columns	
Voll	Value of lost load in /MWh.
Demand_Segment	Number of demand curtailment/lost load segments with different cost and capacity of curtailable demand for each segment. User-specified demand segments. Integer values starting with 1 in the first row. Additional segements added in subsequent rows.
Cost_of_Demand_Curtailment_per_MW	Cost of non-served energy/demand curtailment (for each segment), reported as a fraction of value of lost load. If *Demand_Segment = 1*, then this parameter is a scalar and equal to one. In general this parameter is a vector of length equal to the length of Demand_Segment.
Max_Demand_Curtailment	Maximum time-dependent demand curtailable in each segment, reported as % of the demand in each zone and each period. *If Demand_Segment = 1*, then this parameter is a scalar and equal to one. In general this parameter is a vector of length given by length of Demand_segment.

Time_Index	Index defining time step in the model.
Load_MW_z*	Load profile of a zone z* in MW; if multiple zones, this parameter will be a matrix with columns equal to number of zones (each column named appropriate zone number appended to parameter) and rows equal to number of time periods of grid operations being modeled.
Rep_Periods	Number of representative periods (e.g. weeks, days) that are modeled to approximate annual grid operations. This is always a single entry. For a full-year model, this is `1`.
Timesteps_per_Rep_Period	Number of timesteps per representative period (e.g. 168 if period is set as a week using hour-long time steps). This is always a single entry: all representative periods have the same length. For a full-year model, this entry is equal to the number of time steps.
Sub_Weights	Number of annual time steps (e.g. hours) represented by each timestep in a representative period. The length of this column is equal to the number of representative periods. The sum of the elements should be equal to the total number of time steps in a model time horizon (e.g. 8760 hours if modeling 365 days or 8736 if modeling 52 weeks).

2.1.4 Generator_variability.csv

This file contains the time-series of capacity factors / availability of each resource included in the `Generators_data.csv` file for each time step (e.g. hour) modeled.

• first column: The first column contains the time index of each row (starting in the second row) from 1 to N.

• Second column onwards: Resources are listed from the second column onward with headers matching each resource name in the `Generators_data.csv` file in any order. The availability for each resource at each time step is defined as a fraction of installed capacity and should be between 0 and 1. Note that for this reason, resource names specified in `Generators_data.csv` must be unique. Note that for Hydro reservoir resources (i.e. `HYDRO = 1` in the `Generators_data.csv`), values in this file correspond to inflows (in MWhs) to the hydro reservoir as a fraction of installed power capacity, rather than hourly capacity factor.

2.1.5 Generators_data.csv

This file contains cost and performance parameters for various generators and other resources (storage, flexible demand, etc) included in the model formulation.

Table 5: Mandatory columns in the Generators_data.csv file

Column Name	Description
Resource	This column contains **unique** names of resources available to the model. Resources can include generators, storage, and flexible or time shiftable demand/loads.

Zone	Integer representing zone number where the resource is located.
Technology type flags	
New_Build	{-1, 0, 1}, Flag for resource (storage, generation) eligibility for capacity expansion.
	New_Build = 1: eligible for capacity expansion and retirement.
	New_Build = 0: not eligible for capacity expansion, eligible for retirement.
	New_Build = -1: not eligible for capacity expansion or retirement.
THERM	{0, 1, 2}, Flag to indicate membership in set of thermal resources (e.g. nuclear, combined heat and power, natural gas combined cycle, coal power plant)
	THERM = 0: Not part of set (default)
	THERM = 1: If the power plant relies on thermal energy input and subject unit commitment constraints/decisions if `UCommit >= 1` (e.g. cycling decisions/costs/constraints).
	THERM = 2: If the power plant relies on thermal energy input and is subject to simplified economic dispatch constraints (ramping limits and minimum output level but no cycling decisions/costs/constraints).
Cap_size	Size (MW) of a single generating unit. This is used only for resources with integer unit commitment (`THERM = 1`) - not relevant for other resources.
VRE	{0, 1}, Flag to indicate membership in set of dispatchable (or curtailable) variable renewable energy resources (onshore wind, offshore wind, utility-scale solar PV, and distributed solar PV subject to dispatch signals).
	VRE = 0: Not part of set (default)
	VRE = 1: Dispatchable variable renewable energy (VRE) resources.
Num_VRE_bins	Number of resource availability profiles considered for each VRE resource per zone. This parameter is used to decide the number of capacity investment decision variables related to a single variable renewable energy technology in each zone.
	Num_VRE_bins = 1: using a single resource availability profile per technology per zone. 1 capacity investment decision variable and 1 generator RID tracking technology power output (and in each zone).
	Num_VRE_bins > 1: using multiple resource availability profiles per technology per zone. Num_VRE_bins capacity investment decision variables and 1 generator RID used to define technology power output at each time step (and in each zone). Example: Suppose we are modeling 3 bins of wind profiles for each zone. Then include 3 rows with wind resource names as Wind_1, Wind_2,

