United States Environmental Protection Agency

EPA/600/B-25/077 | May 2025 | www.epa.gov/research

Generative City-based Optimization Model for Energy Technologies:

COMET Documentation and User Guide

Office of Research and Development

Center for Environmental Measurement and Modeling

May 2025

Generative City-based Optimization Model for Energy Technologies:

COMET Documentation and User Guide

By

Kathleen Vaillancourt, Marie Pied, Romain Chaffanjon, Florian Pedroli

ESMIA Consultants (sub-contractor to Abt Global)

David Cooley, Nikolaas Dietsch

Abt Global

For

Ozge Kaplan

U.S. EPA Office of Research and Development
Center for Environmental Measurement and Modeling

**This report was produced with the assistance of UNC Institute for The Environment under Contract 68HERD21A0002 Task Order #68HERH22F0164. While UNC is the prime contractor for this TO, most of the technical work performed by Abt Global and their sub-contractor, ESMIA consultants.*

Disclaimer

The U.S. Environmental Protection Agency, through its Office of Research and Development, funded and conducted the research described herein under an approved Quality Assurance Project Plan (J-AESMD-0033640-QP-1-1). The Contractor's role did not include establishing Agency policy. It has been subjected to the Agency's peer and administrative review and has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use. While UNC is the prime contractor for this TO, most of the technical work performed by Abt Global and their sub-contractor, ESMIA consultants.

Suggested Citation

Kaplan, O., K. Vaillancourt, M. Pied, R. Chaffanjon, F. Pedroli, D. Cooley, and N. Dietsch. Generative City-based Optimization Model for Energy Technologies: COMET Documentation and user guide. U.S. Environmental Protection Agency, Washington, DC, USA, EPA/600/B-25/077. DOI: 10.23719/1532263

About ESMIA

ESMIA is a leader in integrated E3 (energy-economy-environment) systems modeling and analysis for strategic decision-making on complex issues such as energy transition and energy security. ESMIA offers a scientific approach, guided by sophisticated models and high-quality data, that supports the achievement of energy and environmental goals without compromising economic growth. ESMIA provides a comprehensive suite of modeling services, not limited to: i) developing state-of-the-art turnkey E3 models for organizations worldwide, ii) assisting clients in building their custom models, and iii) delivering studies on energy scenarios and energy policies leveraging its proprietary models.

ESMIA serves organizations such as federal and provincial/state departments, city administrations, consulting firms, private companies, and industry associations. To maintain its leadership, ESMIA has forged strong partnerships with leading research institutes and laboratories in North America to advance modeling methodologies. ESMIA works closely with the International Energy Agency (IEA) through its participation in the Energy Systems Technology Analysis Program (ETSAP) Technology Collaboration Program (TCP).

Project team

	POSITION AT ESMIA	ROLE
Kathleen Vaillancourt PhD. MBA	President	Project manager / Technical lead
Marie Pied M.Sc.	Lead energy system modeler	Lead modeler / Documentation
Romain Chaffanjon M.Sc.	Energy system modeler	Lead modeler
Florian Pedroli M.Sc.	Energy system modeler	Modeler / Dashboard / Documentation

Table of contents

List of figures

List of tables

List of acronyms

Acronym	Definition
CBECS	Commercial Buildings Energy Consumption Survey
COMET	City-based Optimization Model for Energy Technologies
eGRID	Emissions & Generation Resource Integrated Database
EIE	Google Environment Insights Explorer
ETSAP	Energy Technology Systems Analysis Program
FAF	Freight Analysis Framework
GC	Generative COMET
GHG	Greenhouse gas
IEA	International Energy Agency
MACC	Marginal abatement cost curve
MOVES	Motor Vehicles Emissions Simulator
NEMS	National Energy Modeling System
RECS	Residential Energy Consumption Survey
SEDS	State Energy Data System
TIMES	The Integrated MARKAL EFOM System
US EPA	United States Environmental Protection Agency

SECTION 1

1. Model documentation

1.1. Introduction

The United States Environmental Protection Agency (U.S. EPA) conducts extensive research to enhance the scientific foundation for national environmental decision-making. In this context, the EPA has developed COMET (City-based Optimization Model for Energy Technologies) to capture the whole energy system at the city level over a user-defined analyses timeline, from the extraction of primary resources to conversion into useful energy to meet end-use service demands. COMET accounts for the investment and operation costs, as well as greenhouse gas (GHG) emissions and other air pollutants, of alternative technology pathways meeting long-term energy demands in the buildings (BLD), transportation (TRA), and waste sectors (WAS). In this way, COMET enables users to explore, compare, and optimize energy technology solutions over the coming decades, especially for medium- and large-sized cities working to achieve energy optimization objectives and emissions reduction targets. Model results reveal how the energy system can be balanced under different scenarios and assumptions, as well as how system costs and emissions change with respect to those scenarios. With this information, city officials and their stakeholders working to pursue energy planning within their jurisdictions are positioned to make more informed policy and program decisions.

U.S. EPA's efforts to develop and deploy the COMET model have proceeded in two phases. In Phase I, the project team updated the COMET model and developed a version tailored to New York City. This version enabled a comprehensive analysis of energy planning strategies for New York City, and resulted in the publication of the first New York City Climate Budget (in April 2024). The project team forecasted three baseline scenarios through 2050, and modeled city actions that are: i) expected to achieve significant GHG reductions, ii) already committed to through local laws or executive orders, and iii) included in the budget or capital plan.

In Phase II, the team developed an open-source version of COMET designed for use by local planning, energy, or environmental agencies in any city. The Generative COMET model incorporates innovative features that streamline the calibration process to align with official energy or GHG inventories, based on the level of detail available in city-level data. It provides cities with a versatile framework for analyzing energy and GHG emission scenarios, with different levels of data granularity and city-specific conditions. Its modular and adaptable structure enables cities of all sizes to explore customized strategies for meeting energy demands and achieving GHG reduction targets effectively.

This *Documentation and User Guide* provides comprehensive technical guidance for users of the Generative COMET model working to directly apply, better understand, or interpret results. These users are likely to include city planners, environmental analysts, and policymakers who are familiar with energy system modeling and are seeking a robust tool for evaluating urban climate and energy strategies. The contents of this document include detailed descriptions of the model structure and functionalities, along with a high-level overview of the model's architecture and data sources, calibration processes, and potential applications to support decision-making (Section 1). The User Guide (Section 2) includes operational details, including model installation and calibration, as well as model solving and reporting of results. It concludes with a description of model limitations and perspectives for future development (Section 3).

1.2. Background

1.2.1. The COMET model

Broadly speaking, the objective of energy, economy, and environment (E3) models is to analyze the interplay between energy systems, economic growth, and environmental impacts. E3 analytic tools include top-down macroeconomic models that focus on economic interactions and policies at the regional or national level, and bottom-up models that provide detailed analyses of specific technologies and their contribution to achieving relevant policy goals. Bottom-up energy system models can be further classified into two broad categories: 1) energy system simulation models that explore possible future states based on predefined conditions, and 2) optimization models that identify the most cost-effective pathways to achieve specified objectives, such as minimizing emissions or costs.

