# Chapter 3
# Methodology

**Sven Teske, Jaysson Guerrero Orbe, Jihane Assaf, Souran Chatterjee, Benedek Kiss, and Diana Ürge-Vorsatz**

**Abstract** The OneEarth Climate Model (OECM), its background, and program architecture are described. How the OECM is broken down into two independent modules to calculate demand and supply is explored. The basic program logic of the MATLAB-based bottom-up demand module, with high technical resolution, is described for various sectors, including the input and output parameters. The description includes numerous figures and tables for both demand and supply modules. The sub-sectors used for the OECM 1.5 °C pathway are listed, including outputs and the areas of use.

The second part of the chapter documents the high-efficiency building (HEB) model of the Central European University, which was used for the global and regional bottom-up analyses of the building sector. Its methodology, including the programme architecture, the workflow, and the equations used, is provided.

**Keywords** Methodology · OneEarth Climate Model (OECM) · MATLAB · High-efficiency building (HEB) model

The Paris Climate Agreement (UNFCCC, 2015) 'notes that … emission reduction efforts will be required … to hold the increase in the global average temperature to below 2 °C above pre-industrial levels…'. The Intergovernmental Panel on Climate Change (IPCC) further quantified the carbon budget to achieve this target in its Sixth Assessment Report of the Working Group (IPCC, 2021). According to the IPCC, a global carbon budget of 400 $GtCO_2$ is required to limit the temperature rise to 1.5 °C, with 67% likelihood, by 2050.

---

S. Teske (✉) · J. G. Orbe · J. Assaf
University of Technology Sydney – Institute for Sustainable Futures (UTS-ISF), Sydney, NSW, Australia
e-mail: sven.teske@uts.edu.au

S. Chatterjee · B. Kiss · D. Ürge-Vorsatz
Central European University, Department of Environmental Sciences and Policy, Budapest, Hungary

S. Teske (ed.), *Achieving the Paris Climate Agreement Goals*,
https://doi.org/10.1007/978-3-030-99177-7_3

To implement these targets, energy and climate mitigation pathways are required. Numerous computer models for the analysis and development of energy and emission pathways have been developed over the last few decades. Many different calculation methods have been established, which mainly differ in the principal task of the model and the level of detail in the GHG emissions and/or energy systems calculated. The various methods of climate-economy modelling use different ways to describe the economy- and climate-relevant parameters as parts of a highly interconnected process (Nikas et al., 2019). In this context, the economy includes all aspects of the energy system and the policy framework, whereas the climate module reflects various GHG emissions from energy-related and non-energy-related processes, such as land use.

A comprehensive review of energy models, focusing on the usability of those models for decision-making, found 'that a better understanding of user needs and closer co-operation between modellers and users is imperative to truly improve models and unlock their full potential to support the transition towards climate neutrality …' (Süsser et al., 2022).

## 3.1 The OneEarth Climate Model

The UN-convened Net-Zero Asset Owner Alliance (NZAOA) is an international group of institutional investors committed to transitioning their investment portfolios to net-zero emissions by 2050 (NZAOA, 2021). Detailed industry sector-based energy scenarios are required to implement those net-zero commitments. On the basis of the OneEarth Climate Model (OECM; Teske et al. 2019a, b), the Institute for Sustainable Futures, University of Technology Sydney (UTS/ISF), in close co-operation with institutional investors, has developed an integrated energy assessment model for industry-specific 1.5 °C pathways, with high technical resolution, for the finance sector. In this article, we describe the detailed methodology and the architecture of the energy model in the 2021 edition of the advanced OneEarth Climate Model (OECM 2.0).

### *3.1.1 The OneEarth Climate Model Architecture*

The OneEarth Climate Model has been developed on the basis of established computer models. The energy system analysis tool consisted of three independent modules:

1. Energy system model (EM): a mathematical accounting system for the energy sector (Simon et al., 2018)
2. Transport scenario model TRAEM (transport energy model) with high technical resolution (Pagenkopf et al., 2019)

3. Power system analysis model [R]E 24/7, which simulates the electricity system on an hourly basis and at geographic resolution to assess the requirements for infrastructure, such as grid connections between different regions and electricity storage types, depending on the demand profile and power generation characteristics of the system (Teske, 2015)

The advanced OneEarth Climate Model, OECM 2.0, merges the energy system model (EM), the transport energy model (TRAEM), and the power system model [R]E 24/7 into one MATLAB-based energy system module. The Global Industry Classification Standard (GICS) was used to define sub-areas of the economy. The global finance industry must increasingly undertake mandatory climate change stress tests for GICS-classified industry sectors in order to develop energy and emission benchmarks to implement the Paris climate protection agreement. This requires very high technical resolution for the calculation and projection of future energy demands and the supply of electricity, (process) heat, and fuels that are necessary for the steel and chemical industries. An energy model with high technical resolution must be able to calculate the energy demand based on either projections of the sector-specific gross domestic product (GDP) or market forecasts of material flows, such as the demand for steel, aluminium, or cement in tonnes per year.

To decarbonize the energy supply, fossil fuels must be phased out and replaced by a renewable energy supply. However, the supply of high-temperature process heat for various production processes cannot yet be fully electrical, and a simple fuel switch from oil, gas, or coal to biomass is also impossible, given the limited availability of sustainable bioenergy (Seidenberger et al., 2010; Farjana et al., 2018). To develop a detailed sector-specific solution, the temperature level required must be considered when developing an energy scenario. An energy model with such high technical resolution can provide detailed results for various industry sectors but requires a highly complex and data-intensive model architecture. Separate modules for the calculation of different sectors of the energy system are not practicable for such high technical resolution because high electrification rates lead to increased sector coupling, and the interactions between sectors cannot be captured if the energy model uses separate modules.

Furthermore, the geographic distribution of the energy demand and supply must be accommodated to calculate the import and export of energy, especially for energy-intensive industries. Finally, the simulation of 100% renewable energy systems requires high time resolution to accommodate the high proportions of variable solar and wind energy.

The MATLAB model has an object-oriented structure and two modules—to calculate demand and supply—that can be operated independently of each other. Therefore, an energy demand analysis independent of the specific supply options or the development of a supply concept based on demand from an external source is possible.

### *3.1.2 The OECM Demand Module*

The demand module uses a bottom-up approach to calculate the energy demand for a process (e.g. steel production) or a consumer (e.g. a household) in a region (e.g. a city, island, or country) over a period of time. One of the most important elements in this approach is the strict separation of the original need (e.g. to get from home to work), how this need can be satisfied (e.g. with a tram), and the kind of energy required to provide this service (in this case, electricity). This basic logic is the foundation for the energy demand calculations across all sectors: *buildings*, *transport*, *services*, and *industry*. Furthermore, the energy services required are defined: electricity, heat (broken down into four heat levels: <100 °C, 100–500 °C, 500–1000 °C, and > 1000 °C), and fuels for processes that cannot (yet) be electrified. Synthetic fuels, such as hydrogen, are part of both the demand module, because electricity is required to produce it, and the supply module.

The energy requirements are assigned to specific locations. This modular structure allows regions to be defined and, if necessary, the supply from other areas to be calculated.

Demand and generation modules are independent and can be used individually or sequentially. Energy demands can be calculated either as synthetic load profiles, which are then summed to annual energy demands, or as annual consumption only, without hourly resolution. Whether or not hourly resolution is selected depends to a large extent on the availability of data. Load profiles, such as those for the chemical industry, are difficult to obtain and are sometimes even confidential.

#### 3.1.2.1 Input Parameters

As in basic energy models, the main drivers of the energy demand are the development of the population and of economic activity, measured in GDP. Figure 3.1 shows the basic methodology of the OECM demand module. Tier 1 inputs are population and GDP by region and sector. 'Population' defines the number of individual energy services, which determines the energy required per capita, and 'economic activity' (in GDP) defines the number of services and/or products manufactured and

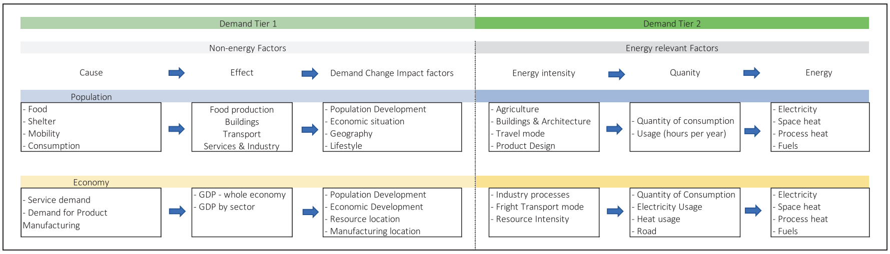


**Fig. 3.1** Tier 1 and tier 2 input parameters for the assessment of energy demand

sold. Tier 1 demand parameters are determined by the effect that a specific service requires. For population, the demand parameters are defined by the need for food, shelter (buildings), and mobility and—depending on the economic situation and/or lifestyle of the population—the demand for goods and services.

Economic activity (measured in GDP) is a secondary input and is directly and indirectly dependent upon the size of the population. However, a large population does not automatically lead to high economic activity. Both population and projected GDP are inputs from external sources, such as the United Nations or the World Bank. Tier 1 input parameters themselves are strictly non-technical. The need to produce food can be satisfied without electricity or (fossil) fuels, just as a service can be provided with physical strength.

Tier 2 demand parameters are energy-relevant factors and describe technical applications, their energy intensities, and the extent to which the application is used. For example, if lighting is required, the technical application 'light bulb' is chosen to satisfy the demand.

In this example, the energy intensity is the capacity of a light bulb, e.g. 100 W. The use of the application (e.g. for 5 h per day) defines the daily demand (5 h × 100 W = 500 Wh per day). The quantity of consumption per year is 365 days at 500 Wh per day = 1825 Wh or 182.5 kWh per year. This very basic and simple principle is used for every application in each of the main sectors: *residential + buildings*, *industry*, and *transport*. These sectors are broken down into multiple sub-sectors, such as aviation, navigation, rail, and road for *transport*, and further into applications, such as vehicle types. The modular programming allows the addition of as many sub-sectors and applications as required.

#### 3.1.2.2 Structure of the Demand Module

Each of the three sectors, *residential and buildings*, *industry*, and *transport*, has standardized sub-structures and applications. The residential sector *R* (first layer) has a list of household types (second layer), and each household type has a standard set of services (third layer), such as 'lighting', 'cooling', and 'entertainment'. Finally, the applications for each of the services are defined (fourth layer), such as refrigerator or freezer for 'cooling'. The energy intensity of each application can be altered to reflect the status quo in a certain region and/or to reflect improvements in energy efficiency. An illustrative example of the layers of the residential sector is shown in Fig. 3.2.

