

Mathematical formulation of the World Bank Electricity Planning
Modeling (EPM): 2023 update
ESMAP

Contents

1.1	Introduction	3
1.1.1	Modeling Assumptions	4
1.1.2	What can the model do?.....	4
1.2	Notation	5
1.2.1	Indices and Sets.....	5
1.2.2	Variables.....	6
1.2.3	Parameters.....	8
1.3	Model formulation.....	12
1.4	Description of the model	20
1.4.1	Indices and Sets.....	20
1.4.2	Objective function.....	21
1.4.3	Load approximation	21
1.4.4	Value of Lost Load.....	22
1.4.5	Transmission network constraints	23
1.4.6	System requirements	24
1.4.7	Generation constraints	25
1.4.8	Renewable generation modeling	26
1.4.9	Concentrated Solar Power (CSP) modeling.....	26
1.4.10	Time consistency of power system additions and retirements	27
1.4.11	Storage modeling	28
1.4.12	Investment constraints	30
1.4.13	Environmental policy	31
1.5	Limitations of the model.....	36

1 Introduction

The World Bank's Electricity Planning Model (EPM) is a long-term, multi-year, multi-zone capacity expansion and dispatch model. The objective of the model is to minimize the sum of fixed (including annualized capital costs) and variable generation costs (discounted for time) for all zones and all years considered, subject to:

- Demand equals the sum of generation and non-served energy,
- Available capacity is existing capacity plus new capacity minus retired capacity,
- Generation does not exceed the max and min output limits of the units,
- Generation is constrained by ramping limits,
- Spinning reserves are committed every hour to compensate forecasting errors,
- renewable generation constrained by wind and solar hourly availability,
- Excess energy can be stored in storage units to be released later or traded between the other zones, and
- Transmission network topology and transmission line thermal limits.

The model is an abstract representation of the real power systems with certain limitations described in more detail in section 5.

EPM is used mostly to perform least cost expansion plans as well as dispatch analyses. Developed by ESMAP since 2015, EPM has been used to study dozens of countries. EPM uses a core mixed integer programming multi-zone model, implemented in the General Algebraic Modelling System ([GAMS](#)), to minimize the total discounted costs of capex and opex. The model optimizes the generation (and transmission if included) expansion over future years of the modeled time period as well as the underlying dispatch of generation and power flows on transmission lines for existing and new transmission assets. In addition to the core decision variables on generation and transmission capacity addition, dispatch, and flow, the model also co-optimizes spinning reserve provision from the generators. EPM had been parameterized for 87 client countries and economies as of end-January 2022 (Table 1) and is being used for many types of analyses.

Table 1: Countries and Economies with EPM Models

Region/Country	Number of economies	Country/economy names ^a
Central Asia + Pakistan	7	Afghanistan, Kazakhstan, Kyrgyz Rep., Tajikistan, Turkmenistan, Uzbekistan, Pakistan
Eastern Africa Power Pool + Zambia	14	Burundi; Congo, Dem. Rep.; Egypt, Arab Rep.; Libya; Djibouti; Eritrea; Ethiopia; Kenya; Rwanda; Tanzania; Uganda; South Sudan; Sudan, Zambia
Southern Africa Power Pool	12	Angola; Botswana; Congo, Dem. Rep.; Eswatini; Lesotho; Mozambique; Malawi; Namibia; South Africa; Tanzania; Zambia; Zimbabwe
Pan Arab Electricity Model	17	Algeria; Bahrain; Egypt, Arab Rep.; Iraq; Jordan; Kuwait; Lebanon; Libya; Morocco; Oman; Qatar; Saudi Arabia; Sudan; Syria; Tunisia; United Arab Emirates; West Bank and Gaza; Yemen
South Asia	5	Bangladesh, Bhutan, India, Nepal, Sri Lanka (BBINS)

West Africa Power Pool	15	Benin, Burkina Faso, Côte d'Ivoire, The Gambia, Guinea, Guinea Bissau, Ghana, Niger, Nigeria, Liberia, Mali, Mauritania, Togo, Senegal, Sierra Leone
Country-specific models ^b	18	Albania, Bosnia and Herzegovina, Bulgaria, Cameroon, Central African Republic, Chad, Comoros, Kosovo, Lao PDR, Madagascar, Mongolia, Myanmar, Papua New Guinea, Poland, Solomon Islands, Turkey, Ukraine, Vietnam
TOTAL	88	

a. The Democratic Republic of Congo is a member of both Southern and Eastern Africa Power Pools. Zambia is not a member of the Eastern Africa Power Pool but has been added to that group of countries.

b. These represent 18 stand-alone models for the individual countries.

This version is an update of the [previous documentation](#). The main changes relate to the addition of new technologies (CCS, PV with storage) or to the improvement of existing technologies (storage in particular), and to the addition of new features regarding imports and exports with zones outside of the model and of transmission expansion capabilities, of VRE representation, etc. Additionally, the code has been updated to be more efficient (without changing the equations but rewriting it in a more efficient GAMS syntax).

1.1 Modeling Assumptions

The model is derived based on the following assumptions:

1. Market participants are not strategic and they behave in a perfectly competitive manner, i.e., the power plant owners submit their true costs as bids.
2. The forecasted demand is considered perfectly inelastic, which implies that the maximization of the social welfare can be replaced by minimization of the system cost.
3. The trade among regions/countries is economically efficient (optimal), which translates to a single objective of minimization of cost for the whole region/country.
4. The pricing is efficient and does not provide incentives to market participants to deviate from the optimal behavior.

1.2 What can the model do?

- Identify investment needs (generation mostly) to meet the demand under various policy scenarios
- Determine where and by how much renewable resources should be deployed to maximize their value to the system.
- Determine the optimal capacity additions over time to complement renewable generation accounting for existing generating units, energy storage, demand-side response and/or a carbon constraint.
- Determine the optimal retirement schedule of the existing units over time.
- Assess the utilization of the transmission lines (important to design trade contracts).
- Determine the optimal transmission capacity additions.

- Determine hourly electricity prices with trade for different countries and zones, which is essential to value the energy traded.
- Determine the operations and expansion cost to the consumer.
- Determine the impact of different market conditions (e.g., fuel prices, fuel subsidies, carbon limits, etc.) and technology cost assumptions on the optimal capacity expansion plan and the optimal energy mix.
- Determine the cost of implementing specific environmental policies: renewable portfolio standard, cap on carbon emissions, tax on carbon emissions, and carbon emissions rate.

2 Notation

2.1 Indices and Sets

$c \in C$	where C is the set of countries
$d \in D$	where D is the set of types of days or weeks
$f \in F$	where F is the set of fuels
$g \in G$	where G is the set of generators that can be built or the set of technology-specific types of aggregated generators
$q \in Q$	where Q is the set of seasons or quarters
$h \in H$	where H is the set of electrolyzers that can be built
$y \in Y$	where Y is the set of years considered in the planning model
$z, z2 \in Z$	where Z is the set of zones/regions modeled
$sc \in S$	where S is the set of flags and penalties used to include/exclude certain features of the model
Subsets considered	
$EG, NG \in G$	where EG and NG is a partition of set G and the former (EG) contains generators existing at the starting year of the planning horizon and the latter (NG) contains candidate generators ¹
$DC \text{ and } NDC \in G$	where DC and NDC are partitions of set G separating generators with a CAPEX cost constant over the modeling horizon and those generators that have costs varying over time. This feature is mainly developed to account for technologies that show cost reductions due to technological improvements and

¹ The generators already planned are included in any of the two sets depending on criteria such as their capacity, status of their construction process etc.

	manufacturing advances.
$EH, NH \in H$	where EH and NH is a partition of set H and the former (HG) contains electrolyzers existing at the starting year of the planning horizon and the latter (NH) contains candidate electrolyzers ²
$DH \text{ and } NDH \in H$	where DH and NDH are partitions of set H separating electrolyzers with a CAPEX cost constant over the modeling horizon and those electrolyzers that have costs varying over time. This feature is mainly developed to account for technologies that show cost reductions due to technological improvements and manufacturing advances.
$so \in G$	PV unit linked to a storage unit as a partition of set G (this is the PV part of PV linked with storage)
$stp \in st$	Storage unit linked to a PV plant as a partition of set G (this is the storage part of PV linked with storage)
$MD \in D$	where MD is a subset of days the planner expects the minimum load levels to be binding
$RE \in F$	where RE is a subset of set F considered as renewable according to regulator's criteria ³
$RG \in G$	where MD is a subset of days the planner expects the minimum load levels to be binding
$map_{c,z}$	includes the subset of zones that correspond to each country
$map_{g,f}$	includes valid combinations of fuels and generators; subset of the set $G \times F$
$map_{eg,ng}$	includes combinations of existing and candidate generators; subset of set $EG \times NG$
$map_{g,z}$	includes the subset of generators that correspond to each zone
$map_{h,z}$	includes the subset of electrolyzers that correspond to each zone

2.2 Variables

Non-negative decision variables	
$AdditionalTransfer_{z,z2,y}$	Additional number of lines between z and $z2$ in year y
$build_{g,y}$	Generator investment in MW
$buildH2_{h,y}$	Electrolyzer investment in MW

² The electrolyzers already planned are included in any of the two sets depending on criteria such as their capacity, status of their construction process etc.

³ Type of resources considered as renewables might be different from country to country or state to state. For example, some states do not include hydropower towards their renewable targets (e.g. California does not count large hydropower towards the RPS (Renewable portfolio standard) while others such as Oregon do [13].