The core COMET model stands out as a bottom-up energy system optimization model uniquely tailored for medium- (population 200,000 to 500,000) and large-sized (population more than 500,000) cities. It employs a techno-economic modeling approach to represent the energy sector of a region or country, incorporating detailed technical parameters (such as emission coefficients) and economic characteristics of various energy technologies. In this way, COMET provides a detailed representation of the energy demand and supply sectors, including extraction, transformation, distribution, end uses, and trade of various energy forms and materials. It simulates market competition of energy carriers and energy technologies to satisfy useful energy demands specified exogenously over a given time horizon. This comprehensive modeling framework enables city officials to evaluate the full range of mitigation options and design optimal strategies that balance a city's energy requirements, emissions goals, and economic priorities.

This model structure enables an in-depth, techno-economic analysis of specific energy technologies within sectors like buildings, transportation, and waste management, allowing for a detailed understanding of how each technology contributes to overall energy demand and emissions reductions. By relying on socio-economic data to define energy demand—such as the number of households, industrial production levels, and travel distances—COMET can project energy needs that align closely with actual urban conditions.

As an optimization model, COMET's primary goal is to determine the least-cost pathways for meeting energy demands while satisfying environmental constraints, such as emissions reduction targets. COMET specifies an optimal mix of technologies, identifying the most cost-effective options to balance energy supply and demand in the long term. This structure supports city officials and other decision-makers by clarifying economic and environmental trade-offs, informing decision-making on technology pathways, and supporting cost-effective urban decarbonization strategies over the 2050 time horizon.

1.2.2. The TIMES approach

The Generative COMET model applies the TIMES[1] modeling framework that is developed and supported by the Energy Technology Systems Analysis Program (ETSAP) of the International Energy Agency (IEA). TIMES is currently used in approximately 70 countries to analyze complex energy-related issues, including technology transitions, energy security, and climate policy strategies. TIMES can assist in the design of least-cost pathways for sustainable energy systems and is ideally suited for the preparation of Low-Emissions Development Strategies (LEDS) and Intended Nationally Determined Contributions (INDC) and Nationally Determined Contributions (NDC) roadmaps. It is usually applied to the analysis of the entire energy sector but may also be applied to study single sectors such as the electricity and district heat sector.

TIMES consists of generic variables and equations constructed from the specification of sets and parameter values depicting an energy system for each distinct region in a model. To construct a TIMES model, a preprocessor first translates all data defined by the modeler into special internal data structures representing the coefficients of the TIMES matrix applied to each variable for each equation in which the variable may appear. This step is called Matrix Generation. Once the model is solved (optimized) a Report Writer assembles the results of the run for analysis by the modeler. The matrix generation, report writer, and control files are written in GAMS (the General Algebraic Modelling System). GAMS is a powerful high-level language specifically designed to facilitate the process of building large-scale optimization models. GAMS accomplishes this by relying heavily on the concepts of sets, compound indexed parameters, dynamic looping and conditional

[1] TIMES = The Integrated MARKAL EFOM System

controls, variables and equations. Thus, there is a very strong synergy between the philosophy of GAMS and the overall concept of the Reference Energy System (RES) specification embodied in TIMES, making GAMS very well suited to the TIMES paradigm. Furthermore, by nature of its underlying design philosophy, the GAMS code is very similar to the mathematical description of the equations. Thus, the approach taken to implement a TIMES model is to "convert" the input data by means of a (rather complex) preprocessor while taking care of the necessary exceptions to properly construct the matrix coefficients for the Linear Programming (LP) model. In addition, GAMS platform integrates seamlessly with a wide range of commercially available optimizers such as CPLEX and/or XPRESS. To build, run, and analyze a TIMES model, several software tools have been developed in the past or are currently under development, so that the modeler does not need to provide the input information needed to build a TIMES model directly in GAMS. These tools are the model interfaces VEDA2.0[2]. EPA/ORD currently holds licenses to utilize VEDA2.0 to build TIMES models. The TIMES model generator has extensive documentation and demo models to build instances of TIMES models[3].

In TIMES, a complete scenario consists of four types of inputs: energy service demand curves, primary resource supply curves, a policy setting, and the descriptions of a complete set of technologies. The basis of a TIMES model is a network diagram called a Reference Energy System (RES), which depicts an energy system from resource supply to end-use demand (Figure 1) The RES constructs an energy system up from a list of technology types, energy carriers, and user demands. The four technology types represented are resource, process, conversion, and demand technologies as defined in detail below:

1) Resource technologies represent the extraction cost and availability of resources such as coal, oil, and natural gas.
2) Conversion technologies represent the conversion of fuel inputs into electricity.
3) Process technologies represent other means of converting resources into end-use fuels including refineries and coal-to-liquid processes.
4) Demand technologies represent the technologies that meet specific user demands, such as vehicles, air conditioners, and water heaters.

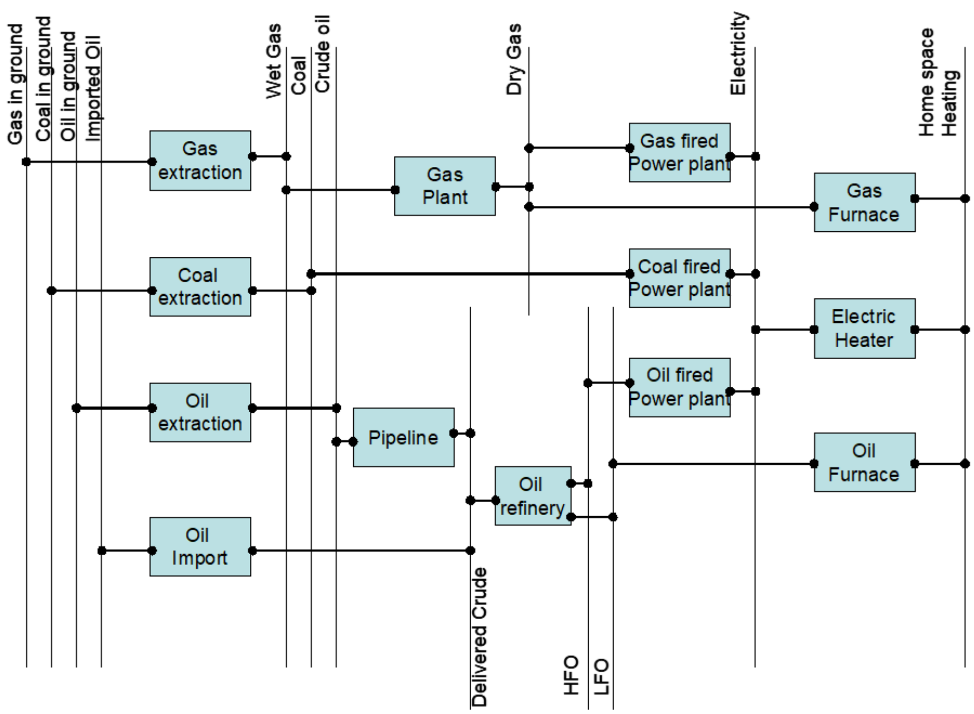
These technologies feed into a final stage consisting of end-use demands for useful energy services. End-use demands include items such as residential lighting, commercial air conditioning, and automobile passenger miles traveled. Estimates of end-use energy service demands (e.g.,

[2] https://iea-etsap.org/docs/Documentation_for_the_TIMES_Model-Part-IV.pdf
https://veda-documentation.readthedocs.io/en/latest/
[3] https://iea-etsap.org/index.php/documentation

vehicle miles traveled; residential lighting, steam heat requirements in the paper industry; etc.) are provided by the user for each region to drive the reference scenario. In addition, the user provides estimates of the existing stocks of energy related equipment in all sectors, and the characteristics of available future technologies, as well as present and future sources of primary energy supply and their potentials.