Figure 3.3 shows an example of the model structure of the *industry* sector. In the second layer are different industries—the OECM uses the GICS classification system for industry sub-sectors. The quantity of energy for each of the sub-sectors is driven by either GDP or the projected quantity of product, such as the tonnes of steel produced per year. The market shares of specific manufacturing processes are defined, and each process has a specific energy intensity for electricity, (process) heat, and/or fuels.

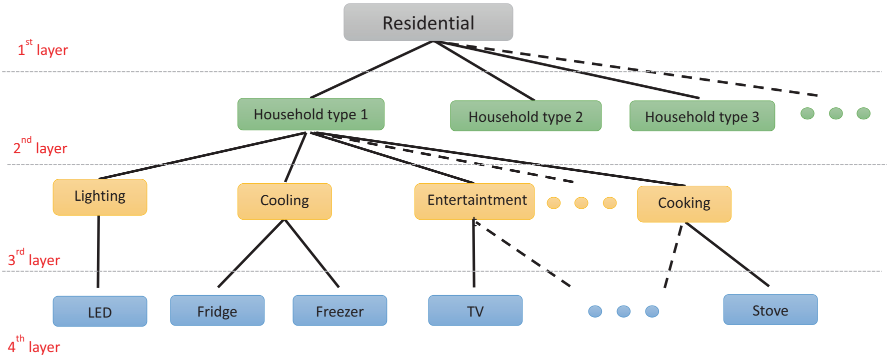


**Fig. 3.2** Residential sector sub-structures

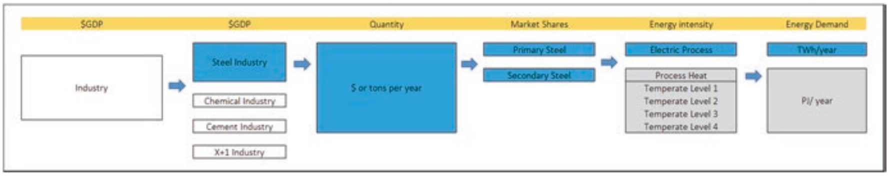


**Fig. 3.3** Calculation of the *industry* energy demand

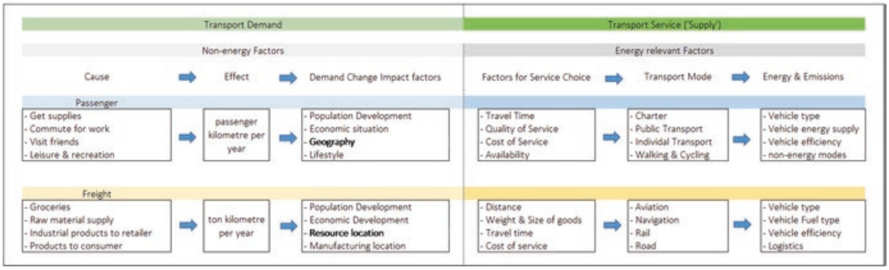


**Fig. 3.4** Calculation of *transport* energy demand

Figure 3.4 shows the structure for the t*ransport* sector. Again, the demand is driven by 'non-energy' factors, such as passenger-kilometres and freight-kilometres, and energy-related factors, such as the transport mode and the energy intensity of the different vehicle options.

### 3.1.2.3 Demand Module Architecture in MATLAB

The demand module is implemented in MATLAB, a widely used programming language for mathematics and science computing. MATLAB allows the integration of a range of tools and databases and has the flexibility to add and develop new functions. Specifically, the model has been developed using an object-oriented programming approach, allowing extensibility and modularity.

Figure 3.5 shows the demand module developed in MATLAB. The demand module encompasses eight classes: (1) demand class, (2) household class, (3) household application class, (4) sub-sector, (5) industry class, (6) industry application class, (7) transport modes, and (8) vehicles class (Fig. 3.6).

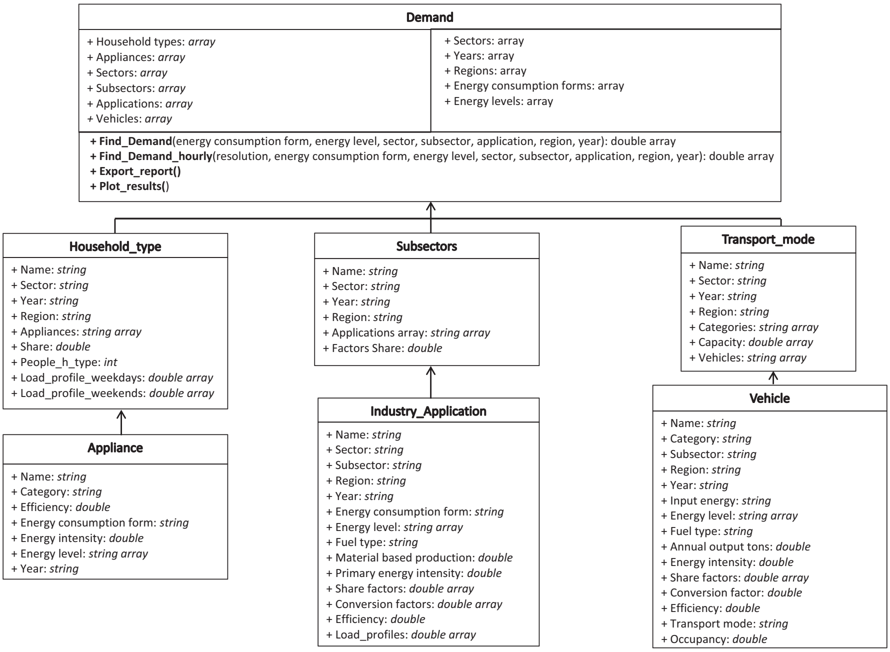
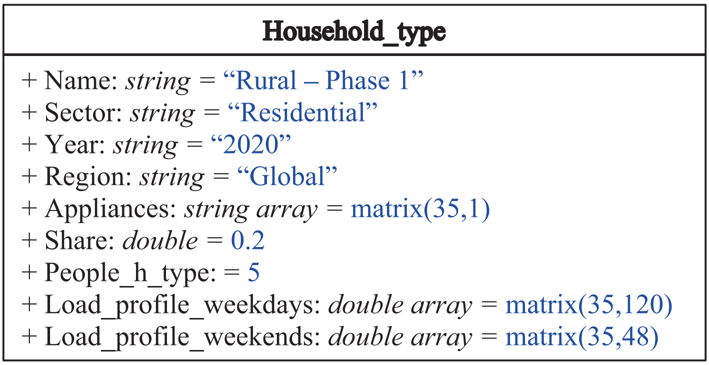


**Fig. 3.5** A unified modelling language (UML) diagram of the demand module in MATLAB, showing its classes, attributes, methods, and associations


**Household_type**

+ Name: *string* = "Rural – Phase 1"
+ Sector: *string* = "Residential"
+ Year: *string* = "2020"
+ Region: *string* = "Global"
+ Appliances: *string array* = matrix(35,1)
+ Share: *double* = 0.2
+ People_h_type: = 5
+ Load_profile_weekdays: *double array* = matrix(35,120)
+ Load_profile_weekends: *double array* = matrix(35,48)

**Fig. 3.6** An example of a *household* type object, showing the assigned attributes

- *Demand class*: This is the main class, which describes the *residential*, *industry*, and *transport* sectors, which are defined by household type, sub-sector, and transport mode classes, respectively. The attributes that define this class include a range of years, energy consumption forms, energy levels, list of sectors, household types, sub-sectors, applications, appliances, and vehicles. The demand class also has two main types of methods: (i) calculation demand methods and (ii) printing results methods. The calculation methods use equations and algorithms to calculate and find the demand. For example, the 'Find Demand' method can be used to find a wide range of calculations and outputs, e.g. the electricity demand of a group of households for a specified year. The calculation method can calculate the demand for single or aggregated sectors, sub-sectors, or applications, for a single year or a range of years, unique or multiple forms of energy consumption, and single or various types of vehicle categories. The printing result methods can be used to export the results into an Excel spreadsheet or to plot the results using the MATLAB interface. Therefore, the outputs of the demand module can be either predefined graphs, tables, or data for a standardized report. See Table 3.1 for a brief description of each method in this class.
- *Household and appliance classes*: These classes are used to define the *residential* sector. The appliance objects are embedded within the household-type objects. Attributes include names, sectors, and regions, which are defined as string inputs (i.e. text or character inputs) or numerical inputs, which are defined as int (i.e. integers) or double (i.e. numeric variables holding numbers with decimal points). Attributes can also include arrays of strings or double values. Array variables are helpful in input time series data, such as load profiles. Because households and appliances have their own classes, this architecture is flexible and allows the addition of households with different attributes and different types of appliances.
- *Sub-sector and industry application classes*: These classes are used to define the *industry* sector. The industry application objects are embedded within the sub-sector objects. As shown in Fig. 3.7, these classes have their own lists of attributes. Therefore, the module developed can accommodate different types of

**Table 3.1** Methods within the demand class

| Type of method | Method | Description |
|---|---|---|
| Calculation | Find Demand() and Find_Demand_hourly() | These methods calculate the annual or hourly aggregated energy demand for the specified region and energy form (i.e. power, heat, or hydrogen). The calculations can be aggregated by sector, sub-sector, transport mode, or any other object class |
| Printing results | Export_report() | This method exports the specified results to external Excel spreadsheets and can be used to print results on predefined report tables |
| Printing results | Plot_results() | This method can be used to plot results using the MATLAB interface |

**Industry_Application**

+ Name: *string* = “Primary steel production”
+ Sector: *string* = “Industrial”
+ Subsector: *string* = “Iron & Steel”
+ Region: *string* = “Global”
+ Year: *string* = “2020”
+ *Energy consumption form: string = “Electricity & Heat”*
+ *Energy level: string array = matrix(2,5)*
+ Material based production: *double* = $1.178 \times 10^3$
+ Primary energy intensity: *double* = $1.184 \times 10^3$
+ Share factors: *double array* = matrix(1,3)
+ Conversion factors: *double array* = matrix(1,3)
+ Efficiency: *double* = 0.98
+ Load_profiles: *double array* = matrix(1,168)

**Fig. 3.7** An example of an *industry* application object, showing the assigned attributes

sub-sectors (e.g. steel, cement, etc.) and incorporate various types of applications under each sub-sector.

- *Transport modes and vehicle classes*: These classes are used to define the *transport* sector. The vehicle objects are embedded within the transport mode objects. Therefore, multiple types of transport modes can be defined, such as aviation and navigation, as well as various types of vehicles, such as planes and cruise ships.