$builtCapVar_{g,y}$	Integer variable to model discrete unit capacity
$builtCapVarH2_{h,y}$	Integer variable to model discrete unit capacity
$builtTransmission_{z,z2,y}$	Integer variable to model number of new transmission lines added between z and $z2$ in year y
$cap_{g,y}$	Generator capacity available at year y in MW
$capH2_{h,y}$	Electrolyzer capacity available at year y in MW
$CO2backstop_c,y$	Annual country CO2 emissions above the system emissions constraint $cap_{Sy_emission_cap_y}$
$emissions_{z,y}$	Total system emissions of carbon dioxide in tons
$emissions_Zo_{z,y}$	Emissions of carbon dioxide in tons per zone z
$exportPrice_{z,q,d,t,y}$	External price driven export in MW
$fuelH2Quarter_{z,q,y}$	Hydrogen fuel consumption for electricity production (hydrogen produced intended for re-circulation in the power sector in (MMBTU per quarter and year)
$fuelH2_{z,q,y}$	Hydrogen fuel consumption for electricity production (hydrogen produced intended for re-circulation in the power sector in (MMBTU per year)
$fuel_{z,f,y}$	Fuel consumption in MMBTU
$gen_{g,f,q,d,t,y}$	Generator output in MW
$genCSP_{g,z,q,d,t,y}$	Power output of the solar panel in MW
$importPrice_{z,q,d,t,y}$	External price driven import in MW
$H2PwrIn_{h,q,d,t,y}$	Power drawn by electrolyzer h in MW
$retire_{g,y}$	Capacity in MW retired
$retireCapVar_{g,y}$	Supplementary integer variable to model discrete unit capacity
$retireCapVarH2_{h,y}$	Supplementary integer variable to model discrete unit capacity
$reserve_{g,q,d,t,y}$	Spinning reserve requirement met in MW
$storage_{st,q,d,t,y}$	Level of energy in MWh stored in storage unit st
$storage_inj_{st,q,d,t,y}$	Power injected in MW in storage unit st is charged during hour (q, d, t)
$storage_capacity_{st,q,d,t,y}$	Total deployed energy storage capacity in MWh for storage unit st
$storageCSP_{g,z,q,d,t,y}$	Level of energy in MWh stored in CSP unit at zone z
$storageCSPinj_{g,z,q,d,t,y}$	Power level in MW at which the CSP storage unit is charged during hour (q, d, t)
$storageCSPout_{g,z,q,d,t,y}$	Power level in MW at which the CSP storage unit is discharged during hour (q, d, t)
$surplus_{z,q,d,t,y}$	Surplus generation in MW
$sysCO2backstop_y$	Annual system CO2 emissions above the system emissions

	constraint cap $Sy_emission_cap_y$
$trans_{z,z2,q,d,t,y}$	Active power in MW flowing from z to $z2$
$unmetDem_{z,q,d,t,y}$	Unmet demand in MW (or equivalently violation of the load balance constraint)
$unmetH2External_{z,q,y}$	Unmet quantity of hydrogen in MW (or equivalently violation of the external demand for hydrogen balance constraint)
$unmetRes_{z,y}$	Violation of the planning reserve constraint in MW
$unmetSResZo_{z,q,d,t,y}$	Violation of the zonal/regional spinning reserve constraint in MW
$unmetSResSY_{q,d,t,y}$	Violation of the system-level spinning reserve constraint in MW
Variables for modeling objective function	
$anncapex_{g,y}$	Annualized capex for new generation capacity installed with varying total constructions cost when built in year y carried through planning horizon
$anncapexH2_{h,y}$	Annualized capex for new electrolyzer capacity installed with varying total constructions cost when built in year y carried through planning horizon
$carboncost_{z,y}$	Carbon tax payments by generators
$CO2backStop_y$	Amount of emissions above meeting target in tons
$CO2backCost_y$	Penalty for missing to comply with carbon constraints in USD
$CostTransmissionAdditions_{z1,z2,y}$	Cost of new transmission capacity in USD
$Curtailment_{z,y}$	Curtailed energy in MWh
$fixedcost_{z,y}$	Fixed Operation and Maintenance Cost along with capital payments in constant prices
$npvcost$	Net present value of power system cost over the whole planning horizon; objective function that optimization model tries to minimize
$reservecost_{z,y}$	Cost to procure spinning reserves
$totalcost_{z,y}$	Annual system cost in constant prices
$SurplusCost_{z,y}$	Cost of surplus energy to the system in USE
$Surplus_{z,y}$	Surplus energy in MW
$usecost_{z,y}$	Damage/economic loss in constant prices because of unmet demand
$usrcost_{z,y}$	Penalty in constant prices for unmet spinning reserve requirements
$variablecost_{z,y}$	Variable cost including fuel and variable operation and maintenance cost in constant prices

2.3 Parameters

$Availability_{g,q}$	Availability of unit g to generate power in quarter q
----------------------	---

$Availability_{h,q}$	Availability of electrolyzer h to produce hydrogen in quarter q
$Annual_built_limit_y$	Maximum amount of MW allowed to be built per year
$CapCost_{NG,y}$	Capital cost of unit g in USD \$ or other monetary unit per MW
$CapCostH2_{NH,y}$	Capital cost of electrolyzer h in USD \$ or other monetary unit per MW
$CapCredit_{g,y}$	Percentage of generator's stable (firm) capacity per year
$CapexTrajectories_{DC,y}$	Capital cost variation index for new generators to indicate how new construction costs change during the planning horizon
$CapexTrajectories_{DH,y}$	Capital cost variation index for new electrolyzers to indicate how new construction costs change during the planning horizon
$Carbon_emission_f$	Equivalent tons of CO_2 emitted per MMBTU of fuel consumed
$Carbon_tax_y$	Carbon price in USD\$ per equivalent tons of CO_2
$CostH2Unserved$	Penalty for unserved hydrogen (\$/mmBTU)
$Commission_year_g$	Earliest commission year for generators
$Commission_yearH2_h$	Earliest commission year for electrolyzers
$CO2backstopPenalty$	Penalty per ton of CO_2 above expected target
$CostOfCO2backstop$	Cost of the last technology the system can build to reduce CO_2 to the emissions caps set in the parameters, in \$ per equivalent tons
CRF_{NG}	Capital Recovery factor for new generators
CRF_{NH}	Capital Recovery factor for new electrolyzers
CRF_{NL}	Capital Recovery factor for new transmission lines
$CSP_storage$	Maximum CSP storage capacity in MWh
$CurtailmentPenalty$	Penalty per MWh of VRE curtailed
$Demand_{z,q,d,t,y}$	Hourly load level in MW in hour t , day d , quarter q and year y
$DiscreteCap_g$	Discrete capacity of the unit for generator g
$DRate_y$	Discount rate; real or nominal if cost parameters in real or nominal terms respectively
$Duration_{q,d,t,y}$	Duration of each time slice (block) in hours
$EEfactor_{z,y}$	Demand scaling factor for energy efficiency measures
$EfficiencyH2_h$	Electrolyzer efficiency in mmBTU of hydrogen produced per MWh of electricity
$ExternalH2_{z,q,y}$	External demand for hydrogen per quarter and year (hydrogen produced for use outside the power sector) (mmBTU)
$FieldEfficiency_{CSP}$	Efficiency of the CSP solar field
$FixedOM_{g,y}$	Fixed Operation and Maintenance Cost in USD \$ or other monetary unit per MW per generator capacity
$FixedOM_H2_{g,y}$	Fixed Operation and Maintenance Cost in USD \$ or other monetary unit per MW of electrolyzer capacity

$FuelPrice_{f,y,z}$	Fuel price in USD \$/MMBTU
$GenCap_g$	Generator's existing/available capacity in MW at initial year
$H2Cap_h$	Electrolyzer's existing/available capacity in MW at initial year
$GenCost_{g,f,y}$	Generation variable cost (fuel and VOM) in USD \$ or other monetary unit per MWh
Gen_zone_g	Contains the zone index of the zone the generator belongs to
$HeatRate_{g,f}$	Heat Rate in MMBTU/MWh
$Life_{NG}$	Operating life for new generators
$LifeH2_{NH}$	Operating life for new electrolyzers
$LineCAPEX_{z1,z2}$	CAPEX of line connecting zones z1 and z2
$LossFactor_{z,z2,y}$	"Linearized" loss factor in % of active power flowing on transmission line
$MaxCapital$	Maximum amount of annualized capital payments in USD\$ billion over the horizon
$MaxExport$	Maximum share of the hourly price driven export in terms of demand
$MaxFuelOff_{f,y}$	Maximum amount of fuel f (in million MMBTU) that can be consumed in year y
$MaxImport$	Maximum share of the hourly price driven import in terms of demand
$MaxNewCap_{NG}$	Maximum capacity to be built over the horizon in MW
$MaxTransLine_{z,z2}$	Maximum number of transmission lines built over the horizon
$MinCapFac_g$	Minimum capacity factor (to reflect minimum load requirements)
$OverLoadFactor_g$	Overload factor of generator g , as %, of capacity
PRM_z	Planning reserve margin per zone z
$PwrREH2_{z,q,d,t,y}$	Renewable power used for hydrogen production in MW
$PwrREGrid_{z,q,d,t,y}$	Renewable power which was used for electricity end use sectors (MW)
$RampDn_g$	Ramp-down capability of generator g , as %, of capacity installed
$RampUp_g$	Ramp-up capability of generator g , as %, of capacity installed
$RampDnH2_h$	Ramp-down capability of electrolyzer h , as %, of capacity installed
$RampUpH2_h$	Ramp-up capability of generator h , as %, of capacity installed
$REPwr2H2_{RE,f,q,d,t,y}$	Renewable generation that is finally headed for hydrogen production (MW)
$REPwr2Grid_{RE,f,q,d,t,y}$	Renewable generation that is finally headed toward electricity end use sectors (MW)
$ResCost_g$	Cost to provide reserves in USD \$ or other monetary unit per MWh
$ResOffer_g$	Percentage of the generator's unit qualify as a reserve offer

<i>RESVoLL</i>	Violation penalty of planning reserve requirement in \$/other monetary unit per MW
<i>Retirement_year_{EG}</i>	Latest retirement year for existing generators
<i>Retirement_year_{EH}</i>	Latest retirement year for existing electrolyzers
<i>ReturnRate_y</i>	Discount factor at the starting year of stage ending at year <i>y</i>
<i>RPprofile_{g,RE,q,d,y,t}</i>	Renewable generation profile in % of installed (rated) capacity
<i>SolarMultipleCSP</i>	CSP output to solar field ratio
<i>SResSY_y</i>	System-level spinning reserve constraint in MW
<i>SResZo_{z,y}</i>	Zonal/regional spinning reserve constraint in MW
<i>StageDuration_y</i>	Duration of a stage represented by year <i>y</i> in years
<i>StartYear</i>	First year of the horizon
<i>StartYearGen_g</i>	First year of generator's operation
<i>Storage_efficiency_{st,y}</i>	Efficiency of storage unit <i>st</i> (per charging cycle)
<i>Maximum_Storage_Energy_{st,y}</i>	Maximum energy capability of storage unit
<i>SurplusPenalty</i>	Penalty per MWh of surplus energy
<i>Sys_emission_cap_y</i>	Cap on <i>CO</i> ₂ emissions within the system at year <i>y</i> in equivalent tons
<i>Topology_{z,z2}</i>	Network topology: contains 0 for non-existing lines and 1 or -1 to define the direction of positive flow over the line
<i>TransLimit_{z,z2,q,y}</i>	Transfer limits by quarter (seasonal) and year between zones
<i>TransCapPerLine_{z,z2}</i>	Transmission capacity of each new line in MW
<i>TurbineEfficiency_{CSP}</i>	Efficiency of the CSP power block
<i>VarOM_{g,y}</i>	Variable Operation and Maintenance Cost in USD \$ or other monetary unit per MWh electricity produced
<i>VarOM_H2_{h,y}</i>	Variable Operation and Maintenance Cost in USD \$ or other monetary unit per mmBTU of hydrogen produced
<i>VOLL</i>	Penalty/Economic loss consider per MWh of unmet demand
<i>VREforecastingerror</i>	Percentage error in VRE forecast [used to estimated required amount of spinning reserve]
<i>WACC</i>	Weighted Average Cost of Capital
<i>WeightYear_y</i>	Weight on years
<i>Co _emission_cap_{z,y}</i>	Cap on <i>CO</i> ₂ emissions within country <i>c</i> and year <i>y</i> in equivalent tons
<i>zone_index_z</i>	Index of zone <i>z</i> , unique number assigned to zone <i>z</i>