Using these as inputs, the TIMES model aims to supply energy services at minimum global cost (more accurately at minimum loss of total surplus) by simultaneously making decisions on equipment investment and operation; primary energy supply; and energy trade for each region. For example, if there is an increase in residential lighting energy service relative to the reference scenario (perhaps due to a decline in the cost of residential lighting, or due to a different assumption on GDP growth), either existing generation equipment must be used more intensively or new – possibly more efficient – equipment must be installed. The choice by the model of the generation equipment (type and fuel) is based on the analysis of the characteristics of alternative generation technologies, on the economics of the energy supply, and on environmental criteria.

Figure 1. Example of a Reference Energy System (RES)



TIMES is thus a vertically integrated model of the entire extended energy system. The scope of the model extends beyond purely energy-oriented issues, to the representation of environmental emissions, and perhaps materials, related to the energy system. In addition, the model is suited to the analysis of energy-environmental policies, which may be represented with accuracy thanks to the explicitness of the representation of technologies and fuels in all sectors.

In TIMES, the quantities and prices of the various commodities are in equilibrium, i.e., their prices and quantities in each "time period" are such that the suppliers produce exactly the quantities demanded by the consumers. This equilibrium has the property that the total economic surplus is maximized.

It is useful to distinguish between a model's structure and a particular instance of its implementation. A model's structure exemplifies its fundamental approach for representing a problem—it does not change from one implementation to the next. Therefore, all TIMES models exploit an identical underlying structure.

Thus, the structure of a TIMES model is ultimately defined by variables and equations created from the union of the underlying TIMES equations and the data input provided by the user. This information collectively defines each TIMES regional model database, and therefore the resulting mathematical representation of the RES for each region.

Data Requirements

The user input sets contain the fundamental information regarding the structure and the characteristics of the underlying energy system model. The user input sets can be grouped according to the type of information related to them:

- One dimensional sets defining the components of the energy system: regions, commodities, processes;
- Sets defining the Reference Energy System (RES) within each region;
- Sets defining the inter-connections (trade) between regions;
- Sets defining the time structure of the model: periods, time slices, time slice hierarchy;
- Sets defining various properties of processes or commodities.

The following is a list of the classifications of data needed to build instances of TIMES models, and the most common data parameters for each classification. For the purposes of brevity, TIMES documentation files include all the necessary information regarding input data needs to build a basic TIMES model. Following are parameters needed to build a typical energy system model using TIMES.

- Energy Service Demands
    - Demand projections for buildings and transportation sectors
    - Season/time-of-day pattern of the demand
- Energy Carrier Profiles
    - Input energy
    - Output energy
- Costs
    - Resource supply
    - Investment in new capacity
- Fixed and variable operations and maintenance (O&M)
    - Fuel delivery
    - "Hurdle" rates
- Technology Profiles
    - Resource supply steps and cumulative resource limits
    - Existing installed capacity and limits on new investment
    - Fuels in and out
    - Efficiency and Availability
- Environmental Indicators
    - Unit emissions per resource
    - Emission constraints/taxes per pollutant
    - Unit emissions per resource
    - Emission constraints/taxes by pollutant
- System and other parameters
    - Electric reserve margin
    - Season/time-of-day fractions describing the electrical load
    - Discount rate

Furthermore, the TIMES models include time periods for modeling horizon. TIMES is 'demand driven' in that feasible solutions are obtained only if all the specified end-use demands for energy services are satisfied for every time-period. Table 1 summarizes the parameters needed to build a typical energy system model using TIMES.

TIMES also distinguishes between two types of units for characterizing energy system technologies, activity, and capacity. Activity represents the use of a technology. Most technology activity is measured in petajoules (PJ). Capacity represents the size (installed capacity) of the technology stock and is measured according to the ability to provide for some amount of activity per unit time. Accordingly, capacities for most technologies are measured in petajoules per year

(PJ/yr). Electricity generation technology capacities are measured in gigawatts (GW), and transportation technology activities are measured in billions of miles per year.

Table 1. Model variables and corresponding data requirements

Variable Type	Input Requirements
End-Use Energy Service Demands	Projections for energy service demands for: **TRANSPORTATION**: Light-duty vehicle demand (bn-vmt-yr), bus transportation demand (bn-vmt-yr), heavy-duty short-haul truck transportation demand (bn-vmt-yr), passenger rail transportation demand (pn-passs-miles), medium-duty truck transportation demand (bn-vmt-yr), **RESIDENTIAL BUILDINGS:** space cooling (PJ/yr), space heating (PJ/yr), water heating (PJ/yr), lighting (billion lumens/yr), other electricity demand (PJ/yr), other natural gas demand (PJ/yr), **COMMERCIAL BUILDINGS**: space cooling (PJ/yr), space heating (PJ/yr), water heating (PJ/yr), lighting (billion lumens/yr), other electricity demand (PJ/yr), other natural gas demand (PJ/yr), *All demands include load shape for electric demand profiles.*
Energy Carriers *Any kind of entity that is a form of energy that is produced or consumed in the energy system (e.g., coal, refined oil, natural gas, gasoline, electricity)*	- Transmission efficiency - Transmission capacity - Investment cost - Operation and maintenance cost - Electricity Transmission and distribution cost - Reserve margin for electricity
Resource Technologies *Technologies that characterize raw fuels exported or imported into the energy system*	- Resource supply cost for each supply step - Cumulative Resource limits for an energy carrier for each period - Cumulative Resource limits for an energy carrier over the entire modeling horizon - cost and capacity limits of Resource transportation - cost of extraction and production of Resource
Process, Conversion, and Demand Technologies *Any kind of technology that can change the location, form, and/or structure of the energy carriers*	- New capacity investment cost - Fixed operation and maintenance cost - Variable operation and maintenance cost as a function of activity - Fuel delivery charges - Technical efficiency as a ratio between input and output - Technology investment availability year - Availability factor - Capacity utilization factors - Base year installed capacity - Upper bound on new capacity investment (if exists) - Upper bound on incremental new investment (growth rate) - Upper bound on total capacity installed over the modeling horizon - "Hurdle" rate for a technology
Emissions	- Emissions factor per unit of fuel consumed - Emissions factor for per unit of activity - Emissions factor for per unit of installed capacity - Upper bound for emission for each period

Variable Type	Input Requirements
	- Emission constraints over the entire modeling horizon - Emission constraints for any given sector

Scenario Framework

A scenario approach is appropriate to the assessment of long-term technological development in the energy system. Extended research, policy, and assessment horizons make business-as-usual extrapolations, conventionally used in shorter-term energy futures analyses, inappropriate. The technology innovation process is inherently uncertain and unpredictable. Over a period of decades, we simply cannot know which technologies will achieve fundamental breakthroughs and which will not. Changes in economic structures, consumer preferences, resource supplies, and other variables similarly lead to inherent unpredictability. With these factors in consideration, COMET-NYC was constructed under the energy system optimization modeling principles presented in DeCarolis et al. (2017). These principles include considerations such as minimizing model bias, setting clear spatio-temporal boundaries and goals, maximizing model and data transparency, and quality assurance.