Figures 3.6 and 3.7 show the high-level class definitions for *residential* and *industry* sub-sector objects, respectively. The blue-marked text indicates the defined value for each attribute. For example, one household object with five residents is defined by the name ‘Rural–Phase 1’ and has a list of 35 appliance objects, defined with a string array. It is assigned a share factor for 2020 of 0.2, which means that 20% of the households in that specific region and year are defined by this type of household and its attributes. Furthermore, 24 h load profiles are defined for each application for every day, with numerical arrays. For example, weekend load profiles have a size of 35 rows and 48 columns, representing 35 applications and 24 time slots for each weekend day.

The object-oriented architecture allows all these input attributes to be updated or modified easily. These attributes can also be read from a predefined Excel spreadsheet. This facilitates a data input process that follows the array structure, such as the load profile.

Figure 3.7 shows an example of an industrial application object that belongs to the sub-sector *iron and steel*. In this case, the energy consumption form is defined as electricity and heat, which means that it considers the electrical and heat demand. The ‘share factors’ represent the portions of the demand assigned to electricity and heat. The energy-level array also allows the predefined network to which the application is connected to be defined, as well as the temperature levels. In this particular case, the demand is defined based on the total annual primary energy intensity and the material-based production, which are 1184 GJ/tonnes and 1178 Mt, respectively, for the specified region and year. The input and output units must be predefined when the MATLAB modules are initialized. Other attributes that can be assigned are conversion factors, such as from primary energy to the final energy via an efficiency factor.

Additional attributes and methods can be defined for each class if required and the data are available. Therefore, the demand module class can be extended by defining new classes, attributes, and methods.

## *3.1.3 The OECM Supply Module*

The supply module consists of three main elements: supply technologies, storage technologies, and the infrastructure for the power supply (capacities of power lines). For the generation of electricity and heat, the programme considers all the technologies of the energy market, from both renewable and non-renewable sources. In addition to the generation of pure electricity and heat, the entire range of combined heat and power systems is included.

Storage technologies include batteries and the use of hydrogen from electrolysers. The calculation of heat storage is possible, but has not yet been used in the OECM scenarios.

A dispatch strategy is defined for electricity and heat generation that reflects market and policy factors. Whether electricity from photovoltaics and onshore and offshore wind turbines have priority dispatch ahead of fossil-fuel power plants and how storage systems are used can be determined. Each technology has a specific conversion efficiency.

Heat generation technologies are also defined by the temperature levels they can provide. For example, residential solar collectors can only supply low-temperature heat and will therefore not be considered for high-temperature process heat (Table 3.2).

The regional energy demand—as defined in the previous section—can be met by neighbouring regions, with importation from or, in the case of oversupply, exportation to them. The extent to which electricity can be imported or exported from one region to another is defined by the capacity of regional interconnections, which represent the available power line capacities.

### 3.1.3.1 The OECM Dispatch Module

The methodology of the dispatch module of the MATLAB-based OECM is based on the previous version of the model (Teske et al. 2019a). The key inputs are related to the supply technologies, storage types, dispatch strategy, and the

**Table 3.2** Example of generation and storage technologies

| Generation | | | Storage | | |
|---|---|---|---|---|---|
| Power plants | Combined heat and power plants | Heating plants | Electrical | Thermal | Hydrogen |
| Hard coal | Hard coal | Coal | Lithium battery | Water tank | Tank |
| Lignite | Lignite | Lignite | Pumped hydro | Molten salt | |
| Gas | Gas | Gas | | | |
| Oil | Oil | Oil | | | |
| Diesel | Biomass | Biomass | | | |
| Biomass | Geothermal | Solar collectors | | | |
| Hydro | Hydrogen | Geothermal | | | |
| Wind | | Hydrogen | | | |
| Photovoltaic | | | | | |
| Concentrated solar (CSP) | | | | | |
| Geothermal | | | | | |
| Solar thermal | | | | | |
| Ocean energy | | | | | |
| Hydrogen | | | | | |

**Table 3.3** Input parameters for the dispatch model

| Input parameter | | |
|---|---|---|
| $L_{Cluster}$ | Load cluster | [MW] |
| $L_{Interconnection}$ | Maximum power-line capacity (import/export) | [MW] |
| $L_{Initial}$ | | [MW] |
| $Cap_{Var.RE}$ | Installed capacity for *variable renewables* | [MW] |
| $Meteo_{Norm}$ | Meteorological data for solar and wind | [MW/MW$_{INST}$] |
| $L_{Post\_Var.RE}$ | Load after *variable renewable* supply | [MW] |
| $Cap_{Storage}$ | *Storage* capacity | [MW] |
| $CapFact_{Max\_Storage}$ | Maximum capacity factor for storage technologies | [h/yr] |
| $L_{Post\_Storage}$ | Load after *storage* supply | [MW] |
| $Cap_{Dispatch}$ | Capacity of *dispatch power plants* | [MW] |
| $CapFact_{Max\_Dispatch}$ | Maximum capacity factor for *dispatch power plants* | [h/a] |
| $L_{Post\_Dispatch}$ | Load after *dispatch power plant* supply | [MW] |
| $Cap_{Interconnection}$ | *Interconnection* capacity | [MW] |

interconnections among regions for possible power exchange (Table 3.3). Different supply technologies can be selected, each with its technical characteristics, including its efficiency, available installed capacity, fuel type, and regional meteorological data (solar radiation or wind speed). Meteorological data define the capacity factors of solar and wind energy generators as their levels of availability at 1-h resolution for an entire year (Table 3.4).

The supply technologies can be either dispatchable (e.g. gas power plants) or non-dispatchable (e.g. solar photovoltaic without storage). The model allows the

**Table 3.4** Output parameters for the dispatch model

| Output parameter | | |
|---|---|---|
| $L_{Initial}$ | Initial load (cluster) | [MW] |
| $L_{Post\_Var.RE}$ | Load after *variable renewable* supply | [MW] |
| $S_{EXECC\_VAR.RE}$ | Access supply *renewables* | [MW] |
| $L_{Post\_Storage}$ | Load after *storage* supply | [MW] |
| $S_{Storage}$ | Storage requirement/curtailment | [MW] |
| $CapFact_{Actual\_Storage}$ | Utilization factor for storage | [h/a] |
| $L_{Post\_Dispatch}$ | Load after *dispatch power plant* supply | [MW] |
| $S_{Dispatch}$ | Dispatch requirement | [MW] |
| $CapFact_{Actual\_Dispatch}$ | Utilization factor for *dispatch power plants* | [h/a] |
| $L_{Post\_Interconnection}$ | Load after *interconnection* supply | [MW] |
| $S_{Interconnection}$ | Interconnection requirement | [MW] |
| $CapFact_{Actual\_Interconnection}$: | Utilization factor for *interconnection* | [h/a] |

**Table 3.5** Technology groups for the selection of dispatch order

| Technology options | Input: assumed order marked with (1) to (4) |
|---|---|
| 1. Variable renewables | Variable renewables (1) |
| 2. Storage | Dispatch generation (3) |
| 3. Dispatch generation | Storage (2) |
| 4. Interconnector | Interconnector (4) |

**Table 3.6** Technology options—variable renewable energy

| Variable renewable power technology options | Input: assumed order of generation priority marked with (1) to (5) |
|---|---|
| 1. Photovoltaic—rooftop | Photovoltaic—utility scale (2) |
| 2. Photovoltaic—utility scale | Photovoltaic—rooftop (1) |
| 3. Wind—onshore | Wind—offshore (4) |
| 4. Wind—offshore | Wind—onshore (3) |
| 5. CSP (dispatchable) | CSP (5) |

order in which the supply technologies and storage functions are utilized to be adjusted to satisfy the demand. However, storage and interconnections cannot be selected as the first elements of supply (Table 3.5).

Tables 3.6, 3.7, and 3.8 provide an overview of the possible supply technologies and examples of different dispatch scenarios. Although concentrated solar power (CSP) plants with storage are dispatchable to some extent—depending on the storage size and the available solar radiation—they are part of the renewable variable group in the MATLAB model. Although the model allows the dispatch order to be changed, the 100% renewable energy analysis always follows the same dispatch logic. The model identifies excess renewable production, which is defined as any potential wind and solar photovoltaic generation greater than the actual hourly demand in MW during a specific hour. To avoid curtailment, the surplus renewable electricity must be

**Table 3.7** Technology options—dispatch generation

| Dispatch generation technology options | Input: assumed order of generation priority marked with (1) to (13) |
|---|---|
| 1. Bioenergy | Hydropower (3) |
| 2. Geothermal | Bioenergy (1) |
| 3. Hydropower | CoGen bioenergy (7) |
| 4. Ocean | Geothermal (2) |
| 5. Oil | CoGen geothermal (8) |
| 6. Gas | Ocean (4) |
| 7. CoGen bioenergy | Gas (6) |
| 8. CoGen geothermal | CoGen gas (9) |
| 9. CoGen gas | Coal (11) |
| 10. CoGen coal | CoGen coal (10) |
| 11. Coal | Brown coal (12) |
| 12. Brown coal | Oil (5) |
| 13. Nuclear | Nuclear (13) |

**Table 3.8** Technology options—storage technologies

| Storage technology option | Input: assumed priority order for storage technologies marked with (1) to (3) |
|---|---|
| 1. Battery | Hydro pump (2) |
| 2. Hydro pump | Battery (1) |
| 3. Hydrogen | Hydrogen (3) |

stored with some form of electric storage technology or exported to a different cluster or region. Within the model, the excess renewable production accumulates through the dispatch order. If storage is present, it will charge the storage within the limits of the input capacity. If no storage is present, this potential excess renewable production is reported as 'potential curtailment' (pre-storage) (Table 3.9).

*Limitations*: It is important to note that calculating the possible interconnection capacities for transmission grids between subregions does not replace technical grid simulations. Grid services, such as the inductive power supply, frequency control, and stability, should be analysed, although this is beyond the scope of the OECM analysis. The results of [R]E 24/7 provide a first rough estimate of whether increased use of storage or increased interconnection capacities or a mix of both will reduce systems costs.