3 Model formulation

Objective function and its components

$$npvcost = \sum_{z,y} ReturnRate_y * WeightYear_y * totalcost_{z,y} \quad (1)$$

$$\begin{aligned} totalcost_{z,y} = & fixedcost_{z,y} + variablecost_{z,y} + reservecost_{z,y} + usecost_{z,y} + usrcost_{z,y} \\ & + carboncost_{z,y} + CO2backCost_y + TradeCost_{z,y} \\ & + CostTransmissionAdditions_{z1,z2,,y} + CurtailmentCost_{z,y} \\ & + SurplusCost_{z,y} + usrcostH2_{z,y} \end{aligned} \quad (2)$$

$$\begin{aligned} fixedcost_{z,y} = & \sum_{g \in NG \cap NDC} CRF_{NG} * CapCost_{NG,y} * cap_{g,y} + \sum_{g \in NG \cap DC} AnnCapex_{NG,y} + \\ & \sum_g FixedOM_{g,y} * cap_{g,y} + \sum_{h \in NH \cap NDH} CRF_{NH} * CapCostH2_{NH,y} * \\ & cap_{h,y} + \sum_{h \in NH \cap DH} AnnCapexH2_{NH,y} + \sum_h FixedOMH2_{h,y} * capH2_{h,y} \end{aligned} \quad (3)$$

$$AnnCapex_{g \in DC,y} = CFR_{g \in DC} * CapCost_{g \in DC} * CapexTrajectories_{g \in DC,y} * build_{g \in DC,y}$$

$$AnnCapexH2_{h \in DH,y} = CFR_{h \in DH} * CapCost_{h \in DH} * CapexTrajectories_{h \in DH,y} * build_{h \in DH,y}$$

$$\begin{aligned} variablecost_{z,y} = & \sum_{g \in G,f,q,d,t} GenCost_{g,f,y} * Duration_{q,d,t,y} * gen_{g,f,q,d,t,y} \\ & + \sum_{h \in H,q,d,t} VOM_H2_{h,y} * EfficiencyH2_h * vH2PwrIn_{h,g,q,d,t,y} \\ & * pHours_{q,d,t,y} \end{aligned} \quad (4)$$

$$reservecost_{z,y} = \sum_{g \in Z,q,d,t} ResCost_g * Duration_{q,d,t,y} * reserve_{g,q,d,t,y} \quad (5)$$

$$usecost_{z,y} = \sum_{q,d,t} VOLL * Duration_{q,d,t,y} * unmetDem_{z,q,d,t,y} \quad (6)$$

$$\begin{aligned} usrcost_{z,y} = & \sum_{q,d,t} RESVoLL * unmetRes_{z,y} \\ & + \sum_{z,q,d,t,y} Duration_{q,d,t,y} * SRESVoLL * unmetSResZo_{z,q,d,t,y} \\ & + \sum_{q,d,t} Duration_{q,d,t,y} * SRESVoLL * unmetSResSY_{q,d,t,y} \end{aligned} \quad (7)$$

$$\begin{aligned}
carboncost_{z,y} = & \sum_{g \in Z,f,q,d,t} Duration_{q,d,t,y} * carbon_{tax_y} * HeatRate_{g,f} * carbon_{emission_f} \\
& * gen_{g,f,q,d,t,y} + \sum_{z \in C} CO2backstop_{z,y} * CostOfCO2backstop \\
& + SysCO2backstop_y * CostOfCO2backstop
\end{aligned} \tag{8}$$

$$CO2backCost_y = CO2backstopPenalty * CO2backstop_y \tag{9}$$

$$\begin{aligned}
YearlyTradeCost_{z,y} \\
= & \sum_{q,d,t} (ImportPrice_{z,q,d,t,y} - ExportPrice_{z,q,d,t,y}) * TradePrice_{z,q,d,t,y} \\
& * Duration_{q,d,t,y}
\end{aligned} \tag{10}$$

$$CurtailmentCost_{z,y} = CurtailmentPenalty * Curtailment_{z,y} \tag{11}$$

$$\begin{aligned}
CostTransmissionAdditions_{z1,z2,y} \\
= & \sum_z AdditionalTransfer_{z,z2,y} * CRF_{NL} * LineCAPEX_{z1,z2}
\end{aligned} \tag{12}$$

$$SurplusCost_{z,y} = SurplusPenalty * Surplus_{z,y} \tag{13}$$

$$usrcostH2_{z,y} = \sum_q CostH2Unserved * unmetH2External_{z,q,y}$$

Transmission network constraints

$$\begin{aligned} \sum_{g \in Z, f} gen_{g,f,q,d,t,y} - \sum_{z2} trans_{z,z2,q,d,t,y} + \sum_{z2} (1 - LossFactor_{z,z2,y}) * trans_{z2,z,q,d,t,y} \\ + storage_{out,z,q,d,t,y} - storage_{inj,z,q,d,t,y} + unmetDem_{z,q,d,t,y} \\ - surplus_{z,q,d,t,y} + importPrice_{z,q,d,t,y} - exportPrice_{z,q,d,t,y} \\ - \sum_{h \in map(h,z)} H2PwrIn_{h,q,d,t,y} = Demand_{z,q,d,t,y} * Eefactor_{z,y} \end{aligned} \quad (14)$$

$$trans_{z,z2,q,d,t,y} \leq TransLimit_{z,z2,q,y} + AdditionalTransfer_{z,z2,y} * TransCapPerLine_{z,z2} \quad (15)$$

$$AdditionalTransfer_{z,z2,y} = AdditionalTransfer_{z,z2,y-1} + BuildTransmission_{z,z2,y}$$

$$AdditionalTransfer_{z,z2,y} = AdditionalTransfer_{z2,z,y}$$

$$\begin{aligned} \sum_{z \in C, q, d, t} (importPrice_{z,q,d,t,y} * Duration_{q,d,t,y}) \\ \leq \sum_{z \in C, q, d, t} (Demand_{z,q,d,t,y} * Duration_{q,d,t,y} * Eefactor_{z,y}) \\ * MaxPriceImShare_{y,c} \end{aligned} \quad (16)$$

$$\sum_{z \in C} (importPrice_{z,q,d,t,y}) \leq \sum_{z \in C} (Demand_{z,q,d,t,y} * MaxImport) \quad (17)$$

$$\sum_{z \in C} (exportPrice_{z,q,d,t,y}) \leq \sum_{z \in C} (Demand_{z,q,d,t,y} * MaxExport) \quad (18)$$

System requirements

$$\sum_g in_z reserve_{g,q,d,t,y} + unmetSResSY_{z,q,d,t,y} + \sum_{z2} (TransLimit_{z,z2,q,y} + AdditionalTransfer_{z,z2,y} * TransCapPerLine_{z,z2} - trans_{z2,z,q,d,t,y}) \geq SResSY_{z,y} + 0.2 * \quad (19)$$

$$\begin{aligned} \sum_{VRE \text{ in } z} gen_{VRE,f,q,d,t,y} \quad \forall z, q, d, t, y \\ reserve_{g,q,d,t,y} \leq cap_{g,y} * ResOffer_g \end{aligned} \quad (20)$$

$$\begin{aligned} \sum_{g \in Z} cap_{g,y} * CapCredit_{g,y} + unmetRes_{z,y} \\ + \sum_{z2} \sum_q (TransLimit_{z,z2,q,y} + AdditionalTransfer_{z,z2,y} \\ * TransCapPerLine_{z,z2} - trans_{z2,z,q,d,t,y}) \\ \geq (1 + PRM_z) * \max_{q,d,t} Demand_{z,q,d,t,y} \quad \forall z, y \end{aligned} \quad (21)$$

Generation constraints

$$\sum_f gen_{g,f,q,d,t,y} + reserve_{g,q,d,t,y} \leq (1 + OverLoadFactor_g) * cap_{g,y} \quad (22)$$

$$reserve_{g,q,d,t,y} \leq cap_{g,y} * ResOffer_g \quad (23)$$

$$\sum_f gen_{g,f,q,d,t-1,y} - \sum_f gen_{g,f,q,d,t,y} \leq cap_{g,y} * RampDn_g \quad \forall t > 1 \quad (24)$$

$$\sum_f gen_{g,f,q,d,t,y} - \sum_f gen_{g,f,q,d,t-1,y} \leq cap_{g,y} * RampUp_g \quad \forall t > 1 \quad (25)$$

$$\sum_f gen_{g,f,q,d,t,y} \geq MinCapFac_g * cap_{g,y} \quad \forall d \in M \quad (26)$$

$$\sum_{f,d,t} Duration_{q,d,t,y} * gen_{g,f,q,d,t,y} \leq Availability_{g,q} * \sum_{d,t} Duration_{q,d,t,y} * cap_{g,y} \quad (27)$$

Renewable generation

$$gen_{g,f,q,d,t,y} \leq RPprofile_{g,RE,q,d,y,t} * cap_{g,y} \quad \forall RE \notin CSP \quad (28)$$

Concentrated Solar Power (CSP) Generation

$$storageCSP_{g,z,q,d,t,y} \leq cap_{g,y} * CSP_storage \quad \forall map(g, CSP) \quad (29)$$

$$genCSP_{g,z,q,d,t,y} = RPprofile_{z,RE \in CSP,q,d,t} * cap_{g,y} * \frac{SolarMultipleCSP}{TurbineEfficiency_{CSP} * FieldEfficiency_{CSP}} \quad (30)$$

$$\sum_{f \in CSP} gen_{g,f,q,d,t,y} \leq cap_{g,y} \quad (31)$$

$$\sum_{f \in CSP} genCSP_{g,z,q,d,t,y} * FieldEfficiency_{CSP} - storageCSPinj_{g,z,q,d,t,y} + storageCSPout_{g,z,q,d,t,y} = \frac{gen_{g,f,q,d,t,y}}{TurbineEfficiency_{CSP}} \quad \forall g, z, q, d, t, y \quad (32)$$

$$storageCSP_{g,z,q,d,t,y} = storageCSP_{g,z,q,d,t-1,y} + storageCSPinj_{g,z,q,d,t,y} - storageCSPout_{g,z,q,d,t,y} \quad (33)$$