The scenario approach to assessing technology futures requires that the menu of technology options being built into the models be appropriately connected to a set of driving forces to produce informative, internally consistent scenarios. Driving forces are the key elements that influence how the future turns out. Any scenario approach must identify the key driving forces that are expected to have an impact on the issues under consideration. Scenarios are then built from combinations of values or realizations of these driving forces. Major driving forces for the energy system technology futures include:

- Economic growth
- Population growth
- Changes in the structure of the economy, work, and recreation
- Land use and transportation policy
- Air pollution and environmental policy
- Oil and natural gas supply
- Consumer attitudes
- Rates and patterns of technological change

Future Technologies and Scenarios

For the technologies of interest in our scenario assessment, we collected estimates of technology costs, performance, and availability. Because many of these technologies are still under development, these data will be estimates of future cost, performance, and availability. There is therefore considerable uncertainty about these parameters. Indeed, it is this uncertainty that is the motivation for and source of our scenario assessment.

We are looking for a range of values that covers plausible future outcomes. Therefore, for each parameter it is best to gather data from several different sources and to provide some evaluation of the reliability of each source and the assumptions supporting each estimate. Having this well-documented range of values will allow us to construct scenarios that explore the range of possible futures. The researcher will determine, based on an analysis of available data, what data parameters will be used initially to input the technology into the database.

As an application of the TIMES approach, the Generative COMET model is enhanced by each of these features.

1.3. Model purpose and general architecture

1.3.1. Model purpose and philosophy

The purpose of the Generative COMET model is to provide cities with a versatile, comprehensive tool for evaluating integrated energy, emissions, and economic strategies at an urban scale. It can be applied in any city to provide a detailed analysis of local-scale energy and emissions strategies, highlighting the cumulative impacts of policy actions on their energy systems.

A key strength of the open-source Generative COMET is its flexible calibration capabilities, which allow it to leverage location-specific datasets—such as building topologies (e.g., residential 1-4 family housing units), local transportation patterns, and emissions data—to tailor the analysis to each city's characteristics. When detailed local data are unavailable, the model can draw upon regional or national databases, ensuring accurate yet adaptable simulations. This flexible calibration process enables cities to model scenarios that reflect their specific energy, infrastructure, and policy environments, giving planners actionable insights into the impacts of climate policies and technological adoption.

1.3.2. Spatial and temporal resolution

When modeling a city, two geographic areas are represented:

- CITY: The city boundaries, defining the area to be modeled.
- EGRID: The electricity grid supplying power to the city, corresponding to the EPA's Emissions & Generation Resource Integrated Database (eGRID)[4] subregion in which the city is located. (More information is provided in Section 1.10.)

The scope of the model covers both regions. This implies that all environmental impacts occurring within the EGRID subregion are assigned solely to the city. This assumption allows the model to increase electricity production outside the city for purposes of meeting the city's electricity demand. To correct for this assumption, the subset of electricity produced in the subregion that is consumed by the city is calculated and applied.

The base year of the model is 2019, which is used for calibration (see Section 1.4). From this starting point, the modeling horizon is divided into user-defined intervals, each represented by a midpoint or milestone year for result reporting. The Generative COMET currently defines the following milestone years: 2020, 2021, 2022, 2023, 2024, 2025, 2027, 2030, 2032, 2035, 2040, 2045, 2050, 2055, 2060, 2065, 2070, 2075, 2080, 2085, 2090, 2095, 2100.

Beyond these time periods, users can define intra-annual "time slices." In the Generative COMET, these represent three seasons: summer (S), winter (W) and intermediate (I), which are further broken down into four intraday periods: day am (DAM), day pm (DPM), night (N) and peak (P). In total, each year consists of 12 time slices. Table 2 summarizes these time slices and their respective fractions within the year.

Table 2. Definition of time slices

Abbreviation	Description	Time Fraction
S-DAM	Summer – Day AM	0.0822
S-DPM	Summer – Day PM	0.0957
S-N	Summer – Night	0.1532
S-P	Summer – Peak	0.0032
W-DAM	Winter – Day AM	0.0975
W-DPM	Winter – Day PM	0.1087
W-N	Winter – Night	0.1253
W-P	Winter – Peak	0.0027

[4] US EPA - United States Environmental Protection Agency (2025). Emissions & Generation Resource Integrated Database (eGRID). Information available online.

I-DAM	Intermediate – Day AM	0.0815
I-DPM	Intermediate – Day PM	0.1087
I-N	Intermediate – Night	0.1381
I-P	Intermediate – Peak	0.0032

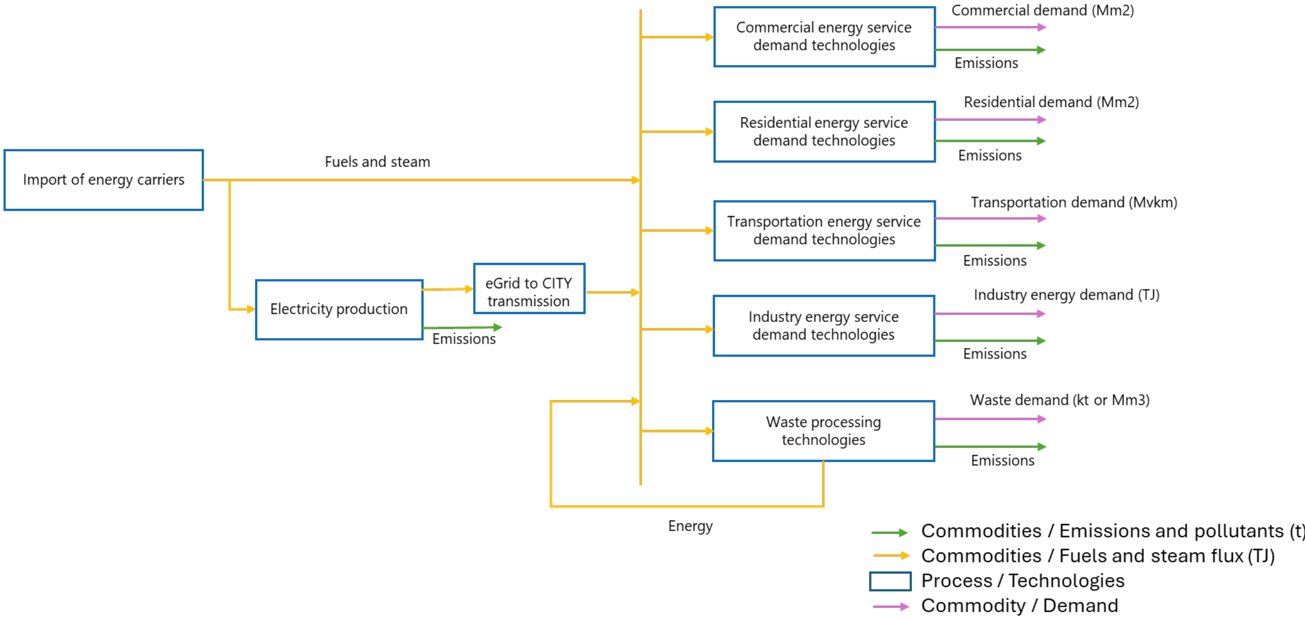
Initially, the model applied a six-time slice definition (three seasons—winter, summer, intermediate—and two intraday periods—day, night), but this limited the assessment of the relative competitiveness of electricity generation technologies. A nine-time slice model version introduced a peak load slice, reducing the overestimation of reserve margins. The current 12-time structure, which divides daytime hours into AM and PM, was adopted to more effectively capture load fluctuations, allowing different generation technologies to compete based on their intrinsic cost characteristics rather than external constraints.