#### 3.1.3.2 Regional Interconnections

Interconnection capacities are set as a function of the total generation capacity within a cluster. Interconnections between defined regions are the only ones considered, and all intra-regional interconnections or line constraints are excluded. Therefore, a region is considered a 'copper plate'—and a transmission system

**Table 3.9** Dispatch module—inputs, intermediate outputs, and outputs

| Inputs, intermediate outputs, outputs | | |
|---|---|---|
| Inputs | Maximum capacity for interconnections among regions | [MW] |
| Inputs | Initial load (cluster or region) | [MW] |
| Inputs | Technical specifications of supply technologies and storage strategies | |
| Inputs | Meteorological data | |
| Intermediate output | Dispatch order of technologies | |
| Intermediate output | Load after *variable renewable* supply | [MW] |
| Intermediate output | Load after *storage* supply | [MW] |
| Intermediate output | Load after *dispatch power plant* supply | [MW] |
| Intermediate output | Load after *interconnection* supply | [MW] |
| Output | Deficit and curtailment | [MWh] |
| Output | *Renewable penetration* | [MWh] |

where electricity can flow unconstrained from any generation site to any demand site is found in most energy modelling tools (Avrin, 2016). This simplification is required to achieve a short calculation time while maintaining high technical and time resolution. The algorithm devised for the function of the interconnectors is based on the following information for each region:

- Unmet load in the region
- Excess generation in other regions
- Interconnection capacity between the undersupplied region and each of the other regions
- Priority of the closest region(s) in exporting power to the undersupplied region

The excess generation capacity and unmet load are calculated by running the model without the interconnections to determine the excess or shortfall in generation when the load within the region is met. These excesses and shortfalls are calculated at the point in the dispatch cascade at which the interconnectors provide or consume power, for example, after the variable renewables and dispatchable generators and before the storage technologies.

The interconnection capacity between regions is defined based on a percentage of the maximum regional load. The capacity is defined in a matrix, both to and from each region to every other region. A priority order for each region to every other region is given based on proximity, so that if a region has an unmet load, it will be served sequentially with the excess generation of loads in other regions in their defined order of proximity.

For every hour and every region in each cluster (a cluster is a group of regions), the possible interconnections required for the importation or exportation of energy to balance the load are calculated. Each region is considered in turn, and the algorithm attempts to meet the unmet load with excess generation by other regions, keeping track of the residual excess loads and the interconnector capacities. Each region's internal load is met first, before its generation resources are considered for other interconnected regions.

For regions sending generation capacity to other regions, the interconnector element behaves as an increase in load, whereas for regions accepting power from neighbouring regions, the interconnector element behaves as an additional generator, from the model's perspective.

Once the total inflow and outflow of the interconnectors are calculated, the hourly values for the total supply in each region are updated, together with any residual deficit in supply or any curtailed (= forced to shut down) electricity generator that does not have priority dispatch.

Similar to the supply technologies, different storage technologies (electrical, thermal, or hydrogen) can be defined and selected, together with their technical characteristics, such as their round-trip efficiency, new or installed capacities in each year of the modelled period, lifetime, maximum depth of discharge, maximum energy out in a time step, and costs. When the total energy delivered by the supply technologies in a region does not meet the demand, energy is discharged from storage (if the storage technology has energy available), following the constraints of the storage operation (maximum energy out per time step, maximum depth of discharge, maximum depth of charge, state of charge) and the order of operation for the defined storage technologies. In the case of a demand deficit after storage, electricity from other regions will be imported. When there is surplus energy generation, the surplus will charge any storage appliances (if available), also according to the same constraints of energy storage operation and sequential order.

#### 3.1.3.3 Supply Module Architecture in MATLAB

Analogous to the demand module, inputs can be made directly into the supply module via MATLAB or a standardized Excel sheet. The supply module in MATLAB is also based on an object-oriented structure, in which classes and the objects belonging to those classes are built based on attributes and methods.

Figure 3.8 shows the UML class diagram for the supply module developed in MATLAB. Specifically, the supply module has three main classes:

1. *Supply class*: This is the main class and it is built on the supply and storage technology objects. Attributes that describe the supply class include years, region, energy supply form, fuel, and generation and storage technologies. The supply class has two main types of methods: (i) calculation supply methods and (ii) printing result methods. The calculation methods implement equations and algorithms to calculate the dispatch and fuel consumption. Table 3.10 presents a brief description of each method.
2. *Supply technology class*: This class is used to define supply technologies. Attributes include name, type, efficiency, year, region, and energy supply form and are defined as text inputs. Additional attributes are defined as numerical inputs, such as lifetime, cost, and capacity factors. The structure adopted allows

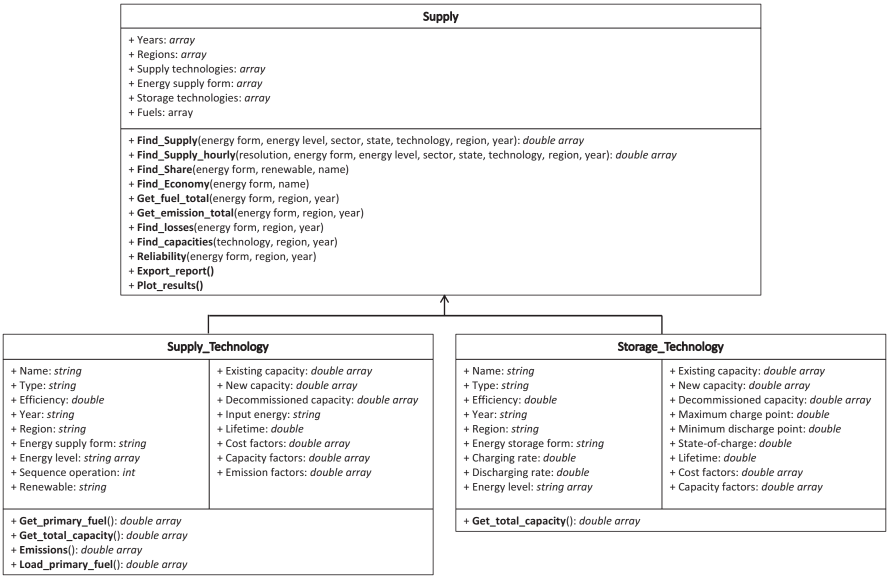


**Fig. 3.8** A UML diagram of the supply module in MATLAB, showing its classes, attributes, methods, and associations

the addition of new attributes if required. This class has methods that are used by the main supply class to calculate the primary fuel, emissions, or installed capacity of a specific technology.

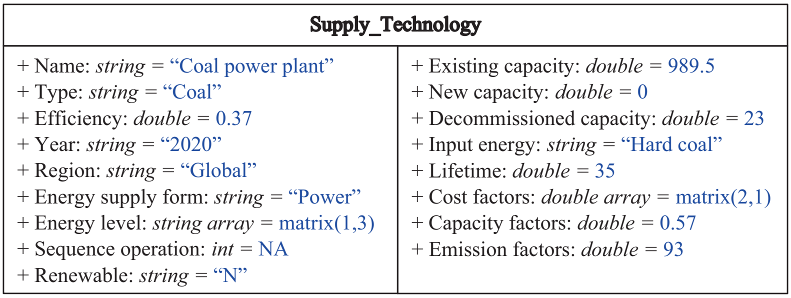
3. *Storage technology class*: This class is used to define storage technologies. The attributes include name, type, efficiency, year, region, and energy storage form and are defined as text inputs. Other numerical attributes include charging and discharging rates, capacity, cost factors, and state of charge.

Figures 3.9 and 3.10 show the high-level class definitions for supply technologies and storage objects, respectively. The text in blue indicates the defined value for each attribute. For example, the supply technology object in Fig. 3.9 has the name ‘coal power plant’, its input energy is defined as hard coal, and the object is associated with the electricity energy form. The attributes in Fig. 3.9 consider the year 2020 and a global scenario. For example, the existing capacity is defined as 989.5 GW and the decommissioned capacity is 23 GW. The lifetime of this object is 35 years.

An example of a storage object is shown in Fig. 3.10. The attributes of this object include text inputs, such as its name ‘battery lithium’ and its type ‘electrical’. This object has numerical attributes such as the efficiency (equal to 0.95 for this object) and the charging and discharging rates (fixed at 5 kW). Note that the units for each attribute are defined when the module is initialized in MATLAB.

**Table 3.10** Methods within the supply class

| Type of method | Method | Description |
|---|---|---|
| Calculation | Find_Supply() and Find_Supply_hourly() | These methods calculate the annual or hourly aggregated energy supply for the specified region and the energy form (i.e. power, heat, or hydrogen). The calculations can be made by individual or group supply technology type or storage type. These methods can also be used to calculate the emissions and primary fuel associated with each supply technology |
| Calculation | Find_Share() | This method calculates the share factor results for predefined supply scenarios; for example, the share factors of power generated from renewable energy sources and non-renewable sources. Another example is the portion of the transport sector that requires electricity or hydrogen |
| Calculation | Find_Economy() | This method calculates the costs associated with supply technologies |
| Calculation | Get_fuel_total() | This method calculates the total fuel or total primary fuel required for demand and supply |
| Calculation | Get_emission_total() | This method calculates the total emissions, considering all the demand sectors and supply technologies |
| Calculation | Find_losses() | This method calculates the losses for a specified energy form. For example, it can be used to calculate the electricity losses or heat losses arising from transport and distribution |
| Calculation | Find_capacities() | This method calculates the installed capacity for a specified technology when decommissioning or new capacity parameters have been defined |
| Calculation | Reliability() | This method calculates the total energy deficit and curtailment based on the total demand and generation, for the specified energy form |
| Printing results | Export_report() | This method exports the specified results to external Excel spreadsheets and can be used to print results on predefined report tables |
| Printing results | Plot_results() | This method can be used to plot results using the MATLAB interface |



**Supply_Technology**

+ Name: *string* = "Coal power plant"
+ Type: *string* = "Coal"
+ Efficiency: *double* = 0.37
+ Year: *string* = "2020"
+ Region: *string* = "Global"
+ Energy supply form: *string* = "Power"
+ Energy level: *string array* = matrix(1,3)
+ Sequence operation: *int* = NA
+ Renewable: *string* = "N"
+ Existing capacity: *double* = 989.5
+ New capacity: *double* = 0
+ Decommissioned capacity: *double* = 23
+ Input energy: *string* = "Hard coal"
+ Lifetime: *double* = 35
+ Cost factors: *double array* = matrix(2,1)
+ Capacity factors: *double* = 0.57
+ Emission factors: *double* = 93

**Fig. 3.9** An example of a supply technology object, showing the assigned attributes

| **Storage_Technology** | |
|---|---|
| + Name: *string* = "Battery Lithium" | + Existing capacity: *double* = 15 |
| + Type: *string* = "Electrical" | + New capacity: *double* = 0 |
| + Efficiency: *double* = 0.95 | + Decommissioned capacity: *double* = 0 |
| + Year: *string* = 2020 | + Maximum charge point: *double* = 13 |
| + Region: *string* = "Global" | + Minimum discharge point: *double* = 2 |
| + Charging rate: *double* = 5 | + State-of-charge: *double* = 0.2 |
| + Discharging rate: *double* = 5 | + Lifetime: *double* = 20 |
| + Energy level: *string array* = matrix(1,3) | + Cost factors: *double array* = matrix(1,2) |

**Fig. 3.10** An example of a storage technology object, showing the assigned attributes

The supply module architecture developed is flexible to accommodate different types of supply and storage technologies. Additional attributes or methods can be easily added to the model.