Time consistency of power system additions and retirements

$$cap_{g \in EG,y} = cap_{EG,y-1} + build_{EG,y} - retire_{EG,y} \quad \forall (y, g \in EG), \forall ord(y) > 1 \quad (34)$$

$$cap_{g \in EG,y} = GenCap_g \quad \forall (y, g \in EG): (ord(y) = 1 \wedge StartYear > Commission_year_g) \quad (35)$$

$$cap_{g \in EG,y} = GenCap_g \quad \forall (y, g \in EG): (Commission_year_g \geq StartYear \wedge ord(y) \geq Commission_year_g \wedge ord(y) < Retirement_year_{EG}) \quad (36)$$

$$cap_{g \in EG,y} = 0 \quad \forall (y, g \in EG): (ord(y) \geq Retirement_year_{EG}) \quad (37)$$

$$cap_{g \in NG,y} = cap_{NG,y-1} + build_{NG,y} \quad \forall ord(y) > 1 \quad (38)$$

$$cap_{g \in NG,y} = build_{NG,y} \quad \forall ord(y) = 1 \quad (39)$$

$$cap_g \leq GenCap_g \quad (40)$$

$$cap_{g,y} = 0 \quad \forall (y, g): (ord(y) < Commission_year_g) \quad (41)$$

Storage constraints

$$storage_{st,z,q,d,t=1,y} = Storage_efficiency_{st} * storage_inj_{g,z,q,d,t=1,y} - gen_{st,z,q,d,t=1,y} \quad \forall t = 1 \quad (42)$$

$$storage_{st,z,q,d,t>1,y} = storage_{st,z,q,d,t-1,y} + Storage_efficiency_{st} * storage_inj_{st,z,q,d,t-1,y} - gen_{st,z,q,d,t-1,y} \quad \forall t > 1 \quad (43)$$

$$storage_inj_{stp,q,d,t,y} = Cap_{so,y} * pStoragePVProfile_{so,q,d,t,y} \quad (44)$$

$$storage_inj_{st,q,d,t,y} \leq Cap_{st,y} \quad (45)$$

$$storage_{st,q,d,t,y} \leq Storage_Capacity_{st,y} \quad (46)$$

$$storage_inj_{st,q,d,t,y} - storage_inj_{z,q,d,t-1,y} \leq Cap_{st,y} * RampDn_{st} \quad \forall t > 1 \quad (47)$$

$$storage_inj_{st,q,d,t-1,y} - storage_inj_{z,q,d,t,y} \leq Cap_{st,y} * RampUp_{st} \quad \forall t > 1 \quad (48)$$

$$Storage_Capacity_{EG,y} = Storage_Capacity_{EG,y-1} + Build_Storage_Capacity_{EG,y-1} - Retire_{Storage_Capacity_{EG,y-1}} \quad \forall y > 1 \quad (49)$$

$$Storage_Capacity_{NG,y} = Storage_Capacity_{NG,y-1} + Build_Storage_Capacity_{NG,y-1} \quad \forall y > 1 \quad (50)$$

$$Storage_Capacity_{NG,y} = Build_Storage_Capacity_{NG,y-1} \quad \forall y = 1 \quad (51)$$

$$StorageCapacity_{st,y} \geq Cap_{st,y} \quad (52)$$

$$StorageCapacity_{st,y} \leq Maximum_Storage_Energy_{st,y} + Maximum_Storage_Energy_{CSP,y} \quad (53)$$

$$\sum_y, Build_Storage_Capacity_{EG,y} \leq Maximum_Storage_Energy_{EG,y} \quad (54)$$

--

Investment constraints

$$\sum_y build_{g \in NG, y} \leq MaxNewCap_{NG} \quad (55)$$

$$\sum_y build_{g \in NG, y} \geq MinNewCap_{NG} \quad (56)$$

$$\sum_y build_{g \in EG, y} \leq GenCap_{g \in EG} \quad \forall g: (Commission_year_g > StartYear) \quad (57)$$

$$build_{g \in NG, y} \leq Annual\ built\ limit_y * WeightYear_y \quad (58)$$

$$build_{g \in NG, y} = DiscreteCap_{g \in NG} * bulitCapVar_{g \in NG, y} \quad \forall g: (DiscreteCap_g \geq 0) \quad (59)$$

$$retire_{g \in NG, y} = DiscreteCap_{g \in NG} * retireCapVar_{g \in NG, y} \quad \forall g: (DiscreteCap_g \geq 0) \quad (60)$$

$$build_{g, y} = 0 \quad \forall y: (ord(y) \leq Commission_year_g) \quad (61)$$

$$fuel_{z, f, y} \leq MaxFuelOff_{f, y} \quad (62)$$

$$fuel_{z, f, y} = \sum_{g \in Z, q, d, t} Duration_{q, d, t, y} * HeatRate_{g, f} * gen_{g, f, q, d, t, y} \quad (63)$$

$$\sum_{y, g \in NG} ReturnRate_y * pweight_y * CRF_{NG} * CapCost_{NG, y} * cap_{g, y} \leq MaxCapital \quad (64)$$

Environmental policy

$$emissions_Zo_{z, y} = \sum_{g \in Z, q, d, t} gen_{g, f, q, d, t, y} * HeatRate_{g, f} * carbon_{emission_f} * Duration_{q, d, t, y} \quad (65)$$

$$emissions_Zo_{z, y} \leq Zo_emission_cap_{y, z} \quad (66)$$

$$emissions_{z, y} = \sum_{g, q, d, t} gen_{g, f, q, d, t, y} * HeatRate_{g, f} * carbon_{emission_f} * Duration_{q, d, t, y} \quad (67)$$

$$emissions_{z, y} \leq Sy_emission_cap_y \quad (68)$$

CCS retrofits

$$build_{ng, y} - retire_{eg, y} \leq 0 \quad \forall ng, eg \in GG \quad (69)$$

Green hydrogen production

$$capH2_{eh, y} = capH2_{eh, y-1} + buildH2_{eh, y} - retireH2_{eh, y} \quad \forall ord(y) > 1 \quad (70)$$

$$capH2_{nh, y} = capH2_{nh, y-1} + buildH2_{nh, y} \quad \forall ord(y) > 1 \quad (71)$$

$$capH2_{h,y} = H2Cap_{h \in EH,y} + buildH2_{h,y} - retireH2_{EH,y} \quad \forall ord(y) = 1 \quad (72)$$

$$\sum_{y \in Y} buildH2_{eh,y} \leq H2Cap_{eh} \quad (73)$$

$$\sum_{y \in Y} buildH2_{nh,y} \leq MaxNewCapH2_{nh} \quad (74)$$

$$buildH2_{nh,y} = UnitSizeH2_{nh} \times buildCapVarH2_{nh,y} \quad (75)$$

$$retireH2_{eh,y} = UnitSizeH2_{eh} \times retireCapVarH2_{eh,y} \quad (76)$$

$$\begin{aligned} \sum_{d,t} (H2PwrIn_{h,q,d,t,y} \times Duration_{q,d,t,y}) \\ \leq AvailabilityH2_{h,q} \times CapH2_{h,y} \times \sum_{d,t} Duration_{q,d,t,y} \end{aligned} \quad (77)$$

$$\begin{aligned} \sum_z (ExternalH2_{z,q,y} - unmetH2External_{z,q,y} + fuelH2Quarter_{z,q,y}) \\ \leq \sum_{z,h \in map(h,z)} (H2PwrIn_{h,q,d,t,y} \times Duration_{q,d,t,y} \times EfficiencyH2_h) \end{aligned} \quad (78)$$

$$\sum_z (fuelH2Quarter_{z,q,y}) = fuelH2_{z,y} \quad (79)$$

$$H2PwrIn_{h,q,d,t,y} \leq CapH2_{h,y} \quad (80)$$

$$H2PwrIn_{h,q,d,t-1,y} - H2PwrIn_{h,q,d,t,y} \leq CapH2_{h,y} \times RampDnH2_h \quad (81)$$

$$H2PwrIn_{h,q,d,t,y} - H2PwrIn_{h,q,d,t-1,y} \leq CapH2_{h,y} \times RampUpH2_h \quad (82)$$

$$Gen_{RE,f,d,t,y} = REPwr2Grid_{RE,f,d,t,y} + REPwr2H2_{RE,f,d,t,y} \quad (83)$$

$$\sum_{z,h \in map(h,z)} (H2PwrIn_{h,q,d,t,y}) = PwrREH2_{z,q,d,t,y} \quad (84)$$

$$\sum_{RE,f,z \in (map(RE,f),map(RE,z))} (REPwr2H2_{RE,f,q,d,t,y}) = PwrREH2_{z,q,d,t,y} \quad (85)$$

$$\sum_{RE,f,z \in (map(RE,f), map(RE,z))} (REPwr2Grid_{RE,f,q,d,t,y}) = PwrREGid_{z,q,d,t,y} \quad (86)$$

4 Description of the model

EPM can be used (i) as a least-cost capacity expansion model to optimize investments in the power sector over a specified horizon that could extend from a few years to decades or as (ii) an economic dispatch model to optimize the operation of a specified power system over a relatively short period compared to Capacity Expansion Model (1-week to 1-year). Like every optimization software the problem is mathematically described through an objective function and several accompanying equations (system constraints). System constraints are used to describe mathematically the physical part of the system, laws of physics, policy targets and other realities specific to the system (for example fuel availability). More detailed description of system constraints can be found on subsequent sections.

4.1 Indices and Sets

All sets used in the formulation can be classified in two major categories: time and power system-related. Four sets belong to the first category: D, Q, T, Y which represent different time scales considered in the model: days, quarters, hours and years. Hour is the smallest unit of time used in this formulation and we could use the same formulation using just one set for time containing as many hours as the set $D \times Q \times T \times Y$. It is convenient though to keep all the four sets since they reveal some fundamental assumptions of the model: (1) days are used to reflect the chronological sequence of the time slices used for ramping and storage constraints as we will further explain in the constraints section; (2) quarters are used to reflect seasonality in the load patterns, the availability of thermal power units and the thermal limits of transmission lines; (3) years are used to represent annual trends on demand growth and keep track of the lifetime of units; and finally (4) hours are commonly used as the smallest time unit in long term models since the day ahead scheduling models schedule generation units on an hourly basis.