	and Wind_3 and a corresponding increasing sequence of RIDs. Set Num_VRE_bins for the generator with smallest RID, Wind_1, to be 3 and set Num_VRE_bins for the other rows corresponding to Wind_2 and Wind_3, to be zero. By setting Num_VRE_bins for Wind_2 and Wind_3, the model eliminates the power outputs variables for these generators. The power output from the technology across all bins is reported in the power output variable for the first generator. This allows for multiple bins without significantly increasing number of model variables (adding each bin only adds one new capacity variable and no operational variables). See documentation for `curtailable_variable_renewable()` for more.
MUST_RUN	{0, 1}, Flag to indicate membership in set of must-run plants (could be used to model behind-the-meter PV not subject to dispatch signals/curtailment, run-of-river hydro that cannot spill water, must-run or self-committed thermal generators, etc).
	MUST_RUN = 0: Not part of set (default)
	MUST_RUN = 1: Must-run (non-dispatchable) resources.
STOR	{0, 1, 2}, Flag to indicate membership in set of storage resources and designate which type of storage resource formulation to employ.
	STOR = 0: Not part of set (default)
	STOR = 1: Discharging power capacity and energy capacity are the investment decision variables; symmetric charge/discharge power capacity with charging capacity equal to discharging capacity (e.g. lithium-ion battery storage).
	STOR = 2: Discharging, charging power capacity and energy capacity are investment variables; asymmetric charge and discharge capacities using distinct processes (e.g. hydrogen electrolysis, storage, and conversion to power using fuel cell or combustion turbine).
FLEX	{0, 1}, Flag to indicate membership in set of flexible demand-side resources (e.g. scheduleable or time shiftable loads such as automated EV charging, smart thermostat systems, irrigating pumping loads etc).
	FLEX = 0: Not part of set (default)
	FLEX = 1: Flexible demand resource.
HYDRO	{0, 1}, Flag to indicate membership in set of reservoir hydro resources.
	HYDRO = 0: Not part of set (default)
	HYDRO = 1: Hydropower with reservoir modeling, including inflows, spillage, ramp rate limits and minimum operating level and efficiency loss associated with discharging. Reservoir capacity can be represented as a ratio or energy to power. This type of plant cannot charge from grid.