The selection of 12 time slices enables a sufficient accurate representation of electricity demand fluctuations across different seasons and times of day, improving the ability to capture the operational dynamics of various generation technologies. By distinguishing between baseload, intermediate, and peaking technologies with greater granularity, the 12-slice structure minimises the need for artificial constraints on technology dispatch. This approach also results in more realistic reserve margin estimations. The introduction of additional slices, particularly peak and near-peak periods, allows peaking technologies to compete effectively and prevents an underrepresentation of their role in meeting short-term demand and potential implications for peaker emissions. Finally, aligning the time slice structure with available data sources, such as the end-use load shapes from the National Energy Modeling System (NEMS), ensures consistent demand-side representation.

1.3.3. Sectoral scope

The model structure breaks down the energy system into seven sectors: commercial, residential, industrial, transportation, electricity production, solid and water waste, and energy carrier imports. The overall architecture of the model is illustrated in Figure 1.

Figure 2. Model structure



*The International System of Units is used in the model, but the result's dashboard includes conversion factors to imperial units.

1.3.4. Data scope

The model represents all energy flows within the city, i.e., all energy consumed or produced within its boundaries. It also tracks the energy used to generate the electricity consumed by electricity customers located in the city, with all energy carrier flows expressed in terajoules (TJ). In terms of emissions, the model estimates GHGs from carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O), as well as "criteria pollutants" such as nitrogen oxides (NO_x), sulfur dioxide (SO_2), and particulate matter ($PM_2.5$ and PM_{10}). All GHGs and criteria pollutants are measured in tonnes (t) and account for both fuel combustion within the city and GHG emissions associated with electricity production for city consumption. All costs are expressed in 2023 thousands of U.S. dollars.

1.3.5. Data workflow

The model requires multiple datasets to accurately represent a city's energy system. To accommodate the variation in data availability across cities, three user experience pathways have been designed:

1. The *simplified approach* requires minimal city-specific information. The model relies on regional and national background datasets to infer city characteristics.
2. The *GHG inventory approach* requires data from a city's GHG inventory (including detailed energy consumption and/or emissions data). This calibration method is likely to be the preferred approach for many users, as it provides accurate results while maintaining a relatively simple user experience.
3. The *advanced approach* allows users to specify detailed settings to accurately match the city's energy system. This option is intended for experienced modelers.

Regardless of the pathway selected, the data are then used to populate two Generative COMET (GC) working files: one describing the current energy system, and another describing the new technologies available. Once these files are populated, the workbooks apply built-in formulas and links to generate sheets formatted for the VEDA platform database system. The updated sheets must then be transferred to the appropriate VEDA files, either manually via copy-paste or automatically using a Python script. Finally, these files are read by VEDA and then used to model the city and run the defined scenarios. This data workflow is summarized in Figure 2 and explained in more details in Section 1.4.

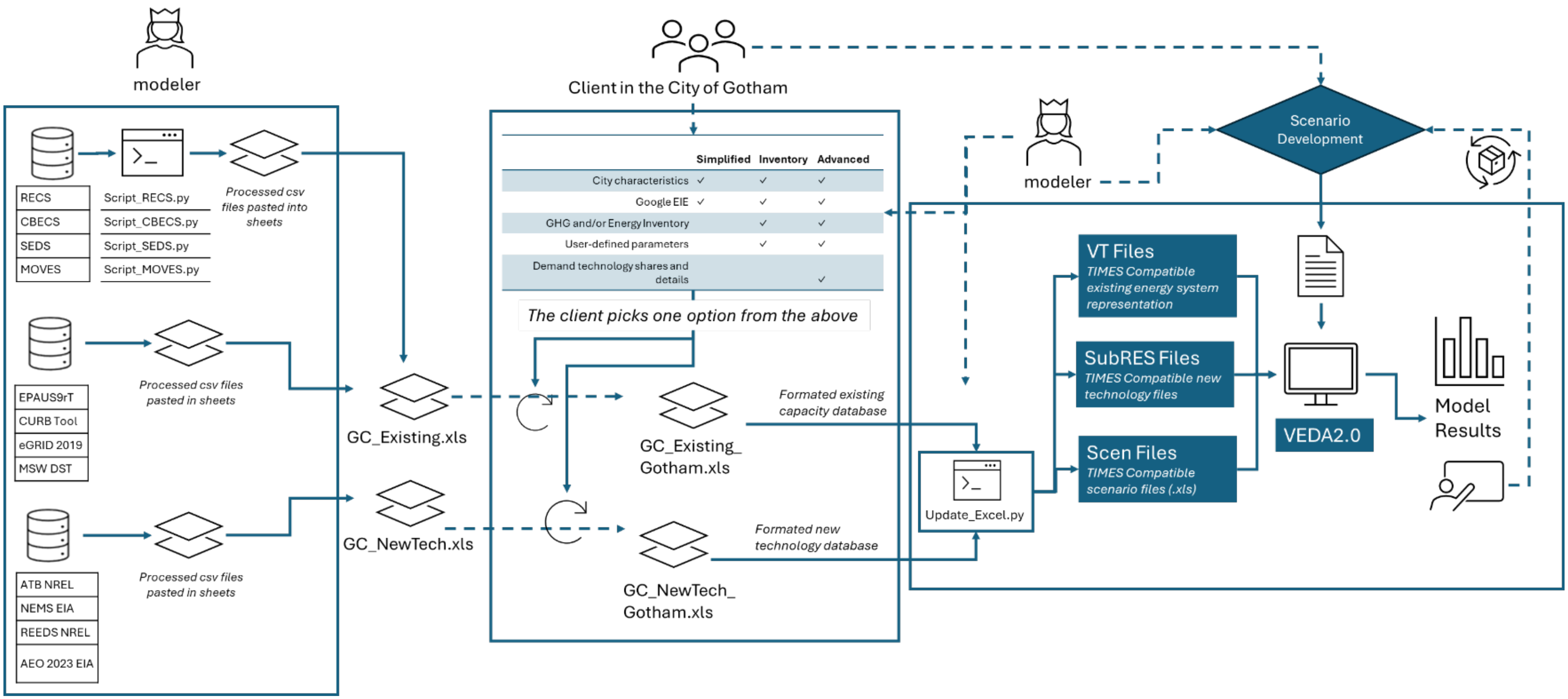


Figure 3. Data workflow for COMET Generative model

The model folder structure is illustrated with Figure 3.