Three sets are power system related. Set G includes all generating units. Depending on the size of the system, we might decide to use set g to model individual units of the power system for a small system or aggregated units that represent multiple units of the same technology for a large system. We use the term technology to refer to different technologies or different fuels used: e.g. coal steam turbines, natural gas combined cycle, natural gas combustion turbines, wind farms, solar photovoltaic panels, geothermal, hydropower and diesel generators. Depending on the resources available in a country, some technologies might not be present. As the model stands now, elements of set G are mapped to

sets F and Z , which stand for fuel and zones respectively. Set Z is one of the major sets used in power systems since the power system is a network and physical laws (widely known as Kirchhoff's laws) govern the flow of power over the transmission lines. Given that, a set such as set Z which captures the spatial dimension of the system is necessary. At the finest granularity, set Z might contain buses of the power system but in case we model larger systems, set Z might contain zones of a power system or even countries. Note that the modeler usually decides on the spatial granularity based on the presence of common regulatory rules or pricing schemes in a zone or/and based on the congestion observed on transmission lines connecting adjacent regions. Finally, set F includes the different fuels used and we model it to keep track of the consumption of different fuels since for certain fuels domestic upper bounds on consumption might apply and/or issues of energy security might be involved in case of imported fuels. In addition, different types of fuels have different carbon content and lead to different emissions of carbon dioxide, which are important to track in case environmental policies exist.

4.2 Objective function

The objective function in this model minimizes the total system cost including violation/penalty terms for constraints that are not met. All generation, transmission and emissions costs of the system are considered including: (1) Fixed costs including annualized capital cost payments for new generators⁴ and any additional transmission capacity and fixed operation and maintenance costs (2) variable costs including the fuel cost and any variable operation and maintenance costs (3) cost to procure spinning reserves (4) carbon tax payments and (5) penalties for unmet demand and unmet capacity and spinning reserve requirements, breaching the emissions and renewable rejection at the system or the zonal level.

In the case of a least-cost expansion model the objective function typically represents total system costs over the study horizon. Depending on the type of the problem (for example linear or non-linear) the user chooses an appropriate solver to minimize the value of the objective function. Total system costs are broken into several components the most important of which are the fixed and variable costs of electricity generation. CAPEX costs typically dominate fixed costs while fuel costs dominate variable costs of generation. In systems with very high shares of renewables VOM costs are higher than fuel costs. Other types of system costs included in the objective function are costs of reserve provision by generators, the cost of carbon (when applicable), the cost of electricity trading (it could be a net benefit if revenues from electricity trade exceed costs) and the cost of transmission enhancements.

When EPM is used as a dispatch model, there are no capex costs in the objective function. In dispatch mode, the model does not consider unit commitment related decisions in its current form, albeit minimum generation constraints and ramp rates are modeled.

4.3 Load approximation

The size of the optimization problem (i.e. the number of decision variables to be determined) depends significantly on the study horizon and the optimization step. An EPM user needs to find the balance

⁴ Capital costs of existing generators are considered sunk costs and are not included in the objective function.

between precision and computational effort (practicality) to complete the analysis. This is especially important when little gains in precision impact significantly the running time. In most cases representing the full year at the hourly level (8760 hours) over a study horizon that extends to multiple years, usually 15 to 20 years, is neither practical nor required to assess in detail the capacity mix of most supply options. For capacity expansion analyses, the load is typically approximated through a number of representative days which are assumed to repeat several times throughout the calendar year. Such representative days are carefully selected to capture the hourly profile (including the peak) of electricity demand on a monthly or quarterly basis. It is common practice to select the following three types of representative days: (i) an average day which is assumed to be representative of the bulk of electricity production (repeats several times over a quarter (or month)), (ii) a typical day of high demand (which also need to include the peak) and (iii) a typical day of low demand. The last two types of days typically represent a much smaller portion of the quarterly (or monthly) demand compared to the average day. The EPM input file includes tabs where a user can provide the hourly profile of each of the selected typical days on the “Demand” tab. Also, a user needs to provide the relative contribution of each typical day on total quarterly/monthly/seasonal demand represented as number of repeats over a typical quarter (or month), in the “Duration” tab.

The above-mentioned representation of the load has been widely used in power sector analysis to estimate the capacity mix of conventional supply options both efficiently and accurately. However, application of this methodology has weaknesses in power systems with very high shares of variable renewable energy (VRE) sources. In such systems detailed representation of chronology might have significant impact on results. VRE generation profiles are inputs on EPM. Selection of a representative day for load need to be accompanied by the respective VRE generation profile that corresponds to the selected day. This requires extensive data processing.

A weakness of selecting representative demand days is that the analysis might undersize electricity storage, the simulation of which requires full description of chronology. This is especially true in the case multi-hour and seasonal storage. For that reason, running EPM on dispatch mode (full chronology) of a full (selected) year might be required to supplement capacity expansion analysis.

An EPM user can represent demand in detail for each year of the study horizon or provide annual projections (single value) of demand and peak demand together with typical day profiles over a base (historical) year. When the latter option is activated, the model uses the above data to calculate the projected demand in detail.

4.4 Value of Lost Load

The Value of Lost Load (VoLL) is an economic penalty used to ensure that the model will invest to meet the demand and avoid unserved energy if it is economic to do so. The VoLL represents the damage on the local economy from missing to supply 1MWh of demand. Without assigning VoLL matching supply and demand would be economically suboptimal as it is cheaper to fall short on supply rather than paying for the cost of electricity generation and associated investments. Thus, the selection of VoLL need to be based on research. There is no universally acceptable Value of Lost Load and different methods have been applied to estimate a reasonable value for the unserved energy. Typical values used in developed economies vary between 4,000 and 40,000\$/MWh while in developing countries between 500 and 10,000\$/MWh [1].

VoLL affects both the total-non served energy in the system and peaking supply investment decisions. In some cases, allowing for some loss of load might be preferable compared to investing into supply options that will only operate for a tiny amount of the annual time. The user will need to calibrate the model considering the local realities and system rules since the maximum allowed number of hours failing to meet demand is constrained by reliability criteria. When a value for VoLL is not available in global literature the EPM user can select a reported VoLL from a country with similar economic characteristics. Alternatively, the assessment of VoLL can be based on analytical expressions or based on the assumption that demand is sensitive to electricity price. In the latter case VoLL can take the value of cost of generation of the most expensive asset in the system which are usually backup diesel generators [2].

Typically, back-up generators are fueled with expensive diesel, and if they are used in the mode just described, the system VoLL would take the value of the variable cost of these generators (~500 \$/MWh).

4.5 Transmission network constraints

Kirchhoff's laws are physical laws governing the flows over transmission lines in a network. According to first Kirchhoff law, also known as KCL (Kirchhoff's Current Law), the sum of injections in a node should equal zero. In our formulation, KCL corresponds to equation (14). There, we can see that the power provided by generators and storage should be equal to demand (minus the unmet demand) plus/min outflows/inflows from the node to adjacent nodes. This equation additionally includes price-driven import $importPrice_{z,q,d,t,y}$ and export $exportPrice_{z,q,d,t,y}$ volumes from external zones. The demand is adjusted with the potential improvements in the energy efficiency using $EEfactor_{z,y}$ parameter.

$$\begin{aligned} \sum_{g \in Z,f} gen_{g,f,q,d,t,y} - \sum_{z2} trans_{z,z2,q,d,t,y} + \sum_{z2} (1 - LossFactor_{z,z2,y}) * trans_{z2,z,q,d,t,y} \\ - storage_{inj_{z,q,d,t,y}} + unmetDem_{z,q,d,t,y} - surplus_{z,q,d,t,y} \\ + importPrice_{z,q,d,t,y} - exportPrice_{z,q,d,t,y} = Demand_{z,q,d,t,y} * EEFactor_{z,y} \end{aligned} \quad (14)$$

Note that the second Kirchhoff Law (or widely known as KVL- Kirchhoff's Voltage Law) is valid for power systems but for reduced power system, it might not apply depending on the method of network reduction followed. In this particular formulation, KVL is not considered.

Another important feature of our model relates to modeling of transmission losses. We model transmission losses as a percent reduction of the imported electricity at each node. In particular, term $\sum_{z2} (1 - LossFactor_{z,z2,y}) * trans_{z2,z,q,d,t,y}$ at equation (14) models injections to node z and we can see how the loss factor reduces the amount of energy imported. On the contrary, the outflow is fully considered at the origin node of the network: $\sum_{z2} trans_{z,z2,q,d,t,y}$.

Another common constraint for transmission networks refers to the capacity limits of transmission lines. In particular, as equation (15) implies, the flow over a specific line cannot exceed a certain limit, which is defined either by thermal limit of the line or upper bounds imposed by reliability considerations. Note that we model flows over a particular transmission line with two positive variables, one for each direction. Please observe that the transmission limit parameter might change per year to reflect planned upgrades and the additions to the transmission network. There are also maximum built limits of additional transmission capacity between zones due to budget or technical constraints. Moreover, the

transmission limit differs per season since ambient temperature affects the capacity available for power transfers.

$$trans_{z,z2,q,d,t,y} \leq TransLimit_{z,z2,q,y} + AdditionalTransfer_{z,z2,y} * TransCapPerLine_{z,z2} \quad (15)$$

The modelled zones, countries and regions are usually not isolated from the external systems and markets not included in the model. Therefore, we allow trading with these zones at time-varying predefined price. Equations (16), (17) and (18) define the limits of this exchange. Equation (16) sets the annual limit on the energy imported from external zones as a percentage of the annual demand using parameter $MaxPriceImShare_{y,c}$. Equations (17) and (18) introduce the hourly limit on the imported and exported power, defined as a percentage of hourly demand.

$$\begin{aligned} \sum_{z \in c, q, d, t} (importPrice_{z,q,d,t,y} * Duration_{q,d,t,y}) \\ \leq \sum_{z \in c, q, d, t} (Demand_{z,q,d,t,y} * Duration_{q,d,t,y} * EEfactor_{z,y}) * MaxPriceImShare_{y,c} \\ \sum_{z \in c} (importPrice_{z,q,d,t,y}) \leq \sum_{z \in c} (Demand_{z,q,d,t,y} * MaxImport) \\ \sum_{z \in c} (exportPrice_{z,q,d,t,y}) \leq \sum_{z \in c} (Demand_{z,q,d,t,y} * MaxExport) \end{aligned}$$

4.6 System requirements

In our formulation, we model two products that the system operator might require generators to provide during operation: (1) energy and (2) spinning reserves. Per NERC's definition, spinning reserves refers to "unloaded generation that is synchronized and ready to serve additional demand" [3]. Note that more products exist, especially in organized U.S. wholesale markets such as non-spinning reserves or flex ramp. Operating reserves (spinning and non-spinning reserves) provide the capability above firm system demand required to provide for regulation, load-forecasting error, equipment forced and scheduled outages and local area protection [3]. Moreover, under the spinning reserves different products might exist with respect to the response times required etc.