LDS	{0, 1}, Flag to indicate the resources eligible for long duration storage constraints with inter period linkage (e.g., reservoir hydro, hydrogen storage)
	LDS = 0: Not part of set (default)
	LDS = 1: Long duration storage resources
Existing technology capacity	
Existing_Cap_MW	The existing capacity of a power plant in MW.
Existing_Cap_MWh	The existing capacity of storage in MWh where `STOR = 1` or `STOR = 2`.
Existing_Charge_Cap_MW	The existing charging capacity for resources where `STOR = 2`.
Capacity/Energy requirements	
Max_Cap_MW	-1 (default) – no limit on maximum discharge capacity of the resource. If non-negative, represents maximum allowed discharge capacity (in MW) of the resource.
Max_Cap_MWh	-1 (default) – no limit on maximum energy capacity of the resource. If non-negative, represents maximum allowed energy capacity (in MWh) of the resource with `STOR = 1` or `STOR = 2`.
Max_Charge_Cap_MW	-1 (default) – no limit on maximum charge capacity of the resource. If non-negative, represents maximum allowed charge capacity (in MW) of the resource with `STOR = 2`.
Min_Cap_MW	-1 (default) – no limit on minimum discharge capacity of the resource. If non-negative, represents minimum allowed discharge capacity (in MW) of the resource.
Min_Cap_MWh	-1 (default) – no limit on minimum energy capacity of the resource. If non-negative, represents minimum allowed energy capacity (in MWh) of the resource with `STOR = 1` or `STOR = 2`.
Min_Charge_Cap_MW	-1 (default) – no limit on minimum charge capacity of the resource. If non-negative, represents minimum allowed charge capacity (in MW) of the resource with `STOR = 2`.
Cost parameters	
Inv_Cost_per_MWyr	Annualized capacity investment cost of a technology (/MW/year).
Inv_Cost_per_MWhyr	Annualized investment cost of the energy capacity for a storage technology (/MW/year), applicable to either `STOR = 1` or `STOR = 2`.
Inv_Cost_Charge_per_MWyr	Annualized capacity investment cost for the charging portion of a storage technology with `STOR = 2` (/MW/year).

Fixed_OM_Cost_per_MWy	Fixed operations and maintenance cost of a technology (/MW/year).
Fixed_OM_Cost_per_MWhyr	Fixed operations and maintenance cost of the energy component of a storage technology (/MWh/year).
Fixed_OM_Cost_charge_per_MWyr	Fixed operations and maintenance cost of the charging component of a storage technology of type `STOR = 2`.
Var_OM_Cost_per_MWh	Variable operations and maintenance cost of a technology (/MWh).
Var_OM_Cost_per_MWhIn	Variable operations and maintenance cost of the charging aspect of a storage technology with `STOR = 2`, or variable operations and maintenance costs associated with flexible demand deferral with `FLEX = 1`. Otherwise 0 (/MWh).
Technical performance parameters	
Heat_Rate_MMBTU_per_MWh	Heat rate of a generator or MMBtu of fuel consumed per MWh of electricity generated for export (net of on-site house loads). The heat rate is the inverse of the efficiency: a lower heat rate is better. Should be consistent with fuel prices in terms of reporting on higher heating value (HHV) or lower heating value (LHV) basis.
Fuel	Fuel needed for a generator. The names should match with the ones in the `Fuels_data.csv`.
Self_Disch	[0,1], The power loss of storage technologies per hour (fraction loss per hour)- only applies to storage techs.
Min_Power	[0,1], The minimum generation level for a unit as a fraction of total capacity. This value cannot be higher than the smallest time-dependent CF value for a resource in `Generators_variability.csv`. Applies to thermal plants, and reservoir hydro resource (`HYDRO = 1`).
Ramp_Up_Percentage	[0,1], Maximum increase in power output from between two periods (typically hours), reported as a fraction of nameplate capacity. Applies to thermal plants, and reservoir hydro resource (`HYDRO = 1`).
Ramp_Dn_Percentage	[0,1], Maximum decrease in power output from between two periods (typically hours), reported as a fraction of nameplate capacity. Applies to thermal plants, and reservoir hydro resource (`HYDRO = 1`).
Eff_Up	[0,1], Efficiency of charging storage – applies to storage technologies (all STOR types).
Eff_Down	[0,1], Efficiency of discharging storage – applies to storage technologies (all STOR types).
Hydro_Energy_to_Power_Ratio	The rated number of hours of reservoir hydro storage at peak discharge power output. Applies to `HYDRO = 1` (hours).
Min_Duration	Specifies the minimum ratio of installed energy to discharged power capacity that can be installed. Applies to STOR types 1 and 2 (hours).
Max_Duration	Specifies the maximum ratio of installed energy to discharged power capacity that can be installed. Applies to STOR types 1 and 2 (hours).