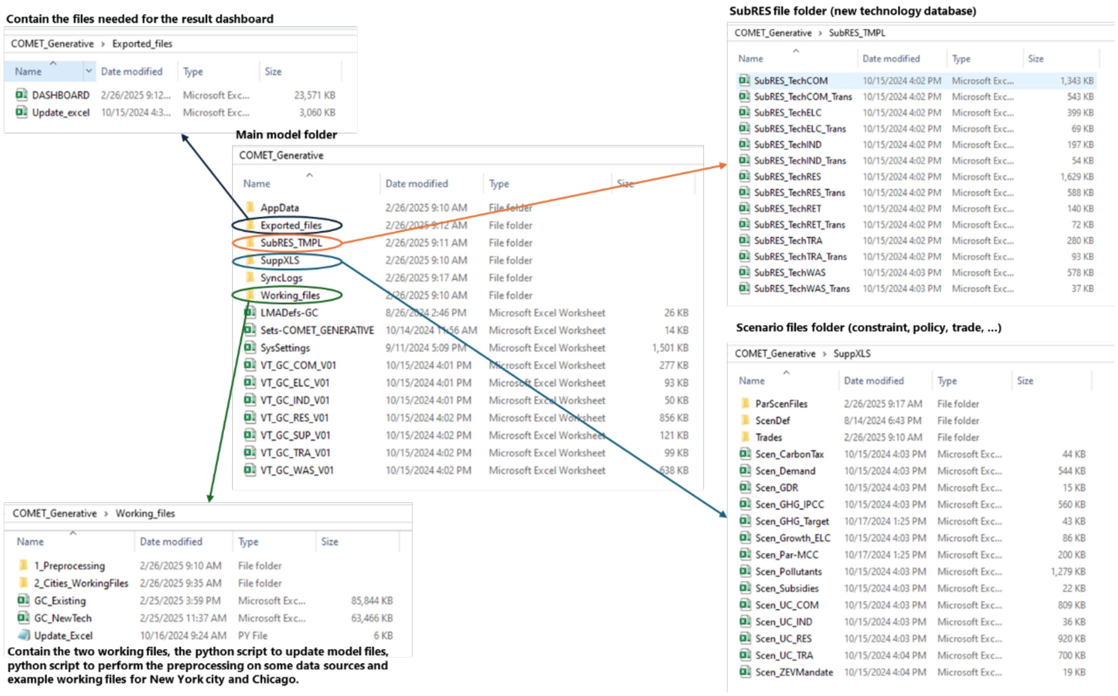


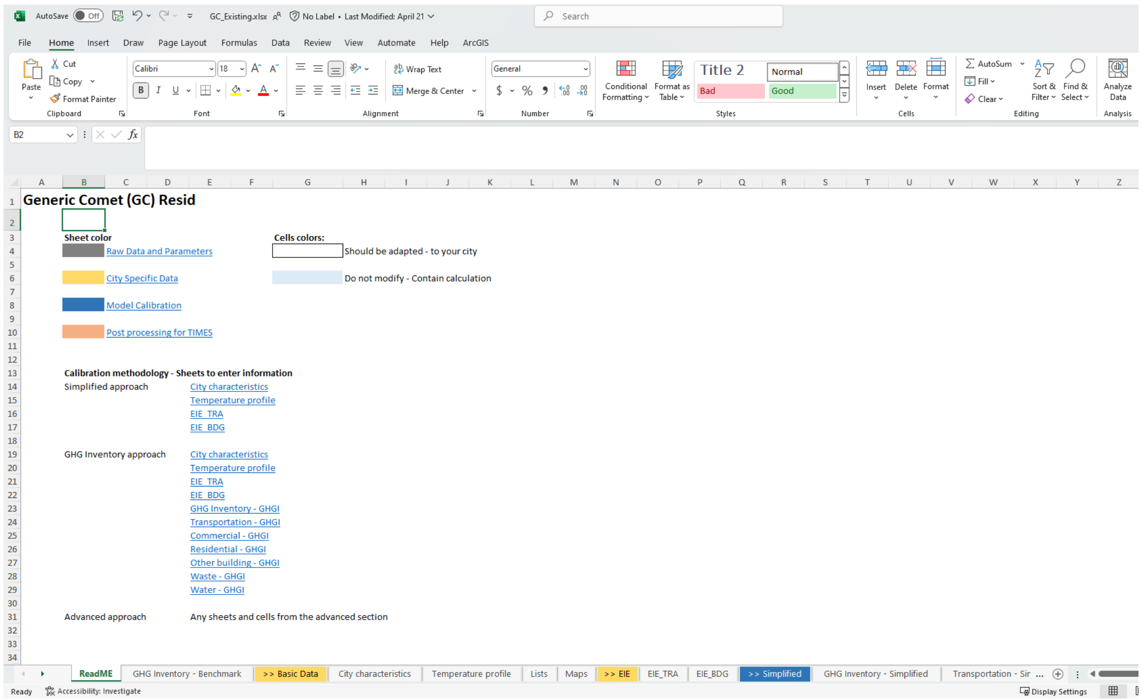
Figure 4. Structure of the model folder

1.4. The Generative COMET (GC) calibration process

A key modeling step is providing data necessary to calibrate the model, particularly for the base year. Calibration involves setting an initial reference year, known as the base year, from which the optimization of subsequent years begins. This process serves as the foundation for projection scenarios. A calibrated model incorporates city-specific data, inventory records, and stock information. The *GC_Existing* file is used to input calibration data and must be filled-in accordingly. This file is in the *Working_files* folder.

Within the *GC_Existing* file, the *ReadMe* sheet provides a color-coded legend and links to different sections of the file (Fig 5). Each section starts with a colored tab in workbook, named using the prefix “>>” followed by the section title. These introduction sheets include relevant data sources for each subsection. For instance, the GHG Inventory section begins with a blue sheet (used for model calibration) labeled >> *GHG Inventory*, which contains a link to New York City's GHG inventory.

Figure 5. ReadMe sheet of the calibration file



1.4.1. Data preprocessing

Various data sources are used in the model, most of which are directly copy-pasted in the *RAW data and Parameters* sections. However, five of them require preprocessing, i.e., a reformatting of the dataset into a specific format for the calibration workbooks (Table 3).

Table 3. Data sources used for calibration

Data source	Workbook	Worksheet
Freight Analysis Framework (FAF) (Version 4, April 2024) [1]	*GC_Existing*	*FAF4*
Commercial Buildings Energy Consumption Survey (CBECS) (2018 version, released December 2022) [2]	*GC_Existing*	*C_SH* *C_LT* *C_AC* *C_OT*
Motor Vehicles Emissions Simulator (MOVES) (version 3, extracted October 2021) [3]	*GC_Existing*	*MOVES_STATES*
Residential Energy Consumption Survey (RECS) (2020 version, released January 2024) [4]	*GC_ Existing*	*R_SH* *R_LT* *R_AC* *R_OT*

State Energy Data System (SEDS) (2022 version, released June 2024) [5]	*GC_Existing*	*Transportation_EB* *Commercial_EB* *Residential_EB* *Industrial_EB* *Degree_Days* *Population*
	GC_NewTech	*Energy_Price*
Google Environment Insights Explorer (Google EIE) [6]	*GC_NewTech*	*EIE_SOL*

These data are common across all cities and have already been preprocessed. Based on a city's location, the relevant sub-datasets are automatically pulled into the model. However, users can update these data with a more recent release if needed. Each data source has its own dedicated folder under the *Workingfile/1_Preprocessing* folder and a corresponding Python script for processing. To update a data source, users must place the new raw data file in the *Complete* folder and run the associated Python script. This script generates one or more CSV files in the *Processed* folder, which must then be copied into the *GC_Existing* working file, specifically in the sheet that matches the file name.

For convenience, users may choose to retain the original filename when updating data files. However, if a different name is chosen (e.g., to track release versions), the corresponding Python script must be updated accordingly. In all cases, before running the script, users should verify that the folder path and file name are correctly set for their computer. This information is conveniently located at the top of each script for easy checking and modification.

All other data used in the file retains the original source format and can be directly copied and pasted from the source into the *GC_Existing* working file.