### *3.1.4 Databases and Model Calibration*

The OECM model uses several databases for energy statistics, energy intensities, technology market shares, and other market or socio-economic parameters. The calculation of the energy balance for the base year is based on the International Energy Agency (IEA) Advanced World Energy Balances (IEA, 2020, 2021).

The energy statistics for a calculated country and/or region are uploaded via an interface module. The data for each year from 2005 onwards until the last year for which data are available are used to calibrate the model. This process is based on the energy system model (EM), developed by the German Aerospace Center DLR, and is implemented in the energy simulation platform Mesap/PlaNet (Schlenzig, 1999; Seven2one, 2012). The market shares are calculated based on the IEA statistics and a technical database for energy intensities for various appliances and applications across all sectors. These data are input and the calibration processes performed with a standardized Excel tool. The calibration method is briefly outlined below using the *transport* sector.

To calibrate the model, the transport demand of the past decade is recalculated on the basis of the available energy statistics. The IEA's Advanced World Energy Balances provides the total final energy demand by transport mode—aviation, navigation, rail, or road—by country, by region, or globally. However, it provides no further specification of the energy use within each of the transport modes. Therefore, further division into passenger and freight transport is calculated using percentage shares. These proportions are determined with a literature search, together with the average energy intensity for each of the transport modes for passenger and freight vehicles.

The annual transport demand in passenger-kilometres per year (pkm/year) or tonne-kilometres per year (tkm/year) is calculated as the annual energy demand

**Table 3.11** Calibration for calculating the transport demand

| Calculation concept | Process | Until 2019 | Units | Comment |
|---|---|---|---|---|
| Transport demand | | | | |
| Aviation, navigation, rail, and road—*past to present* | | | | |
| Annual demand | Data | Database | [PJ/yr] | Data: IEA Advanced World Energy Balances |
| Passenger share | Input | Literature | [%] | Shares of the total energy demand from the literature |
| Freight share | Input | Literature | [%] | Shares of the total energy demand from the literature |
| Average energy intensity—passenger transport | Data | Literature | [MJ/pkm] | Literature review—based on current supply mix |
| Average energy intensity—freight transport | Data | Literature | [MJ/tkm] | Literature review—based on current supply mix |
| Passenger-kilometres | Calculation | = Annual demand/energy intensity | [pkm] | Checked against OECD statistics |
| Tonne-kilometres | Calculation | | [tkm] | Checked against OECD statistics |
| Annual growth/reduction—passenger-kilometres | Calculation | = Annual demand in the previous year/annual demand in the calculated year | [%/yr] | Calculated to understand the trend between 2005 and 2020 |
| Annual growth/reduction—tonne-kilometres | Calculation | | [%/yr] | |
| Population—indicator of passenger transport development | Data | Database | [Million] | Data: UN |
| GDP per capita—indicator of passenger and freight transport development | Data | Database | [$GDP/capita] | Data: World Bank |
| GDP—indicator of freight transport development | Data | Database | [$GDP] | Data: World Bank |

divided by the average energy intensity by mode. These results are then compared with the Organisation for Economic Co-operation and Development (OECD) transport statistics, which provide both parameters——pkm/year and tkm/year. Calibrating the model on the basis of historical data ensures that the basis of the scenario projections for the coming years and decades is correctly mapped and ensures that the changes are calculated most realistically (Tables 3.11 and 3.12).

**Table 3.12** Projection of transport demand based on the changing demand in kilometres

| Process | 2020–2050 | Units | Comment |
|---|---|---|---|
| Aviation, navigation, rail, road—projections | | | |
| Calculation | = (pkm in previous year) × (increase/reduction in % per year) | [pkm] | Starting point: base year 2019 |
| Calculation | = (tkm in previous year) × (increase/reduction in % per year) | [tkm] | Starting point: base year 2019 |
| Input | Input in %/year | [%/yr] | Assumption |
| Input | Input in %/year | [%/yr] | Assumption |
| Calculation | Input in %/year | [million] | Assumption based on UN projection |
| Calculation | = $GDP/population | [$GDP/capita] | |
| Calculation | INPUT in %/year | [$GDP] | Assumption based on the World Bank projection |
| Result | Time series 2020–2050: pkm per year and region | [pkm/yr] | Input for energy demand calculation |
| Result | Time series 2020–2050: tkm per year and region | [tkm/yr] | Input for energy demand calculation |

For the forward projection of the transport demand, the calculation method is reversed: the transport demand for each transport mode is calculated on the basis of the annual change, as a percentage. The calculated total annual pkm and tkm are the inputs for the energy demand calculation.

This methodology for calibration and projection is used across all sectors.

The developed MATLAB tool can access online data and databases through available *application programming interfaces* (APIs). For example, the API for the World Bank Indicators provides access to nearly 16,000 time series indicators, including population estimates and projections (World Bank, 2021). Likewise, the OECD provides access to datasets through an API. This allows a developer to easily call the API and access data using the code lines in MATLAB.

### *3.1.5 Sectors and Sub-sectors*

The OECM was designed to calculate energy pathways for geographic regions, as documented in Chap. 2. The OECM was developed further to meet the requirements of the financial industry and to design energy and emission pathways for clearly defined industry sectors (sectorial pathways). The finance industry uses different classification systems to describe sub-areas of certain branches of industry. The most important system is the Global Industry Classification Standard (GICS; MSCI, 2020). However, the GICS sub-sectors do not match the IEA statistical breakdown of the energy demands of certain industries. Table 3.13 shows examples of the finance sector calculated with the OECM model, the GICS codes, and the statistical

**Table 3.13** Examples of industry sub-sectors based on the Global Industry Classification Standard (GICS)

| Financial sector | GICS | IEA statistical categories | Sector definition |
|---|---|---|---|
| **Agriculture** | 3010 Food and staple retailing | **Farming** | Food and tobacco production, excluding the energy demand for agriculture, as defined under the IEA energy statistic *other sectors*. Additional statistics from industry partners are required because the IEA statistics only provide the accumulated energy demand for agriculture and forestry |
| | 3020 Food, beverages, and tobacco | Food production and supply | |
| **Forestry** | 1510 Materials | **Agricultural and forestry** | Energy demand for all wood and wood products, including pulp and paper and printing. Also includes all energy demands for agricultural services not included in food and tobacco production |
| | 1510 50 Paper and forest products | Paper and forest products | |
| | 1510 5010 Forest products | | |
| | 1510 5020 Paper products | | |
| **Chemicals** | 1510 Materials | **Chemical industry** | Energy demand for all chemical, petrochemical, glass, and ceramic products |
| | 1510 10 Chemicals | Chemical products | |
| | | Petrochemical products | |
| | | Glass and ceramics | |
| **Aluminium** | 1510 40 Metals and mining | **Aluminium** | Energy demand for the production of primary and secondary aluminium, as well as bauxite mining |
| | 1510 4010 Aluminium | | |
| **Textiles and leather** | 2520 Consumer durables and apparel | **Textiles and leather industry** | This sector covers the energy demand for the textiles and leather industry |
| | 2520 30 Textiles, apparel, and luxury goods | | |

information used. Although the OECM model allows all the GICS code sub-sectors to be calculated, the availability of statistics is the factor limiting the resolution of the sectorial pathways. For example, the statistical data for the textile and leather industry are stored in the IEA database, but the database does not separate the two industries further.

### *3.1.6 Cost Calculation*

The costs linked to the energy supply in each year of the modelled period include the investment costs related to 'new capacities' for technologies and storage (including replacement or decommissioning, based on the assumed technical lifetime = vintaging), operation and maintenance (O&M) costs as a percentage of the total installed capacities, and fuel costs. Other inputs for each technology and storage type include the capital cost per unit (\$/kW), O&M costs as a percentage of the capital cost, and unit fuel cost (\$/GJ).

Therefore, for each technology or storage type:

- It is assumed that the change in 'installed capacity' between each of the years modelled is linear and a linear interpolation between these is considered.
- The 'installed capacities' and 'new capacities' are interrelated (one depends upon the other) in each of the modelled and interpolated years, based on the cumulative capacities in the calculated year and the assumed technology lifetime.
- The capital costs per unit and the fuel costs in each of the modelled years are also interpolated linearly between the modelled years. Therefore, if a scenario is calculated in 5-year steps, e.g. the development from 2025 and 2030, the years 2026 to 2029 are calculated as a linear interpolation.
- Replacement capacities, if required, are also included in each year as part of the investment costs.
- The O&M costs in each of the interpolated years are calculated based on the interpolated installed capacities and the annual O&M input costs (as a percentage of the capital cost).
- Annual fuel costs for non-renewable technologies are calculated based on their output energy (running time) and interpolated fuel costs.
- The resulting 'specific costs' (\$/kWh) are also calculated from the interpolated energy supplied in each year.

The total specific costs (\$/kWh) of a scenario, as practically distributed over the interpolated years, allow the incurred costs for a scenario to be determined. *Limitations*: The economic model does not consider the change in the value of money over time. Each year of the modelled period is regarded as if it were the present year, with the multiple costs incurred. Future additions to the model could include the net present costs and the contemporary value of money.

### 3.1.7 OECM 2.0 Output and Area of Use

The added value of OECM 2.0 is its high resolution of the sector-specific parameters for both demand and supply, which are required as key performance indicators (KPIs) by the finance industry. Table 3.14 provides an overview of the main parameters and the areas of their use, with a focus on the needs of institutional investors.

Commodities and GDP are the main drivers of the energy demand for industries. The projection of, for example, the global steel demand in tonnes per year over the next decades is discussed with the industry and/or client. The OECM 2.0 can calculate either a single specific sector only or a whole set of sectors. For the development of global scenarios, various industry projections are combined to estimate both the total energy supply required and the potential energy-related emissions. Therefore, a global carbon budget can be broken down into carbon budgets of specific industries.

Energy intensities are both input data for the base year and a KPI for future projections. The effect of a targeted reduction in the energy intensity in a given year and the resulting energy demand and carbon emissions can be calculated, for example, for the steel industry.