Max_Flexible_Demand_Delay	Maximum number of hours that demand can be deferred or delayed. Applies to resources with FLEX type 1 (hours).
Max_Flexible_Demand_Advance	Maximum number of hours that demand can be scheduled in advance of the original schedule. Applies to resources with FLEX type 1 (hours).
Flexible_Demand_Energy_Eff	[0,1], Energy efficiency associated with time shifting demand. Represents energy losses due to time shifting (or 'snap back' effect of higher consumption due to delay in use) that may apply to some forms of flexible demand. Applies to resources with FLEX type 1 (hours). For example, one may need to pre-cool a building more than normal to advance demand.
Required for writing outputs	
region	Name of the model region
cluster	Number of the cluster when representing multiple clusters of a given technology in a given region.
MultiStage == 1	
Capital_Recovery_Period	Capital recovery period (in years) used for determining overnight capital costs from annualized investment costs.
Lifetime	Lifetime (in years) used for determining endogenous retirements of newly built capacity.
Min_Retired_Cap_MW	Minimum required discharge capacity retirements in the current model period. This field can be used to enforce lifetime retirements of existing capacity.
Min_Retired_Cap_MW	Minimum required discharge capacity retirements in the current model period. This field can be used to enforce lifetime retirements of existing discharge capacity.
Min_Retired_Energy_Cap_MW	Minimum required energy capacity retirements in the current model period. This field can be used to enforce lifetime retirements of existing energy capacity.
Min_Retired_Charge_Cap_MW	Minimum required energy capacity retirements in the current model period. This field can be used to enforce lifetime retirements of existing charge capacity.

Table 6: Settings-specific columns in the Generators_data.csv file

Column Name	Description
UCommit >= 1	The following settings apply only to thermal plants with unit commitment constraints (`THERM = 1`).

Up_Time	Minimum amount of time a resource has to stay in the committed state.
Down_Time	Minimum amount of time a resource has to remain in the shutdown state.
Start_Cost_per_MW	Cost per MW of nameplate capacity to start a generator (/MW per start). Multiplied by the number of generation units (each with a pre-specified nameplate capacity) that is turned on.
Start_Fuel_MMBTU_per_MW	Startup fuel use per MW of nameplate capacity of each generator (MMBtu/MW per start).
Reserves = 1	The following settings apply to thermal, dispatchable VRE, hydro and storage resources
Reg_Cost	Cost of providing regulation reserves (/MW per time step/hour).
Rsv_Cost	Cost of providing upwards spinning or contingency reserves (/MW per time step/hour).
Reg_Max	[0,1], Fraction of nameplate capacity that can committed to provided regulation reserves. .
Rsv_Max	[0,1], Fraction of nameplate capacity that can committed to provided upwards spinning or contingency reserves.
EnergyShareRequirement > 0	
ESR_*	Flag to indicate which resources are considered for the Energy Share Requirement constraint.
	1- included
	0- excluded
CapacityReserveMargin > 0	
CapRes_*	[0,1], Fraction of the resource capacity eligible for contributing to the capacity reserve margin constraint (e.g. derate factor).
ModelingToGenerateAlternatives = 1	
MGA	Eligibility of the technology for Modeling To Generate Alternative (MGA) run.
	1 = Technology is available for the MGA run.
	0 = Technology is unavailable for the MGA run (e.g. storage technologies).
Resource_Type	For the MGA run, we categorize all the resources in a few resource types. We then find maximally different generation portfolio based on these resource types. For example, existing solar and new solar resources could be represented by a resource type names `Solar`. Categorization of resources into resource types is user dependent.