1.4.2. Data processing in the working file

The *GC_Existing* working file is structured as follows (see Figure 5):

- **Raw Data and Parameters Sheets (grey sections)**: These sheets contain all raw data from external sources, as described in the previous section, used for model calibration. These sections are not city-specific and should not be modified unless the user wants to replace the raw data with more recent releases from the same sources. If updated, it is crucial not to alter the format of these sheets, as the data is linked to other sections of the model.

- **City-Specific Data Section (yellow sections)**: These sheets include location-specific data for the city being modeled and update the corresponding sheets accordingly.

- **Model Calibration Section (blue sections):** This section is where the calibration for the base year is conducted, processing all external data. While the model base year can be changed, it requires advanced VEDA-TIMES modeling skills. Three types of model calibration are possible, depending on the amount of data available to the user:

    - The Simplified approach: Only the *City Specific Data* section needs to be filled. External data sources are used to calibrate the energy consumptions in each sector.
    - The GHG inventory approach: In addition to the *City Specific Data* section, the *GHG Inventory* section must be filled with the city's GHG inventory data by entering energy consumption for each sector. This approach allows for the calibration of the base year to replicate the city GHG inventory. External sources are used to disaggregate the aggregated energy consumptions.
    - The Advanced approach: If the user has detailed data available, they can adjust any of the energy consumption data down to the most disaggregated level.

- **Post processing for TIMES (in orange):** These sheets collect all the necessary calibration data to generate the VEDA tables and produce the model.

1.4.3. Model generation

To generate the model, the sheets from the section >>*For Times* must be copied into the corresponding model sheets. There are two solutions to achieve this:

1. Automatically through a Python script: The script *Update_Excel.py* is located in the *Working_files* folder. As with any Python script, users should verify that the folder path is correct. (The current folder path is specified at the top of the script.)

2. Manually: The sheets should be manually copied into the correct model Excel files. Table 4 provides a detailed specification of which sheets need to be copied into which model files. In each model file, the sheets to be updated will be in the first two sections: >> *City Data* and >>*GC_Resid*. The other sections should not be modified, as they are specifically structured for VEDA to interpret them correctly.

Table 4. Existing calibration sheets mapping

	GC_Existing sheets								
Model Workbook	City characteristics	Demand	BDG Intensity	COM RESID	TRA RESID	WAS RESID	IND RESID	ELC RESID	RES RESID
VT_GC_COM_V01	●	●	●	●					
VT_GC_ELC_V01	●	●						●	
VT_GC_IND_V01	●	●					●		
VT_GC_RES_V01	●	●	●						●
VT_GC_SUP_V01	●								
VT_GC_TRA_V01	●	●			●				
VT_GC_WAS_V01	●	●				●			
SubRES_TMPL/SubRES_TechCOM	●								
SubRES_TMPL/SubRES_TechCOM_Trans	●								
SubRES_TMPL/SubRES_TechELC	●								
SubRES_TMPL/SubRES_TechELC_Trans	●								
SubRES_TMPL/SubRES_TechIND	●	●							
SubRES_TMPL/SubRES_TechIND_Trans	●								
SubRES_TMPL/SubRES_TechRET	●	●							
SubRES_TMPL/SubRES_TechRET_Trans	●								
SubRES_TMPL/SubRES_TechRES	●								
SubRES_TMPL/SubRES_TechRES_Trans	●								
SubRES_TMPL/SubRES_TechTRA	●								
SubRES_TMPL/SubRES_TechTRA_Trans	●								
SubRES_TMPL/SubRES_TechWAS	●								
SubRES_TMPL/SubRES_TechWAS_Trans	●								
SuppXLS/Trades/TRADEPARAM	●							●	
SuppXLS/Scen_CarbonTax	●								
SuppXLS/Scen_Demand	●	●							
SuppXLS/Scen_GDR	●								
SuppXLS/Scen_GHG_IPCC	●							●	
SuppXLS/Scen_GHG_Target	●								
SuppXLS/Scen_Growth_ELC	●							●	
SuppXLS/Scen_Pollutants	●				●			●	
SuppXLS/Scen_Subsidies	●								
SuppXLS/Scen_UC_COM	●			●					
SuppXLS/Scen_UC_IND	●						●		
SuppXLS/Scen_UC_RES	●								●
SuppXLS/Scen_UC_TRA	●				●				
SuppXLS/Scen_ZEVMandate	●								

The *VT_GC_** files include information about the existing stocks of each sector. The *SubRES_Tech * files* include the new technologies available in the model for future investments in each sector; they are in the *SubRES_TMPL* folder. The various Scen_*files include other components of the system such as demand projections, the global discount rate (GDR), emission coefficients, market share constraints in each sector and other policy constraints (carbon tax, carbon cap, subsidies, zero emission vehicle mandate). They are in the *SuppXLS* folder. The *TRADEPARAM* file includes the information related to trade movements and is located in the *Trade* folder of the *SuppXLS* folder.

1.5. The GC new technology database

1.5.1. Data processing in the working file

The new technology working file *GC_NewTech* follows the same format as the base year file and contains three main data sections:

- **The *Basic Data* section (in yellow)** includes city-specific data.
  - The *City characteristics* and *Lists* sheets contain the same information as in the *GC_Existing* file and should be copied from that file.
  - The *EIE_SOL* sheet needs to be updated using Google Environment Insights Explorer (Google EIE) [6]. It includes different types of information such as land area, solar potential, energy intensity, number of trips, etc.
- **The *Raw data and parameters* section (in grey)** contain various data required to build the new technology database.
  - Some summary sheets from the *GC_Existing* workbook need to be copied in this new technology workbooks: *BDG intensity*, *COM RESID*, *RES RESID* and *WAS RESID*. (Note: BDG stand for buildings; RESID stands for residual)
  - All the other sections contain data from external sources, with sources listed on the first summary sheet of the section. Users can update this data with more recent releases by copying the data. No pre-processing has been completed, so a simple copy-paste should suffice. However, users should ensure that formats remain consistent in order to maintain links with other sheets.
- **The *For TIMES* section (in orange)** contains the summary tables required to formally build the model. These sheets include all the necessary techno-economic parameters for building the new technologies in the model.

1.5.2. New technology generation

To generate the model, the sheets from the section >>*For Times* in the *GC_Exisiting file* must be copied into the corresponding model sheets. There are two solutions to achieve this:

- Automatically, using the *Update_Excel.py* Python script. This script will update the model files with data from both *GC_Existing* and *GC_NewTech* files.
- Manually, by copying the required sheets in the appropriate files. Table 5 specifies which sheets should be copied and where they should go.