The amount of spinning reserve required depends on several factors that the planner/operator considers such as the load level and the associated forecasting error, the forecasting error attached to the renewable generation and the size of the largest unit committed on the system to be able to accommodate N-1 outages. Equation (19) indicates that spinning reserve can be provided by interconnections. On top of system-wide reserve requirements, zonal requirements apply to accommodate for outages on transmission lines connecting adjacent regions/zones/nodes.

$$\begin{aligned} \sum_{g \text{ in } z} reserve_{g,q,d,t,y} + unmetSResSY_{z,q,d,t,y} + \sum_{z2} (TransLimit_{z,z2,q,y} + \\ AdditionalTransfer_{z,z2,y} * TransCapPerLine_{z,z2} - trans_{z2,z,q,d,t,y}) \geq SResSY_{z,y} + \\ VREforecastingerror * \sum_{VRE \text{ in } z} gen_{VRE,f,q,d,t,y} \quad \forall z, q, d, t, y \end{aligned} \quad (19)$$

Planners usually consider a planning reserve margin (PRM) to account for forecasting error in demand projections. Typical values for the PRM vary between 10-15%. Equation (21) indicates that interconnections can be accounted for as reserve margin. Note that intermittent units do not contribute towards the planning reserve constraint at their full capacity but at a fraction specified by the planner

(using the Capacity Credit parameter), e.g. in the U.S. markets this fraction is calculated based on available historical data as the capacity factor during a set of peak hours [4].⁵

$$\sum_{g \in Z} cap_{g,y} * CapCredit_{g,y} + unmetRes_{z,y} + \sum_{z \in Z} \sum_q TransLimit_{z2,z,q,y} \geq (1 + PRM_z) * \max_{q,d,t} Demand_{z,q,d,t,y} \quad \forall z, y \quad (21)$$

4.7 Generation constraints

Spinning reserve capacity is modeled as it is a very important dimension of a system flexibility and ability to integrate VRE.

$$\sum_f gen_{g,f,q,d,t,y} + reserve_{g,q,d,t,y} \leq (1 + OverLoadFactor_g) * cap_{g,y} \quad (22)$$

Equation (22) assures that the power generated by the unit along with the spinning reserves provided by the same unit do not exceed the unit's capacity.⁶ Note that the capacity is augmented by an overload factor. This factor is typically 10% for those generators that can handle overload conditions for a short period of time, and zero for those generators that cannot handle such conditions.

Given that spinning reserve products are usually defined with respect to response time of generator to a certain dispatch signal, only a certain percentage of the generator's unit qualify as a reserve offer. We capture this characteristic in the model through equation (23).

$$reserve_{g,q,d,t,y} \leq cap_{g,y} * ResOffer_g \quad (23)$$

Ramping constraints acknowledge that the generation units have inertia in changing their outputs and differences in generation outputs between consecutive hours should be constrained by the ramping up and down capabilities of the unit.

$$\sum_f gen_{g,f,q,d,t-1,y} - \sum_f gen_{g,f,q,d,t,y} \leq cap_{g,y} * RampDn_g \quad \forall t > 1 \quad (24)$$

$$\sum_f gen_{g,f,q,d,t,y} - \sum_f gen_{g,f,q,d,t-1,y} \leq cap_{g,y} * RampUp_g \quad \forall t > 1 \quad (25)$$

⁵ For the first model runs, capacity credit was considered at full capacity. This will be modified in the next runs but it is not expected to change significantly the model results.

⁶ Depending on the scope of the project, the dispatch constraint (22) might be slightly different. For example, ReEDS model implemented by NREL [16] treats quick start capacity service provided by a generator in the same way as spinning reserves under constraint (22) and on top of that, it accounts for planned and forced outages by considering average outage rates.

Another important feature of generators is the minimum load. The minimum load can either be determined based on technical specifications provided by the manufacturer or be calculated as an “economic” minimum beyond which the unit can provide energy economically. The minimum load constraint is really important for unit commitment and requires the use of binaries variables that make sure the constraint is enforced when the unit is on. However, in generation expansion models operations are approximated through a simple dispatch model for representative hours of the year. In this case, if deemed important, an approximation of the minimum load constraint can be applied for the thermal units for which it is relevant. The constraint can be applied for a subset of the days modeled and the constraint is activated only for those days or to all days. Constraint (26) is forcing all units to generate power equal to at least their minimum loading levels for specific days in the year.

$$\sum_f gen_{g,f,q,d,t,y} \geq MinCapFac_g * cap_{g,y} \quad \forall d \in M \quad (26)$$

Generating units require maintenance every year. Given that, we should consider the units as unavailable for certain periods during the year. In this particular application, we consider a uniform availability factor per quarter to account for maintenance.

$$\sum_{f,d,t} Duration_{q,d,t,y} * gen_{g,f,q,d,t,y} \leq Availability_{g,q} * \sum_{d,t} Duration_{q,d,t,y} * cap_{g,y} \quad (27)$$

4.8 Renewable generation modeling

Renewable generation differs from conventional units in that its output is, to a certain extent, uncontrollable and intermittent. The power generated by renewables such as wind or solar depends on wind velocity or solar irradiation. Collecting historical data that records weather information (such as wind speed, temperature, wind direction, etc.) or the power generation output by installed renewables at specific locations, analysts usually employ statistical methods such as k-means to reduce the number of hours required to approximate the intermittent nature of renewables [5]. In this particular application, the generation profile for each renewable energy technology (such as wind or solar PV) is defined by the hourly capacity factor, in a year, of a generic power plant of each type, modeled at a specified location. Then, given this hourly profile for a year, we choose the amount of days modeled based on the days selected for the load approximation (see Section 4.3), i.e., the renewable profile during the 12 days in the year selected, for load, is maintained.

$$gen_{g,f,q,d,t,y} \leq RPprofile_{g,RE,q,d,y,t} * cap_{g,y} \quad (28)$$

Note that the renewable profile is highly dependent on the region/location the resource is located. This formulation implicitly models that aspect since g might have different elements for the same generation technology at different locations.

4.9 Concentrated Solar Power (CSP) modeling

CSP technology modeling differs from other renewable technologies due to the complexity derived by its storage capabilities. The CSP configuration considered in this model consists of two integrated subsystems; these include the thermal storage system, and power cycle. Thermal storage is modeled using a simple energy balance approach that includes charging and discharging energy. The power cycle model provides a simple mechanism for modeling the conversion from thermal energy output from the solar field, and thermal storage into electrical energy.

$$storageCSP_{g,z,q,d,t,y} \leq cap_{g,y} * CSP_storage \quad \forall map(g, CSP) \quad (29)$$

Equation (29) indicate that at any time the CSP storage level cannot exceed its storage capability.

$$genCSP_{g,z,q,d,t,y} = RPprofile_{z,RE \in CSP,q,d,t} * cap_{g,y} * \frac{SolarMultipleCSP}{TurbineEfficiency_{CSP} * FieldEfficiency_{CSP}} \quad (30)$$

The power output of the solar panel is calculated by multiplying the nameplate capacity of the CSP power plant, the capacity factor of the system, and the solar multiple, then, dividing this by the turbine and solar field efficiencies (Equation (30)).

$$\sum_{f \in CSP} gen_{g,f,q,d,t,y} \leq cap_{g,y} \quad (31)$$

Equation (31) indicate that all the power output produced by CSP generators at any given zone, cannot exceed the nameplate capacity. Finally, Equations (32) and (33) detail the power balance formulations for the power cycle and thermal storage subsystems.

$$\sum_{f \in CSP} genCSP_{g,z,q,d,t,y} * FieldEfficiency_{CSP} - storageCSPinj_{g,z,q,d,t,y} + storageCSPout_{g,z,q,d,t,y} = \frac{gen_{g,f,q,d,t,y}}{TurbineEfficiency_{CSP}} \quad \forall g, z, q, d, t, y \quad (32)$$

$$storageCSP_{g,z,q,d,t,y} = storageCSP_{g,z,q,d,t-1,y} + storageCSPinj_{g,z,q,d,t,y} - storageCSPout_{g,z,q,d,t,y} \quad (33)$$

4.10 Time consistency of power system additions and retirements

We use constraint (34) to track the capacity in consecutive years. In particular, generation capacity at year y equals capacity at previous year plus any investment minus any retirement at year y .

$$cap_{g \in EG,y} = cap_{EG,y-1} + build_{EG,y} - retire_{EG,y} \quad \forall ord(y) > 1 \quad (34)$$

Several more constraints are formulated to fix the capacity at pre-specified levels in certain years.

- The first constraint sets the capacity of the existing units in the first year to the predefined capacity, if the commissioning year is earlier then the first year of the optimization horizon.

$$cap_{g \in EG, y} = GenCap_g \quad \forall (y, g \in EG): (ord(y) = 1 \wedge StartYear < Commission_year_g) \quad (35)$$

- The second constraint sets the capacity of the existing units to the predefined capacity, for all the years exceeding the commissioning year, but before the defined retirement year.

$$cap_{g \in EG, y} = GenCap_g \quad \forall (y, g \in EG): (Commission_year_g \geq StartYear \wedge ord(y) \geq Commission_year_g \wedge ord(y) < Retirement_year_{EG}) \quad (36)$$

- The third constraint forces the capacity of the existing units to 0, when the retirement year is reached.

$$cap_{g \in EG, y} = 0 \quad \forall (y, g \in EG): (ord(y) \geq Retirement_year_{EG}) \quad (37)$$

- The fourth constraint states that the total capacity of a new generator equals the capacity of that new generator built the previous year plus the capacity to be built the current year:

$$cap_{g \in NG, y} = cap_{NG, y-1} + build_{NG, y} \quad \forall ord(y) > 1 \quad (38)$$

- The fifth constraint forces the capacity at the first year of the horizon for new units to be equal to the capacity installed that year

$$cap_{g \in NG, y} = build_{NG, y} \quad \forall ord(y) = 1 \quad (39)$$

- The sixth constraint limits the capacity to the predefined available or installed capacity of the unit.

$$cap_g \leq GenCap_g \quad (40)$$

- The seventh constraint forces the capacity of planned and candidate units at zero for years preceding the commission year of the unit.

$$cap_{g, y} = 0 \quad \forall (y, g): (ord(y) < Commission_year_g) \quad (41)$$

4.11 Storage modeling

Economically efficient storage in power systems has mainly been pumped hydro storage for a considerable number of years. Nowadays, more storage technologies such as battery energy storage systems are being added to the power system providing energy and/or reserve services.