MinCapReq = 1	
MinCapTag_*	Eligibility of resources to participate in Minimum Technology Carveout constraint. * corresponds to the ith row of the file `Minimum_capacity_requirement.csv`.
MaxCapReq = 1	
MaxCapTag_*	Eligibility of resources to participate in Maximum Technology Carveout constraint. * corresponds to the ith row of the file `Maximum_capacity_requirement.csv`.

2.2 Optional inputs files

2.2.1 Online Time-domain reduction

Modeling grid operations for each hour of the year can be computationally expensive for models with many zones and resources. Time-domain reduction is often employed in capacity expansion models as a way to balance model spatial and temporal resolution as well as representation of dispatch, while ensuring reasonable computational times. GenX allows the option of performing time-domain reduction on the user supplied time-series input data to produce a representative time series at the desired level of temporal resolution. The below table summarizes the list of parameters to be specified by the user to perform the time domain reduction implemented in GenX. These parameters are passed to GenX via the YAML file `time_domain_reduction_settings.yml`.

Table 7: Structure of the time*domain*reduction.yml file

Key	Description
Timesteps_per_period	The number of timesteps (e.g., hours) in each representative period (i.e. 168 for weeks, 24 for days, 72 for three-day periods, etc).
UseExtremePeriods	1 = Include outliers (by performance or load/resource extreme) as their own representative extreme periods. This setting automatically includes periods based on criteria outlined in the dictionary `ExtremePeriods`. Extreme periods can be selected based on following criteria applied to load profiles or solar and wind capacity factors profiles, at either the zonal or system level. A) absolute (timestep with min/max value) statistic (minimum, maximum) and B) integral (period with min/max summed value) statistic (minimum, maximum). For example, the user could want the hour with the most load across the whole system to be included among the extreme periods. They would select Load, System, Absolute, and Max.
	0 = Do not include extreme periods.
ExtremePeriods	If UseExtremePeriods = 1, use this dictionary to select which types of extreme periods to use. Select by profile type (Load, PV, or Wind), geography (Zone or System), grouping by timestep or by period (Absolute or Integral), and statistic (Maximum or Minimum).
ClusterMethod	Either `kmeans` or `kmedoids`, the method used to cluster periods and determine each time step's representative period.

ScalingMethod	Either ‘N' or ‘S', the decision to normalize ([0,1]) or standardize (mean 0, variance 1) the input data prior to clustering.
MinPeriods	The minimum number of representative periods used to represent the input data. If using UseExtremePeriods, this must be greater or equal to the number of selected extreme periods. If `IterativelyAddPeriods` is off, this will be the total number of representative periods.
MaxPeriods	The maximum number of representative periods - both clustered and extreme - that may be used to represent the input data.
IterativelyAddPeriods	1 = Add representative periods until the error threshold between input data and represented data is met or the maximum number of representative periods is reached.
	0 = Use only the minimum number of representative periods. This minimum value includes the selected extreme periods if `UseExtremePeriods` is on.
Threshold	Iterative period addition will end if the period farthest from its representative period (as measured using Euclidean distance) is within this percentage of the total possible error (for normalization) or 95% of the total possible error (± 2 σ for standardization). E.g., for a threshold of 0.01, each period must be within 1% of the spread of possible error before the clustering iterations will terminate (or until the maximum is reached).
IterateMethod	Either ‘cluster' (Default) or ‘extreme', whether to increment the number of clusters to the kmeans/kmedoids method or to set aside the worst-fitting periods as a new extreme periods.
nReps	Default 200, the number of kmeans/kmedoids repetitions at the same setting.
LoadWeight	Default 1, a multiplier on load columns to optionally prioritize better fits for load profiles over resource capacity factor or fuel price profiles.
WeightTotal	Default 8760, the sum to which the relative weights of representative periods will be scaled.
ClusterFuelPrices	Either 1 or 0, whether or not to use the fuel price time series in `Fuels_data.csv` in the clustering process. If 'no', this function will still write `Fuels_data.csv` in the TimeDomainReductionFolder with reshaped fuel prices based on the number and size of the representative periods but will not use the fuel price time series for selection of representative periods.