All sector demands are supplied by the same energy supply structure in terms of electricity, process heat (for each level), and total final energy. Finally, specific emissions, such as $CO_2$ per tonne of steel or per cubic metre of wastewater treatment, are calculated and can be used to set industry targets.

All input and output OECM data are available as MATLAB-based tables or graphs or as standard Excel-based reports.

### 3.1.8 Further Research Demand

Industry-specific energy intensities and energy demands are not available for a variety of industries. In particular, the energy intensities for sub-sectors of the chemical industry are either totally unavailable or confidential. A database of energy intensities is required to develop more detailed scenarios. Although energy intensities can be estimated based on the available data, the input parameters are usually derived from various sources, which may not follow the same methodology. Energy intensities based on GDP, for example, are calculated with either nominal GDP, real GDP, or purchasing power parity GDP. Furthermore, energy intensities can be provided as final energy or primary energy. In some cases, this information is not available at all. A database of industry-specific energy demands and energy intensities, with a consistent methodology, is required to improve the accuracy of calculations in future research.

To capture the complexity of regional and global building demand projection, both in terms of data availability and high technical resolution, the high-efficiency building (HEB) model was used to develop four bottom-up demand scenarios. The HEB was developed by the Central European University (CEU) of Budapest under

**Table 3.14** Energy-related key performance indicators (KPIs) for net-zero target setting, calculated with OECM 2.0

| Sector | Parameter | Units | Base year 2019 | Projections 2025, 2030, 2035, 2040, 2045, 2050 |
|---|---|---|---|---|
| *Commodities* | | | | |
| Water utilities | Water withdrawal | [Billion $m^3$/yr] | Input | Calculated projection with annual growth rates discussed with client |
| Chemical industry | Economic development | [$GDP/yr] | Input | |
| Steel industry | Product-based market projection | [Tonnes steel/yr] | Input | |
| *Energy intensities* | | | | |
| Water utilities | Wastewater treatment | [kWh/$m^3$] | Input | Technical target (KPI) Calculated with annual progress ratio based on technical assessment |
| Chemical industry | Industry-specific energy intensity | [MJ/$GDP] | Input | |
| Steel industry | Energy intensity | [MJ/tonne steel] | Input | |
| *Energy demand* | | | | |
| Water utilities | Final energy demand | [PJ/yr] | Input | Output—industry-specific scenario(s) |
| Chemical industry | Electricity demand | [TWh/yr] | Input | |
| Steel industry | Process heat demand by temperature level | [PJ/yr] | Input | |
| | Total final energy demand | [PJ/yr] | | |
| *Energy supply* | | | | |
| Water utilities | Electricity generation by technology | [TWh/yr] | Input | Output—based on scenario developed Supply for all (sub-)sectors |
| Chemical industry | (Process) heat by technology | [PJ/yr] | Input | |
| Steel industry | Fuel supply by fuel type | [PJ/yr] | Input | |
| | Total final energy supply by fuel type | [PJ/yr] | Input | |
| *Energy-related emissions* | | | | |
| | Electricity—specific $CO_2$ emissions | [g$CO_2$/kWh] | Calculated | Output—KPI for utilities |
| | Electricity—total $CO_2$ emissions | [t $CO_2$/yr] | Calculated | Output—KPI for utilities |
| | (Process) heat—specific $CO_2$ emissions | [g$CO_2$/kWh] | Calculated | Output—KPI for industry |
| | (Process) heat—total $CO_2$ emissions | [t$CO_2$/yr] | Calculated | Output—KPI for industry |

(continued)

**Table 3.14** (continued)

| Sector | Parameter | Units | Base year 2019 | Projections 2025, 2030, 2035, 2040, 2045, 2050 |
|---|---|---|---|---|
| *Product-specific emission* | | | | |
| Water utilities | Emissions intensity | [$kgCO_2/m^3$] | Calculated | KPI—water utilities |
| | Total energy-related $CO_2$ emissions | [$tCO_2$] | Calculated | KPI—water utilities |
| Chemical industry | Emissions intensity | [$kgCO_2$/\$GDP] | Calculated | KPI—chemical industry |
| | Total energy-related $CO_2$ emissions | [$tCO_2$] | Calculated | KPI—chemical industry |
| Steel industry | Emissions intensity | [$kgCO_2$/t steel] | Calculated | KPI—steel industry |
| | Total energy-related $CO_2$ emissions | [$tCO_2$] | Calculated | KPI—steel industry |

the scientific leadership of Prof. Dr. Diana Uerge-Vorsatz. The following section documents the methodology of the HEB based on the paper by Chatterjee, S.; Kiss, B.; and Ürge-Vorsatz, D. (2021). The results are documented in Sects. 7.1 and 7.2.

## 3.2 The High-Efficiency Building (HEB) Model

Modelling the energy demand for buildings is a complex task because the building sector-related energy demand depends on several factors, such as spatial resolution, temporal resolution, building physics, and the different technologies of building construction (Prieto et al., 2019; Chatterjee & Ürge-Vorstaz, 2020). The majority of demand models do not incorporate these factors and therefore provide insights into the future energy demand scenarios of the building sector that can be far from realistic (Prieto et al., 2019; Chatterjee & Ürge-Vorstaz, 2020). Therefore, in this study, we use the HEB model to understand the future energy demand potentials for building in key regions across the globe.

The HEB model was originally developed in 2012 to calculate the energy demand and $CO_2$ emissions of the residential and tertiary building sectors until 2050 under three different scenarios (Ürge-Vorsatz & Tirado Herrero, 2012). Since then, the model has been developed and updated several times. With the latest update, the model calculates the energy demand under four scenarios until 2060 based on the most recent data for macroeconomic indicators and technological development. This model is novel in its methodology compared with earlier global energy analyses and reflects an emerging paradigm—the performance-oriented approach to the energy analysis of buildings. Unlike component-oriented methods, a systemic perspective is taken: the performance of whole systems (e.g. whole buildings) is studied, and these performance values are used as the input in the scenarios. The model calculates the overall energy performance levels of buildings, regardless of the measures applied to achieve them. It also captures the diversity of solutions required in each region by including region-specific assumptions about advanced and suboptimal technology mixes. The elaborated model uses a bottom-up approach, because it includes rather detailed technological information for one sector of the economy. However, it also exploits certain macroeconomic (GDP) and socio-demographic data (population, urbanization rate, floor area per capita, etc.). The key output of the HEB model is floor area projections for different types of residential and tertiary buildings in different regions and their member states, the total energy consumption of residential and tertiary buildings, the energy consumption for heating and cooling, the energy consumption for hot water energy, the total $CO_2$ emissions, the $CO_2$ emissions for heating and cooling, and the $CO_2$ emissions for hot water energy.

### *3.2.1 The High-Efficiency Building Model Methodology*

The HEB model conducts a scenario analysis for the entire building sector, in which the building sector is distinguished by location (rural, urban, and slum), building type (single-family, multifamily, commercial, and public buildings, with subcategories), and building vintage (existing, new, advanced new, retrofitted, and advanced retrofitted). This detailed classification of buildings is undertaken for 11 regions (Ürge-Vorsatz & Tirado Herrero, 2012), extended with country-specific results for the EU-27 countries, China, India, and the USA. Furthermore, within each region, different climate zones are considered to capture the differences in building energy uses and renewable energy generation caused by variations in climate. The climate zones are calculated based on four key climatic factors—heating degree days (HDD), cooling degree days (CDD), relative humidity (RH) in the warmest month, and average temperature in the warmest month (T). These parameters are processed using the GIS5 tool—spatial analysis—and performed with the ArcGIS software. The detailed classification categories are summarized in Table 3.2.

The purpose of the detailed classification of building categories and scenario assessments is to explore the consequences of certain policy directions or decisions that inform policy-making (Table 3.15).

The key input data used in the HEB are region-specific forecasts of GDP, population, rate of urbanization, and the proportion of the population living in urban slums. The time resolution of the model is yearly, so that socio-economic input data can be easily obtained from various credible sources, such as the databases of the World Bank, United Nations Development Programme (UNDP), EUROSTAT, and the OECD. Besides these socio-economic parameters, many others are included, and in the case of data absence, assumptions are made in the HEB model to calculate the final energy demand. Figure 3.11 shows the main workflow of the HEB model.

The HEB model includes several calculation steps, from considering the input data to obtaining the final output. Each of these calculation steps is discussed in the sections below.

**Table 3.15** Building classification scheme of HEB

| Classification scope | Categories | Subscript notation |
|---|---|---|
| Regions | 11 key geographic regions +30 focus countries | *r* |
| Climate zones | 17 different climate zones | *c* |
| Urbanization | Urban/rural areas | *u* |
| Building category | Residential/commercial and public/slums | *b* |
| Building type | Single-family houses (SF)/multifamily houses (MF) (residential sector)<br>Educational/hotel and restaurant/hospital/retail/office/others (commercial and public sector) | *t* |
| Building vintage | Existing/new/advanced new/retrofitted/advanced retrofitted | *v* |

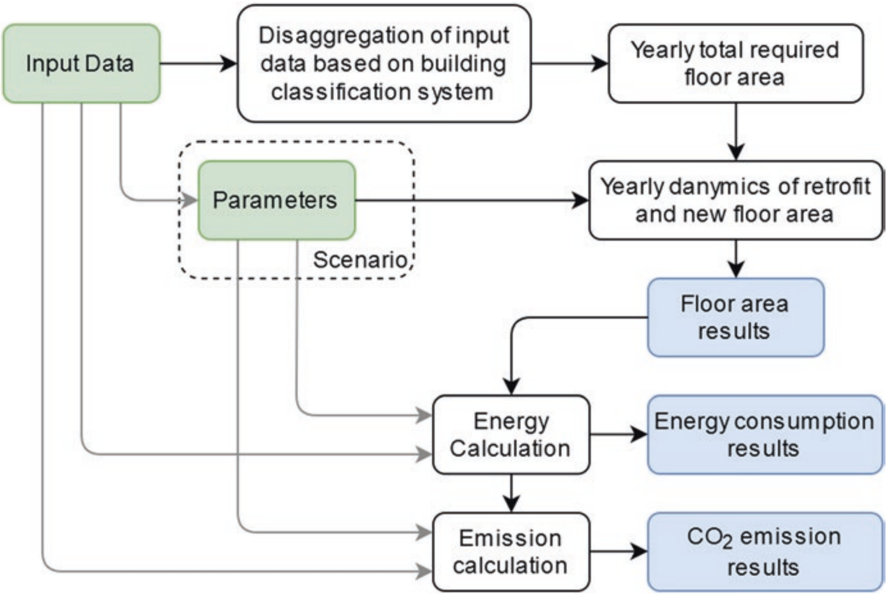


**Fig. 3.11** The main workflow of the HEB. Input data and parameters can be modified by the user (green). Main outputs are the floor areas of different building vintage types and the energy consumption and $CO_2$ emissions of the stock (blue)

### 3.2.2 Disaggregation

In the first step of the calculation, after all the socio-economic input data are obtained, the input is disaggregated into the detailed building classification scheme, and the total floor area required to satisfy the year-specific population and GDP needs (the year is denoted with *Y* in subscript) is determined. The core concept for calculating the floor area differs for residential and commercial buildings:

- For residential buildings, the total occupied floor area correlates with the population, and thus, population forecasts are used to determine the floor areas of buildings in each region.
- For commercial and public buildings, the floor area correlates more strongly with GDP, so GDP forecasts are used as a proxy to determine the total floor space areas of commercial and public buildings.