Storage is modeled differently compared to conventional units since it requires two more variables: (1) one variable to keep track of the storage level (*storage*) and (2) one variable (*storageinj*) to model the injection of energy into the storage unit. The variable *gen* used to track the generator output of conventional units is used to account for the output of a storage unit when it is discharged. Moreover, the chronological sequence of the time slices is important to make sure that the simulated operation is feasible e.g., we cannot discharge a storage unit if charging of the unit has not preceded. Finally, storage of energy requires the conversion of electricity to another form of energy e.g., mechanical for flywheels or chemical for fuel cells and common batteries. The conversion of one form of energy to another involves losses that we should consider in our models. Equations (30) and (31) represent the storage balance equations on the first hour of a representative day and any other hour of a representative day respectively:

$$\begin{aligned} storage_{st,z,q,d,t=1,y} &= Storage_efficiency_{st} * storage\ inj_{g,z,q,d,t=1,y} \\ &- gen_{st,z,q,d,t=1,y} \quad \forall t = 1 \end{aligned} \quad (42)$$

$$\begin{aligned} storage_{st,z,q,d,t>1,y} &= storage_{st,z,q,d,t-1,y} + Storage_efficiency_{st} * storage\ inj_{st,z,q,d,t-1,y} \\ &- gen_{st,z,q,d,t-1,y} \quad \forall t > 1 \end{aligned} \quad (43)$$

The first equation (42) makes sure that the battery is discharged at the start of each representative day. The choice for not making a link between different days and thus having the battery discharged at the start of each day is because the standard formulation of the model works with representative days (e.g. peak load, minimum load, and average load days) instead of actual days. Equation (43) in turn keeps track of the energy stored in the storage unit between consecutive hours of the same day: the energy stored in the unit at time slice t ($t > 1$) equals the energy stored in the unit at time slice $t-1$ plus any injection reduced by the storage efficiency minus any discharge at time t . In case the storage unit is linked to a PV plant, the injection is given by the PV profile:

$$storage\ inj_{stp,q,d,t,y} = Cap_{so,y} * pStoragePVProfile_{so,q,d,t,y} \quad (44)$$

Two additional constraints are used to make sure that the operation of storage in each time slice is feasible considering the peak storage capacity (in MW):

$$storage\ inj_{st,q,d,t,y} \leq Cap_{st,y} \quad (45)$$

$$storage_{st,q,d,t,y} \leq Storage_Capacity_{st,y} \quad (46)$$

The first constraint ensures that the hourly energy injection is lower or equal than the maximum hourly energy injection. The second constraint limits the total amount of energy (in MWh) stored to the maximum storage level (in MWh). Storage outputs are constrained by the general equation (22).

Furthermore, the storage units can be charged or discharged at a rate, which cannot exceed a specific value. To model this behavior, we include constraints (47) and (48) respectively.

$$storage\ inj_{st,q,d,t,y} - storage\ inj_{z,q,d,t-1,y} \leq Cap_{st,y} * RampDn_{st} \quad \forall t > 1 \quad (47)$$

$$storage\ inj_{st,q,d,t-1,y} - storage\ inj_{z,q,d,t,y} \leq Cap_{st,y} * RampUp_{st} \quad \forall t > 1 \quad (48)$$

Three equations keep track of the deployed energy storage capacity (in MWh) of the storage units:

$$Storage_Capacity_{EG,y} = Storage_Capacity_{EG,y-1} + Build_Storage_Capacity_{EG,y-1} - Retire_Storage_Capacity_{EG,y-1} \quad \forall y > 1 \quad (49)$$

$$Storage_Capacity_{NG,y} = Storage_Capacity_{NG,y-1} + Build_Storage_Capacity_{NG,y-1} \quad \forall y > 1 \quad (50)$$

$$Storage_Capacity_{NG,y} = Build_Storage_Capacity_{NG,y-1} \quad \forall y = 1 \quad (51)$$

Equation (49) keeps track of the energy storage capacity for existing storage units whereas equations **Error! Reference source not found.** and (51) are used for new storage units.

Three final constraints limit the total deployed energy storage capacity (in MWh):

$$StorageCapacity_{st,y} \geq Cap_{st,y} \quad (52)$$

$$StorageCapacity_{st,y} \leq Maximum_Storage_Energy_{st,y} + CSP_storage_{CSP,y} \quad (53)$$

$$\sum_y Build_Storage_Capacity_{EG,y} \leq Maximum_Storage_Energy_{EG,y} \quad (54)$$

The first constraint ensures that the deployed energy storage (in MWh) is larger than the peak storage capacity (in MW) at any time.⁷ The second constraint limits the total deployed energy storage (in MWh) to the maximum deployable energy storage (in MWh). Equation (54) constrains the deployment of existing storage units for which the commission year is later than the starting year of the modelling horizon.

4.12 Investment constraints

Planners consider several constraints when they decide on a generation investment plan. Common constraints refer to budget, land use, scheduling of new construction and consumption of specific fuels for energy security considerations.

Constraint (55) usually reflects land use considerations, regulation that imposes an upper bound on capacity of specific technologies or simply resource potential (e.g., for wind there is a finite amount of locations where wind farms can provide the capacity factor modeled).

⁷ This constraint is necessary to ensure that the model does not simply build additional storage units just to meet planning or spinning reserve requirements.

$$\sum_y build_{g \in NG, y} \leq MaxNewCap_{NG} \quad (55)$$

Constraint (58) is usually employed to reflect practical limitations on construction and spread the construction of new units more uniformly over time. For example, it seems pretty unrealistic that the whole system capacity can be built in one year.

$$build_{g \in NG, y} \leq Annual\ built\ limit_y * WeightYear_y \quad (58)$$

Constraint (62) imposes an upper bound on fuel consumption. This upper bound might correspond to the fuel reserves a country might have at its disposal or the capacities of refinement units or importing units such as size of LNG terminals. Constraint (63) simply estimates the fuel consumption. Note that in case we want to reduce the number of variables in our model, we can get rid of the fuel variable since it is defined in terms of the generation variable.

$$fuel_{z,f,y} \leq MaxFuelOff_{f,y} \quad (62)$$

$$fuel_{z,f,y} = \sum_{g \in Z, q, d, t} Duration_{q,d,t,y} * HeatRate_{g,f} * gen_{g,f,q,d,t,y} \quad (63)$$

Constraint (64) represents a budget constraint. It limits the capital expenses withdrawal to be lower than a pre-specified amount. In this formulation, we assume that the *MaxCapital* parameter is similar to the maximum debt payments that a power system planner can do over the horizon.

$$\sum_{y, g \in NG} ReturnRate_y * pweight_y * CRF_{NG} * CapCost_{NG,y} * cap_{g,y} \leq MaxCapital \quad (64)$$

An alternative constraint that addresses the same concern but relies on different information is expressed by constraint (69). In that case, the planner does not know the maximum amount of debt payments that the power plant owners might do but he has a good understanding of the maximum capital available to the system for investment. In that case, the sum of the overnight capital expenditure is not allowed to exceed this known budget.

$$\sum_{g \in NG, y} CapCost_{g,y} \leq MaxBudget \quad (69)$$

4.13 Environmental policy

Environmental concerns often lead policymakers to adoption of caps on the amount of emissions of certain pollutants. Several countries have announced targets to reduce their carbon dioxide emissions below particular amounts. Given that the power system is a big contributor to carbon dioxide emissions, power system CO_2 caps are usually enforced at the country or at multi-country level. Constraints (65) and (66) impose the caps decided by regulators. Note that constraints (67) and (68) are modeled as “hard” constraints with no violation allowed. We would like to note that environmental policy might

impose caps on more pollutants beyond carbon dioxide. For example, in the USA certain caps apply on NO_x and SO_x emissions. The dual variables of constraint (66) and (68) provide valuable information to the planner since they estimate the additional cost the planner would have to bear to decrease the amount of the pollutant by an infinitely small amount. In other words, the dual prices of (66) and (68) provide an approximation of the prices that the planner would pay per unit of pollutant if slightly stricter regulation would be applied.

$$emissions_Co_{c,y} = \sum_{g \in c,q,d,t} gen_{g,f,q,d,t,y} * HeatRate_{g,f} * carbon_{emission_f} * Duration_{q,d,t,y} \quad (65)$$

$$emissions_Co_{c,y} - CO2backstop_{c,y} \leq CO_emission_cap_{c,z} \quad (66)$$

$$emissions_y = \sum_c emissions_Co_{c,y} \quad (67)$$

$$emissions_y - SysCO2backstop_y \leq Sys_emission_cap_y \quad (68)$$

Another policy mechanism related to carbon emissions is a carbon tax (less popular than the cap-and-trade system at present). We model the carbon tax as part of the objective function. Note that the carbon tax does not correspond to an actual cost for the society since it is a transfer payment for emitters to the government. It reflects an actual cost, though, only if it attempts to monetize the public health cost and the damage to the environment. However, it reflects an actual cost for power system since generators would probably have to pay the tax to the government and that explains why it is part of the objective function (8).

$$carboncost_{z,y} = \sum_{g \in Z,f,q,d,t} Duration_{q,d,t,y} * carbon_{tax_y} * HeatRate_{g,f} * carbon_{emission_f} * gen_{g,f,q,d,t,y} \quad (8)$$

4.14 Modelling CCS

CCS technology investments can be modelled on EPM as either greenfield or retrofits of existing coal plants. A new CCS plant is modelled similarly as any other coal plant with its techno-economic parameters adjusted. Typically, CCS plants are modelled to have lower efficiency and higher CAPEX and OPEX compared to conventional coal plants. The environmental impact of CO2 storage is modelled through setting fuel burned on a CCS plant to have reduced emissions factor compared to conventional coal (typically reduction 90% to 95%).

CCS retrofits are modeled as a sudden change on the techno-economic parameters of an existing plant subject to some CAPEX incurred. More specifically, this change is modelled as a decommissioning of the existing plant and a simultaneous commissioning of the retrofitted plan. The retrofit CAPEX and the updated techno-economic parameters need to be carefully selected based on international experience. The mathematical addition of CCS retrofit on EPM is done through equation 69 below:

$$build_{ng,y} - retire_{eg,y} \leq 0 \quad \forall ng, eg \in mapGG_{eg,ng} \quad (69)$$

Set mapGG containing the relevant pairs of the existing plant (candidate for retrofit) and the new plant (retrofitted plant) is automatically generated through the inputs development phase. The relevant connection (pairs of plants) is made by the EPM user on excel tab named LinkedComDecom

4.15 Modelling of Green Hydrogen

Modeling of production of green hydrogen is possible on EPM. When the option is activated investments on electrolyzers are optimized similarly as with supply side technologies. The user provides the relevant technoeconomic inputs for electrolyzers (CAPEX, OPEX, efficiency, project lifetime etc). There are two potential uses of green hydrogen in EPM:

1. Re-circulation in the power sector. In that case green hydrogen acts as a carrier of renewable energy providing long-term storage to the system. Hydrogen can be burned in modified new CCGTs/OCGTs at 100% per volume or burned in existing dual fuel generators (future versions of EPM will include option for pre-specified per volume mix of hydrogen with natural gas to be burned in conventional turbines). The user needs to define hydrogen as an additional fuel with zero cost and zero emissions factor since it is produced within the system and not imported. Costs of hydrogen are indirectly accounted through the additional investments and relevant OPEX within the system. Also, techno-economic characteristics of new hydrogen turbines need to reflect published figures based on international experience
2. Used in other end-use sectors (residential, industry, transport). This is modeled as an external demand of green hydrogen to be satisfied subject to some penalty for each mMBTU of unserved hydrogen demand. EPM simulates power sector operation and investment requirements to satisfy external hydrogen demand but it does not model any subsequent use outside the power sector. The costs for external hydrogen production are accounted for in the model and thus the power sector is assumed to bear the costs of decarbonization of other sectors.