2.2.2 Reserves.csv

This file includes parameter inputs needed to model time-dependent procurement of regulation and spinning reserves. This file is needed if `Reserves` flag is activated in the YAML file `genx_settings.yml`.

Table 8: Structure of the Reserves.csv file

Column Name	Description
Reg_Req_Percent_Load	[0,1], Regulation requirement as a percent of time-dependent load; here load is the total across all model zones.

Reg_Req_Percent_VRE	[0,1], Regulation requirement as a percent of time-dependent wind and solar generation (summed across all model zones).
Rsv_Req_Percent_Load [0,1],	Spinning up or contingency reserve requirement as a percent of time-dependent load (which is summed across all zones).
Rsv_Req_Percent_VRE	[0,1], Spinning up or contingency reserve requirement as a percent of time-dependent wind and solar generation (which is summed across all zones).
Unmet_Rsv_Penalty_Dollar_per_MW	Penalty for not meeting time-dependent spinning reserve requirement (/MW per time step).
Dynamic_Contingency	Flags to include capacity (generation or transmission) contingency to be added to the spinning reserve requirement.
Dynamic_Contingency	= 1: contingency set to be equal to largest installed thermal unit (only applied when `UCommit = 1`).
	= 2: contingency set to be equal to largest committed thermal unit each time period (only applied when `UCommit = 1`).
Static_Contingency_MW	A fixed static contingency in MW added to reserve requirement. Applied when `UCommit = 1` and `DynamicContingency = 0`, or when `UCommit = 2`. Contingency term not included in operating reserve requirement when this value is set to 0 and DynamicContingency is not active.

2.2.3 Energy_share_requirement.csv

This file contains inputs specifying minimum energy share requirement policies, such as Renewable Portfolio Standard (RPS) or Clean Energy Standard (CES) policies. This file is needed if parameter EnergyShareRequirement has a non-zero value in the YAML file `genx_settings.yml`.

Note: this file should use the same region name as specified in the `Generators_data.csv` file.

Table 9: Structure of the Energy_share_requirement.csv file

Column Name	Description
Region_description	Region name
Network_zones	zone number represented as z*
ESR_*	[0,1], Energy share requirements as a share of zonal demand (calculated on an annual basis). * represents the number of the ESR constraint, given by the number of ESR_* columns in the `Energy_share_requirement.csv` file.

2.2.4 CO2_cap.csv

This file contains inputs specifying CO2 emission limits policies (e.g. emissions cap and permit trading programs). This file is needed if `CO2Cap` flag is activated in the YAML file `genx_settings.yml`. `CO2Cap` flag set to 1 represents mass-based (tCO2) emission target. `CO2Cap` flag set to 2 is specified when emission target is given in terms of rate (tCO2/MWh) and is based on total demand met. `CO2Cap` flag set to 3 is specified when emission target is given in terms of rate (tCO2 /MWh) and is based on total generation.

Table 10: Structure of the CO2_cap.csv file

Column Name	Description
Region_description	Region name
Network_zones	zone number represented as z*
CO_2_Cap_Zone_*	If a zone is eligible for the emission limit constraint, then this column is set to 1, else 0.
CO_2_Max_tons_MWh_*	Emission limit in terms of rate
CO_2_Max_Mtons_*	Emission limit in absolute values, in Million of tons
	where in the above inputs, * represents the number of the emission limit constraints. For example, if the model has 2 emission limit constraints applied separately for 2 zones, the above CSV file will have 2 columns for specifying emission limit in terms on rate: CO_2_Max_tons_MWh_1 and CO_2_Max_tons_MWh_2.

2.2.5 Capacity_reserve_margin.csv

This file contains the regional capacity reserve margin requirements. This file is needed if parameter CapacityReserveMargin has a non-zero value in the YAML file `genx_settings.yml`.