The region-specific population data—as the input for the calculation—is further disaggregated into urban and rural populations based on the urbanization rate and into the different climate zones based on GIS data:

$$P_{\mathrm{r,c,u,Y}} = P_{\mathrm{r,Y}} \times U_{\mathrm{r,Y}} \times \mathrm{Sc}_{\mathrm{r,c}} \quad \text{if u = urban} \tag{3.1}$$

and

$$P_{\mathrm{r,c,u,Y}} = P_{\mathrm{r,Y}} \times \left(1 - U_{\mathrm{r,Y}}\right) \times \mathrm{Sc}_{\mathrm{r,c}} \quad \text{if r = rural} \tag{3.2}$$

where

$P_{\mathrm{r,c,u,Y}}$ [capita] is the total urban/rural population of region r and climate zone c in year Y
$P_{\mathrm{r,Y}}$ [capita] is the total population of region r in year Y
$U_{\mathrm{r,Y}}$ [−] is the urbanization rate of region r in year Y
$\mathrm{Sc}_{\mathrm{r,c}}$ [%] is the share of the population within region r living in climate zone c

The urban population is then further disaggregated into the population living in slums (in regions where a significant number of people do not have access to standard living conditions) and the population living in conventional residential buildings. The latter group is split into the populations living in single-family and multifamily houses based on region-specific fixed values:

$$P_{\mathrm{r,c,u,b,Y}} = P_{\mathrm{r,c,u,Y}} \times \mathrm{Ss}_{\mathrm{r,Y}} \quad \text{where u = urban and b = slum} \tag{3.3}$$

and

$$P_{\mathrm{r,c,u,b,Y}} = P_{\mathrm{r,c,u,Y}} \times \left(1 - \mathrm{Ss}_{\mathrm{r,Y}}\right) \quad \text{where u = urban and b = residential} \tag{3.4}$$

then

$$P_{\mathrm{r,c,u,b,t,Y}} = P_{\mathrm{r,c,u,b,Y}} \times \mathrm{Ssf}_{\mathrm{r}} \quad \text{where u = urban, b = residential and t = SF} \tag{3.5}$$

and

$$P_{\mathrm{r,c,u,b,t,Y}} = P_{\mathrm{r,c,u,Y}} \times \left(1 - \mathrm{Ssf}_{\mathrm{r}}\right) \quad \text{where u = urban, b = residential and t = MF} \tag{3.6}$$

where

$P_{\mathrm{r,\,c,\,u,\,b,\,Y}}$ [capita] is the total urban/rural population of region r, climate zone c, and building category b in year Y.
$P_{\mathrm{r,\,c,\,u,\,Y}}$ [capita] is the total urban/rural population of region r and climate zone c in year Y.
$P_{\mathrm{r,\,c,\,u,\,b,\,t,\,Y}}$ [capita] is the total urban/rural population of region r, climate zone c, building category b, and building type t in year Y.
$\mathrm{Ss}_{\mathrm{r,\,Y}}$ [%] is the share of the urban population living in slums in region r and year Y.
$\mathrm{Ssf}_{\mathrm{r}}$ [%] is the share of the urban population living in single-family houses in region r.

The population living in rural areas is assumed to live in single-family houses.

The disaggregation of GDP follows the same pattern, except that the share of GDP that can be associated with rural commercial or public buildings is fixed within the modelling period:

$$\mathrm{GDP}_{\mathrm{r,c,u,Y}} = \mathrm{GDP}_{\mathrm{r,Y}} \times \left(1 - U_{\mathrm{r,Y}}\right) \times \mathrm{Sc}_{\mathrm{r,c}} \quad \text{if u = urban} \tag{3.7}$$

and

$$\mathrm{GDP}_{\mathrm{r,c,u,Y}} = \mathrm{GDP}_{\mathrm{r,Y}} \times U_{\mathrm{r,Y}} \times \mathrm{Sc}_{\mathrm{r,c}} \quad \text{if u = rural} \tag{3.8}$$

where

$\mathrm{GDP}_{\mathrm{r,\,c,\,u,\,Y}}$ [USD] is the total GDP that can be associated with urban/rural commercial or public buildings in region r and climate zone c in year Y.
$\mathrm{GDP}_{\mathrm{r,\,Y}}$ [USD] is the total GDP of region r in year Y.
$U_{\mathrm{r,\,Y}}$ [−] is the urbanization rate of region r in year Y.
$\mathrm{Sc}_{\mathrm{r,\,c}}$ [%] is the share of climate zone c within region r.

The share of different commercial building types is also determined with fixed ratios based on data from the literature:

$$\mathrm{GDP}_{\mathrm{r,c,u,t,Y}} = \mathrm{GDP}_{\mathrm{r,c,u,Y}} \times \mathrm{Scp}_{\mathrm{t}} \tag{3.9}$$

where

$\mathrm{GDP}_{\mathrm{r,\,c,\,u,\,t,\,Y}}$ [USD] is the total GDP that can be associated with urban/rural commercial or public buildings of type t in region r and climate zone c in year Y.
$\mathrm{GDP}_{\mathrm{r,\,c,\,u,\,Y}}$ [USD] is the total GDP that can be associated with urban/rural commercial or public buildings in region r and climate zone c in year Y.
$\mathrm{Scp}_{\mathrm{t}}$ [%] is the share of commercial and public buildings of type t in the commercial and public building stock.

### 3.2.3 Determining the Total Floor Area

Different equations are used for the calculation of the floor area of residential buildings and non-residential buildings. The floor area of residential buildings can be calculated with the following equation, using specific floor area values (the floor area that is occupied by one person):

$$\mathrm{TFA}_{\mathrm{r,c,u,b,t,Y}} = \mathrm{P}_{\mathrm{r,c,u,b,t,Y}} \times \mathrm{SFAc}_{\mathrm{r,u,b,t,Y}} \quad \text{where b} = \text{residential / slum} \tag{3.10}$$

where

$\mathrm{TFA}_{\mathrm{r,\,c,\,u,\,b,\,t,\,Y}}$ [m$^2$] is the total urban/rural floor area of building category b and building type t in region r and climate zone c in year Y.

$\mathrm{P}_{\mathrm{r,\,c,\,u,\,b,\,t,\,Y}}$ [capita] is the total urban/rural population of region r, climate zone c, building category b, and building type t in year Y.

$\mathrm{SFAc}_{\mathrm{r,\,u,\,b,\,t,\,Y}}$ [m$^2$/capita] is the specific floor area of building category b and building type t in region r in year Y.

Similarly, the floor area of commercial and public buildings is calculated using specific floor area values (the floor area that is required to produce one unit of GDP):

$$\mathrm{TFA}_{\mathrm{r,c,u,b,t,Y}} = \mathrm{GDP}_{\mathrm{r,c,u,t,Y}} \times \mathrm{SFAg}_{\mathrm{r,b,Y}} \quad \text{if b} = \text{C \& P} \tag{3.11}$$

where

$\mathrm{TFA}_{\mathrm{r,\,c,\,u,\,b,\,t,\,Y}}$ [m$^2$] is the total urban/rural floor area of commercial or public buildings of building type t in region r and climate zone c in year Y.

$\mathrm{GDP}_{\mathrm{r,\,c,\,u,\,t,\,Y}}$ [USD] is the total GDP that can be associated with urban/rural commercial or public buildings of type t in region r and climate zone c in year Y.

$\mathrm{SFAg}_{\mathrm{r,\,b,\,Y}}$ [m$^2$/USD] is the specific floor area of commercial or public buildings in region r in year Y.

Specific floor area values are determined from statistical data for each region. To take socio-economic development into account, the floor area per capita and the floor area per GDP are modelled as values that change yearly, reaching the average for OECD countries by the end of the modelling period in developing regions.

### 3.2.4 Yearly Dynamics of Floor Area Changes

The yearly dynamics of this floor area model transition the existing building stock into the future state determined by the scenarios. This includes the retrofitting or demolition of existing buildings, as well as the introduction of new buildings to the stock. In some cases, the floor area is left abandoned, which might result from a reduction in the population (e.g. in developed regions) or an increased rate of urbanization due to which buildings located in rural areas are abandoned after a certain time. It is important to capture this phenomenon, because abandoned buildings do not contribute to energy consumption or the emissions of the building stock. This yearly dynamic of the vintage types of buildings is presented in Fig. 3.12.

The demolished floor area is calculated with region-specific demolition rates. After the demolished floor area is subtracted from the existing total, the remaining existing floor area is classified into different building vintages. Similarly, the retrofitted floor area is calculated by applying the yearly changing region-specific retrofitting rate to the total existing building stock. The retrofitted floor area is further classified into two types: advanced retrofitted floor area and normal retrofitted floor area. For each of the regions, the shares of retrofitted and advanced retrofitted floor area differ, and the shares of advance retrofitted, advance new, and retrofitted floor areas also vary under different scenarios. The floor area from new constructions is classified into two building vintages: new and advanced new. Like the retrofitted floor area, the share of advanced new floor area also varies under different scenarios.