Description of green hydrogen formulation on EPM follows below:

Investments on electrolyzers are mathematically constrained to dictate time consistency of capacity additions and retirements, similarly as with generators:

Generation capacity of existing electrolyzer capacity at year y equals capacity at previous year plus any investment minus any retirement at year y where y is not the first year of the study horizon (equation 70)

$$capH2_{eh,y} = capH2_{eh,y-1} + buildH2_{eh,y} - retireH2_{eh,y} \quad \forall ord(y) > 1 \quad (34)$$

Equation (71) is the same constraint as (70) modified for candidate capacity

$$capH2_{nh,y} = capH2_{nh,y-1} + buildH2_{nh,y} \quad \forall ord(y) > 1 \quad (38)$$

Equation (72) constraints total electrolyzer capacity at the end of the first year of the optimization horizon. At the beginning of horizon existing capacity is the one defined in the inputs section and by the end of year the capacity is adjusted by any optimized additions or retirement throughout the year.

$$capH2_{h,y} = H2Cap_{h \in EH,y} + buildH2_{h,y} - retireH2_{EH,y} \quad \forall ord(y) = 1 \quad (34)$$

Total electrolyzer capacity that can be developed into mandatory projects (decision for development has been made) can not exceed the maximum electrolyzer capacity provided as input

$$\sum_{y \in Y} buildH2_{eh,y} \leq H2Cap_{eh} \quad (66)$$

Total new electrolyzer capacity that can be developed in candidate projects (decision to be optimized) cannot exceed the maximum electrolyzer capacity provided as input

$$\sum_{y \in Y} buildH2_{nh,y} \leq MaxNewCapH2_{nh} \quad (66)$$

Electrolyzer candidate investments can also be set as integer decisions. In such cases the defined capacity can either be built in one step or stay out of the plan (equation 75). Similarly retirement of existing electrolyzer plant takes place in a single step (76). In equations (75) and (76), variables $buildH2decision_{nh,y}$ and $retireH2decision_{eh,y}$ can only take positive integer values

$$buildH2_{nh,y} = UnitSizeH2_{nh} \times buildCapVarH2_{nh,y} \quad (75)$$

$$retireH2_{eh,y} = UnitSizeH2_{eh} \times retireCapVarH2_{eh,y} \quad (76)$$

Equation 77 puts a constraint on operational time of electrolyzers so that it doesn't exceed their technical availability

$$\sum_{d,t} (H2PwrIn_{h,q,d,t,y} \times Duration_{q,d,t,y}) \leq AvailabilityH2_{h,q} \times CapH2_{h,y} \times \sum_{d,t} Duration_{q,d,t,y} \quad (77)$$

Equation 88 describes the thermal balance of hydrogen within the system (in mmBTU). Total mmBTU of hydrogen produced are equal to hydrogen demand for other sectors minus unmet hydrogen demand plus hydrogen recirculate in the power sector as fuel.

$$\begin{aligned} \sum_z (ExternalH2_{z,q,y} - unmetH2External_{z,q,y} + fuelH2Quarter_{z,q,y}) \\ \leq \sum_{z,h \in map(h,z)} (H2PwrIn_{h,q,d,t,y} \times Duration_{q,d,t,y} \times EfficiencyH2_h) \end{aligned} \quad (78)$$

Equation 79 makes a connection between quarterly an annual amount of hydrogen fuel produced in the system

$$\sum_z (fuelH2Quarter_{z,q,y}) = fuelH2_{z,y} \quad (79)$$

Equation 80 puts a constraint on total amount of electricity drawn by an electrolyzer. Max power absorbed can't be more than the electrolyzer capacity

$$H2PwrIn_{h,q,d,t,y} \leq CapH2_{h,y} \quad (80)$$

Equations 81 and 82 define ramping constraints for electrolyzer operation

$$H2PwrIn_{h,q,d,t-1,y} - H2PwrIn_{h,q,d,t,y} \leq CapH2_{h,y} \times RampDnH2_h \quad (81)$$

$$H2PwrIn_{h,q,d,t,y} - H2PwrIn_{h,q,d,t-1,y} \leq CapH2_{h,y} \times RampUpH2_h \quad (82)$$

Equations 83 sets an energy balance on renewable electricity. Total renewable electricity generated is equal to renewable electricity consumed by end-use sectors and renewable electricity consumed by electrolyzes

$$Gen_{RE,f,d,t,y} = REPwr2Grid_{RE,f,d,t,y} + REPwr2H2_{RE,f,d,t,y} \quad (83)$$

Equation 84 makes a connection between renewable power produced and electricity drawn by electrolyzers

$$\sum_{z,h \in map(h,z)} (H2PwrIn_{h,q,d,t,y}) = PwrREH2_{z,q,d,t,y} \quad (84)$$

Finally, equations 85 and 86 are mapping renewable electricity produced on a zonal basis which is necessary for the architecture of the hydrogen model

$$\sum_{RE,f,z \in (map(RE,f),map(RE,z))} (REPwr2H2_{RE,f,q,d,t,y}) = PwrREH2_{z,q,d,t,y} \quad (85)$$

$$\sum_{RE,f,z \in (map(RE,f),map(RE,z))} (REPwr2Grid_{RE,f,q,d,t,y}) = PwrREGrid_{z,q,d,t,y} \quad (86)$$

5 Limitations of the model

As any model, this model also has some limitations in most practical system applications. Two major limitations of the model are the following:

(1) Transmission network representation: Although in theory, the model can be used to represent the high voltage transmission network down to individual substations, the size limitation of the underlying LP, typically restricts to a small (below 20) number zones and we assume zero congestion within each zone modeled. Transmission expansion is not always an option. The model does not consider voltage angles and DC approximation to power flow in its current form and as such may have unrealistic flows around a loop.

(2) As noted before, the model does not consider unit commitment (startup and shutdown) decisions. It also does not consider heat rate curves associated with thermal units and uses a single average heat rate for each generating unit.

(3) In its main version, the model is deterministic and no uncertainty with respect to assumed parameters etc. is considered.

(4) Hydropower: reservoir optimization is currently not a feature of EPM hence hydropower potential is represented as daily/seasonal/annual energy limits for individual hydro plant/unit. The main version model does not consider any representation of river chain.

(5) In the current model version, coal plant flexibilization or repurposing is not considered.

(6) When the model is used as a capacity expansion tool, the use of representative days and seasons limits the possibility to capture the VRE variability, hence the VRE integration costs are approximated. Similarly, the flexibility costs of the system are approximated as ancillary services are not all fully modeled.

Bibliography

- [1] A. Van Der Welle and B. Van Der Zwaan, "An overview of selected studies on the value of lost load (VOLL)," *Energy Research Centre of the Netherlands (ECN), Amsterdam*, 2007.
- [2] F. De Sisternes, "Investment model for renewable electricity systems (IMRES): an electricity generation capacity expansion formulation with unit commitment constraints," *MIT, Cambridge MA*, 2013.
- [3] NERC, "Glossary of Terms Used in NERC Reliability Standards," [Online]. Available: [http://www.nerc.com/pa/stand/glossary of terms/glossary_of_terms.pdf](http://www.nerc.com/pa/stand/glossary%20of%20terms/glossary_of_terms.pdf). [Accessed 12 December 2016].
- [4] B. Hobbs and C. Bothwell, "System Adequacy with Intermittent Resources : Capacity Value and Economic Distortions," ISO New England, 2016. [Online]. Available: https://www.iso-ne.com/static-assets/documents/2016/09/PSPC09222016_A4_Cindy-Bothwell-Johns-Hopkins-University-System-Adequacy-with-Intermittent-Resources-Capacity-Value-and-Economic-Distortions.pdf. [Accessed 2 October 2016].
- [5] L. Baringo and A. J. Conejo, "Correlated wind-power production and electric load scenarios for investment decisions," *Applied Energy*, vol. 101, pp. 475-482, 2013.
- [6] CESI, "Pan-Arab Regional Energy Trade Platform (PA-RETP). Power Market Study: Gas Consumption Assessment for the Power Sector accross MENA Region," CESI, 2016.
- [7] CESI-Ramboll, "Feasibility Study of the Electrical Interconnection and Energy Trade between Arab Countries," Arab Fund for Economic and Social Development, 2014.
- [8] The World Bank, "Energy Efficiency Study in Lebanon," The World Bank, Washington, DC, 2009.
- [9] Presidency of the Council of Ministries Central Administration of Statistics, "Energy".
- [10] Palestinian Electricity Regulatory Council, "Annual Report 2011,," Palestinian Electricity Regulatory Council, 2011.
- [11] DSIRE, "N.C. Clean Energy Technology Center," DSIRE, [Online]. Available: <http://programs.dsireusa.org/system/program/tables>. [Accessed 12 December 2016].
- [12] A. Hainoun, "Construction of the hourly load curves and detecting the annual peak load of future Syrian electric power demand using bottom-up approach," *International Journal of Electrical Power & Energy Systems*, vol. 31, no. 1, pp. 1-12, 2009.
- [13] P. Khajavi, H. Monsef and H. Abniki, "Load profile reformation through demand response programs using smart grid," in *Proceedings of the International Symposium Modern Electric Power Systems (MEPS), 2010*, 2010.
- [14] A. Khazaei, M. Ghasempour, H. Hoseinzadeh and H. Hooshmandi, "DEMAND MANAGEMENT PROGRAM FOR LARGE INDUSTRIAL CONSUMERS BY USING THE AMI SYSTEM," in *23rd International*

Conference on Electricity Distribution, 2015.

- [15] A. Salimi-Beni, D. Farrokhzad, M. Fotuhi-Firuzabad and S. Alemohammad, "A new approach to determine base, intermediate and peak-demand in an electric power system," in *International Conference on Power System Technology, 2006. PowerCon 2006.*, 2006.
- [16] W. Short, P. Sullivan, T. Mai, M. Mowers, C. Uriarte, N. Blair, D. Heimiller and A. Martinez, "Regional Energy Deployment System (ReEDS)," NREL, 2011.