Note: this file should use the same region name as specified in the `Generators_data.csv` file

Table 11: Structure of the Capacity_reserve_margin.csv file

Column Name	Description
Region_description	Region name
Network_zones	zone number represented as z*
CapRes_*	[0,1], Capacity reserve margin requirements of a zone, reported as a fraction of demand

2.2.6 Minimum_capacity_requirement.csv

This file contains the minimum capacity carve-out requirement to be imposed (e.g. a storage capacity mandate or offshore wind capacity mandate). This file is needed if the `MinCapReq` flag has a non-zero value in the YAML file `genx_settings.yml`.

Table 12: Structure of the Minimum_capacity_requirement.csv file

Column Name	Description
MinCapReqConstraint	Index of the minimum capacity carve-out requirement.
Constraint_Description	Names of minimum capacity carve-out constraints; not to be read by model, but used as a helpful notation to the model user.
Min_MW	minimum capacity requirement [MW]

Some of the columns specified in the input files in Section 2.2 and 2.1 are not used in the GenX model formulation. These columns are necessary for interpreting the model outputs and used in the output module of the GenX.

2.2.7 Maximum_capacity_requirement.csv

This contains the maximum capacity limits to be imposed (e.g. limits on total deployment of solar, wind, or batteries in the system as a whole or in certain collections of zones). It is required if the `MaxCapReq` flag has a non-zero value in `genx_settings.yml`.

Table 13: Structure of the Maximum_capacity_requirement.csv file

Column Name	Description
MaxCapReqConstraint	Index of the maximum capacity limit.
Constraint_Description	Names of maximum capacity limit; not to be read by model, but used as a helpful notation to the model user.
Max_MW	maximum capacity limit [MW]

Some of the columns specified in the input files in Section 2.2 and 2.1 are not used in the GenX model formulation. These columns are necessary for interpreting the model outputs and used in the output module of the GenX.

2.2.8 Method_of_morris_range.csv

This file contains the settings parameters required to run the Method of Morris algorithm in GenX. This file is needed if the `MethodofMorris` flag is ON in the YAML file `genx_settings.yml`.

Table 14: Structure of the Method_of_morris_range.csv file

Column Name	Description
Resource	This column contains **unique** names of resources available to the model. Resources can include generators, storage, and flexible or time shiftable demand/loads.
Zone	Integer representing zone number where the resource is located.
Lower_bound	Percentage lower deviation from the nominal value
Upper_bound	Percentage upper deviation from the nominal value
Parameter	Column from the `Generators_data.csv` file containing uncertain parameters
Group	Group the uncertain parameters that will be changed all at once while performing the sensitivity analysis. For example, if the fuel price of natural gas is uncertain, all generators consuming natural gas should be in the same group. Group name is user defined
p_steps	Number of steps between upper and lower bound
total_num_trajectory	Total number of trakectories through the design matrix
num_trajectory	Selected number of trajectories throigh the design matrix
len_design_mat	Length of the design matrix
policy	Name of the policy

Notes:

1. Upper and lower bounds are specified in terms of percentage deviation from the nominal value.
2. Percentage variation for uncertain parameters in a given group is identical. For example, if solar cluster 1 and solar cluster 2 both belong to the 'solar' group, their Lower*bound and Upper*bound must be identical
3. P_steps should at least be = 1\%, i.e., Upper_bound – Lower_bound $<$ p_steps
4. P_steps for parameters in one group must be identical
5. Total_num_trajectory should be around 3 to 4 times the total number of uncertain parameters
6. num_trajectory should be approximately equal to the total number of uncertain parameters
7. len_design_mat should be 1.5 to 2 times the total number of uncertain parameters
8. Higher number of num_trajectory and len*design*mat would lead to higher accuracy
9. Upper and lower bounds should be specified for all the resources included in the `Generators_data.csv` file. If a parameter related to a particular resource is not uncertain, specify upper bound = lower bound = 0.