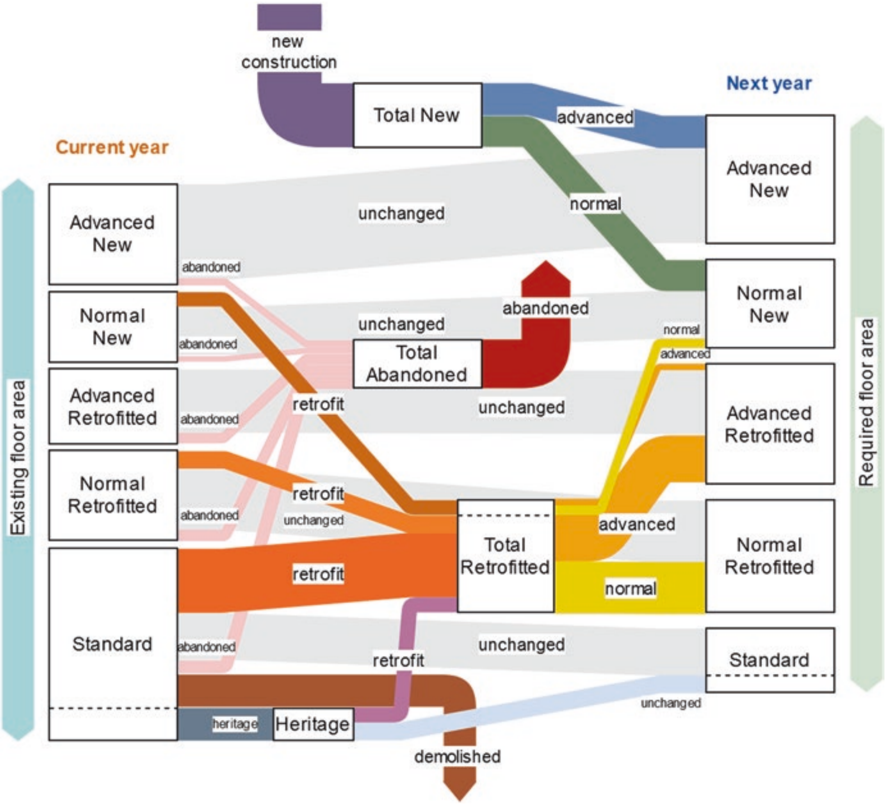


Fig. 3.12 Yearly floor area dynamics in the HEB model

### 3.2.5 Calculating the Energy Consumption of Buildings

The energy consumption for heating and cooling depends on the floor area. Therefore, in the HEB model, energy consumption is calculated after the year-specific floor area is calculated. The key input required to calculate the energy consumption for heating and cooling is the average consumption data for heating and cooling, which are usually obtained from data reported in the literature, for each of the regions, climate zones, and building types, because different building vintages have different consumption requirements. Therefore, different vintage types are modelled by assuming different energy intensities (denoted with subscript v). The values also depend on the scenario (denoted with subscript s). Energy intensity is multiplied by the corresponding floor area to determine the energy consumption for heating and cooling the stock:

$$\mathrm{HCE}_{\mathrm{r,c,u,b,t,Y,v,s}} = \mathrm{TFA}_{\mathrm{r,c,u,b,t,Y}} \cdot \mathrm{EUhc}_{\mathrm{r,c,u,b,t,v,s}} \tag{3.12}$$

where

$\mathrm{HCE}_{\mathrm{r,\,c,\,u,\,b,\,t,\,Y,\,v,\,s}}$ [kWh/year] is the total energy demand for heating and cooling of buildings with vintage type v in scenario s, building type t in region r, and climate zone c in year Y.

$\mathrm{TFA}_{\mathrm{r,\,c,\,u,\,b,\,t,\,Y}}$ [$\mathrm{m}^2$] is the total urban/rural floor area of building category b and building type t in region r and climate zone c in year Y.

$\mathrm{EIhc}_{\mathrm{r,\,c,\,u,\,b,\,t,\,v,\,s}}$ [$\mathrm{kWh/m}^2$/year] is the heating and cooling energy intensity of buildings of vintage type v in scenario s and building type t in region r and climate zone c.

After the detailed energy consumption is calculated, the data can be summed to arrive at the region-specific, yearly aggregated results for a given scenario:

$$\text{Total Energy}_{\mathrm{r,Y,s}} = \sum_{\mathrm{c}}\sum_{\mathrm{u}}\sum_{\mathrm{b}}\sum_{\mathrm{t}}\sum_{\mathrm{v}} \text{Total Energy}_{\mathrm{r,c,u,b,t,Y,v,s}} \tag{3.13}$$

### 3.2.6 Implementation

The most recent version of the HEB model was developed in the Python programming language, using the PyData ecosystem to handle large datasets. This ecosystem ensures quite large flexibility among the modelling parameters, and the diversity of input data and its granularity can be properly handled. This model is not an open-access model, but the Central European University has received funding from the European Union's Horizon 2020 research and innovation programme (under grant agreement no. 837089) in the Sentinel[1] project, to develop HEB further. In this project, the HEB model will be made an open-source model that users can use without cost.

## References

Avrin, A-P. (2016). *Implications of electricity transmission for integrating variable renewable energies – Young scientists summer program – IIASA*. International Institute for Applied Systems Analysis. https://iiasa.ac.at/web/scientificUpdate/2015/cb/Avrin-Anne-Perrine3.html. Accessed 20 Dec 2021.

Chatterjee, S, Ürge-Vorstaz, D. (2020) *D3.1: Observed trends and modelling paradigms. Topic: LC-SC3-CC-2-2018 of the horizon 2020 work program: Modelling in support to the transition to a low-carbon energy system in Europe. Building a low-carbon, climate resilient future: Secure, clean and efficient energy*. SENTINEL. https://sentinel.energy/wp-content/uploads/2021/02/D-3.1-837089-EC.pdf

Chatterjee, S., Kiss, B., & Ürge-Vorsatz, D. (2021). *How far can building energy efficiency bring us towards climate neutrality?* Central European University.

Farjana, S. H., Huda, N., Mahmud, M. A. P., & Saidur, R. (2018). Solar process heat in industrial systems – A global review. *Renewable and Sustainable Energy Reviews, 82*, 2270–2286. https://doi.org/10.1016/j.rser.2017.08.065

IEA. (2020). *World energy balances 2020*. IEA. https://www.iea.org/data-and-statistics?country=WORLD&fuel=Energytransitionindicators&indicator=TFCShareBySector. Accessed 6 Apr 2021.

IEA. (2021). *IEA world energy statistics and balances*. IEA. https://doi.org/10.1787/enestats-data-en. https://www.oecd-ilibrary.org/energy/data/iea-world-energy-statistics-and-balances_enestats-data-en. Accessed 10 Oct 2021.

IPCC. (2021). Summary for policymakers. In V. P. Masson Delmotte, A. Zhai, S. L. Pirani, et al. (Eds.), *Climate change 2021: The physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

MSCI. (2020). *Global industry classification standard (GICS®) methodology*. Guid. Princ. Methodol. GICS. https://www.msci.com/documents/1296102/11185224/GICS+Methodology+2020.pdf/9caadd09-790d-3d60-455b-2a1ed5d1e48c?t=15784059356. Accessed 31 Jan 2021.

Nikas, A., Doukas, H., & Papandreou, A. (2019). A detailed overview and consistent classification of climate-economy models. In *Understanding risks and uncertainties in energy and climate policy: Multidisciplinary methods and tools for a low carbon society* (pp. 1–54). https://doi.org/10.1007/978-3-030-03152-7_1

NZAOA. (2021). *UN-convened Net-Zero Asset Owner Alliance – Institutional investors transitioning their portfolios to net zero GHG emissions by 2050*. https://www.unepfi.org/net-zero-alliance/. Accessed 3 Oct 2021.

Pagenkopf, J., van den Adel, B., Deniz, Ö., & Schmid, S. (2019). Transport transition concepts. In S. Teske (Ed.), *Achieving the Paris climate agreement goals: Global and regional 100% renewable energy scenarios with non-energy GHG pathways for +1.5 C and +2 C* (pp. 131–159). Springer International Publishing. https://doi.org/10.1007/978-3-030-05843-2_6

---

[1] https://sentinel.energy/

Prieto, M. P., de Uribarri, P. M. Á., & Tardioli, G. (2019). Applying modeling and optimization tools to existing city quarters. *Urban Energy Syst Low-Carbon Cities*, 333–414. https://doi.org/10.1016/B978-0-12-811553-4.00010-X

Schlenzig, C. (1999). Energy planning and environmental management with the information and decision support system MESAP. *International Journal of Global Energy Issues, 12*(1/2/3/4/5/6), 81–91. https://ideas.repec.org/a/ids/ijgeni/v12y1999i1-2-3-4-5-6p81-91.html

Seidenberger, T., Thrän, D., Offermann, R., et al. (2010). Global biomass potentials. Investigation and assessment of data, remote sensing in biomass potential research, and country-specific energy crop potentials. *Epic [r]evolution – A Sustain World Energy Outlook, 3*, 166–168.

Seven2one. (2012). Mesap/PlaNet Software Framework. *Mesap4, Release, 4*(14), 1.9.

Simon, S., Naegler, T., & Gils, H. C. (2018). Transformation towards a renewable energy system in Brazil and Mexico—Technological and structural options for Latin America. *Energies, 11*, 907. https://doi.org/10.3390/EN11040907

Süsser, D., Gaschnig, H., Ceglarz, A., Stavrakas, V., Flamos, A., & Lilliestam, J. (2022). Better suited or just more complex? On the fit between user needs and modeller-driven improvements of energy system models. *Energy, 239*(Part B), 121909., ISSN 0360-5442. https://doi.org/10.1016/j.energy.2021.121909

Teske, S. (2015). Thesis, bridging the gap between energy and grid models. In *Developing an integrated infrastructural planning model for 100% renewable energy systems in order to optimize the interaction of flexible power generation, smart grids and storage technologies, chapter 2*. University Flensburg.

Teske, S., Pregger, T., Simon, S., et al. (2019a). Methodology. In S. Teske (Ed.), *Achieving the Paris climate agreement goals: Global and regional 100% renewable energy scenarios with non-energy GHG pathways for +1.5°C and +2°C* (pp. 25–78). Springer International Publishing.

Teske, S., Pregger, T., Naegler, T., Simon, S., Pagenkopf, J., van den Adel, B., & Deniz, Ö. (2019b). Energy scenario results. In *Achieving the Paris climate agreement goals: Global and regional 100% renewable energy scenarios with non-energy GHG pathways for +1.5 C and +2 C* (pp. 175–401). https://doi.org/10.1007/978-3-030-05843-2_8

UNFCCC. (2015). *United Nations Climate Change, website, Report of the Conference of the Parties on its twenty-first session*, December 2015, at https://unfccc.int/sites/default/files/resource/docs/2015/cop21/eng/10a01.pdf

Ürge-Vorsatz, D., & Tirado Herrero, S. (2012). Building synergies between climate change mitigation and energy poverty alleviation. *Energy Policy, 49*, 83–90. https://doi.org/10.1016/J.ENPOL.2011.11.093

World Bank. (2021). *World Bank Indicators API*. Available at https://datahelpdesk.worldbank.org/knowledgebase/articles/889386-developer-information-overview; and OECD API, available at https://data.oecd.org/api/sdmx-json-documentation/