Table 5. New technology calibration sheets mapping

Model Workbook	GC_NewTech sheets							
	ENE TECH	ELC TECH	TRA TECH	COM TECH	RES TECH	RET TECH	WAS TECH	IND TECH
VT_GC_COM_V01								
VT_GC_ELC_V01								
VT_GC_IND_V01								
VT_GC_RES_V01								
VT_GC_SUP_V01	●							
VT_GC_TRA_V01								
VT_GC_WAS_V01								
SubRES_TMPL/SubRES_TechCOM				●				
SubRES_TMPL/SubRES_TechCOM_Trans				●				
SubRES_TMPL/SubRES_TechELC		●						
SubRES_TMPL/SubRES_TechELC_Trans		●						
SubRES_TMPL/SubRES_TechIND								●
SubRES_TMPL/SubRES_TechIND_Trans								●
SubRES_TMPL/SubRES_TechRET						●		
SubRES_TMPL/SubRES_TechRET_Trans						●		
SubRES_TMPL/SubRES_TechRES					●			
SubRES_TMPL/SubRES_TechRES_Trans					●			
SubRES_TMPL/SubRES_TechTRA			●					
SubRES_TMPL/SubRES_TechTRA_Trans			●					
SubRES_TMPL/SubRES_TechWAS							●	
SubRES_TMPL/SubRES_TechWAS_Trans							●	
SuppXLS/Trades/TRADEPARAM								
SuppXLS/Scen_CarbonTax								
SuppXLS/Scen_Demand								
SuppXLS/Scen_GDR								
SuppXLS/Scen_GHG_IPCC	●							
SuppXLS/Scen_GHG_Target								
SuppXLS/Scen_Growth_ELC		●						
SuppXLS/Scen_Pollutants	●	●	●					
SuppXLS/Scen_Subsidies								
SuppXLS/Scen_UC_COM				●				
SuppXLS/Scen_UC_IND								●
SuppXLS/Scen_UC_RES					●			
SuppXLS/Scen_UC_TRA			●					
SuppXLS/Scen_ZEVMandate								

The *VT_GC_** files include the information about the existing stocks of each sector. The *SubRES_Tech * files* include the new technologies available in the model for future investments in each sector; they are located in the *SubRES_TMPL* folder. The various *Scen_** files include other component of the system such as demand projections, the global discount rate (GDR), emission coefficients, market share constraints in each sector and other policy constraints (carbon tax, carbon cap, subsidies, zero emission vehicle mandate). They are located in the *SuppXLS* folder. The *TRADEPARAM* file includes the information related to trade movements, located in the *Trade* folder of the *SuppXLS* folder.

1.6. Residential building sector

1.6.1. Energy service demand

Residential buildings (RES) are divided into 16 demand segments, based on three parameters:

- Type of building: apartment (APA) or house (HOU)
- Income level: low (LO), lower-middle (LM), upper-middle (UM), high (HI)
- Building age/vintage: existing (Old) or new (New)

These acronyms serve as the naming convention for the different demand segments in the model, with "RES" (for residential) at the beginning and DMD (for demand) at the end. For instance, the demand for new, upper-middle-income apartments will be named RESAPAUPNewDMD. The demand values are expressed in millions of square meters (Mm^2). The surface area of existing buildings remains constant over time, while the surface area of new buildings will grow in line with the projected population growth of the city. The population growth rate is a parameter input by the model user, and therefore the source of this data is determined by the modeler. Unless location specific data exist, the user can refer to the US Census data to fill this field.

The model considers ten different end uses for each demand segment:

- Heating and cooling: space heating (SHEUE), space cooling (SCEUE) and water heating (WHEUE)
- Food management: cooking (CKEUE), refrigeration (REEUE) and freezing (FREUE)
- Clothing management: clothes washing (CWEUE) and clothes drying (CDEUE)
- Lighting (LTEUE)
- Other end uses (OEEUE)

Energy consumption for these end uses follows the same naming convention as the demand segments, with "DMD" replaced by the corresponding end use. For example, the energy consumption for cooking in existing low-income houses is named RESHOULOOldCKEUE.

The lighting requirement, LTEUE, is expressed in millions of lumens per millions of square meters (Mlm/Mm^2). All other end uses are expressed in TJ/Mm^2.

For space heating and cooling, end-use consumption is calculated based on 2019 data and projected to 2050 and 2100, considering the city's heating/cooling degree days. The city has flexibility to input their own projections For all other end uses, consumption per square meter is assumed to be constant over time.

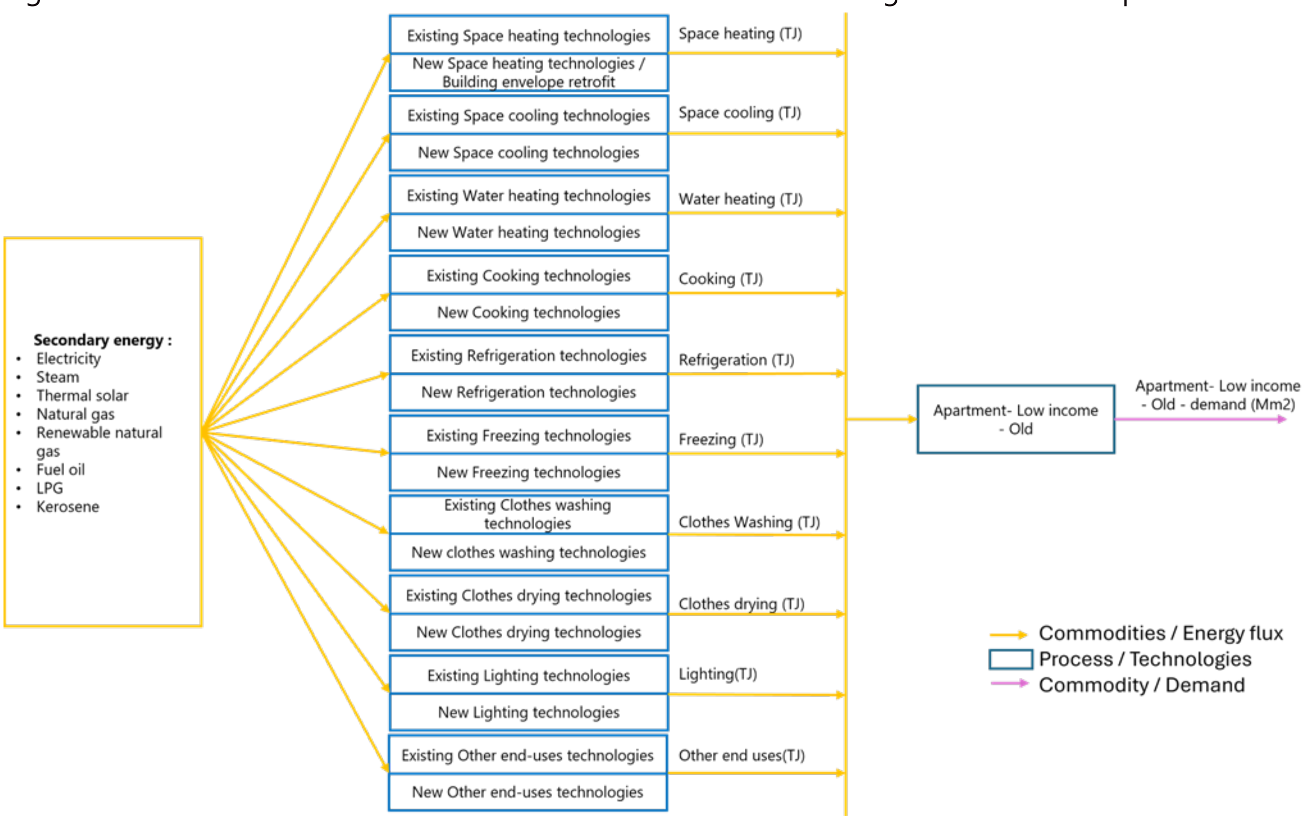
Eight fuel types are available for the residential sector: electricity, natural gas, fuel oil, liquefied petroleum gas (LPG), steam, kerosene, thermal solar, and renewable natural gas.

The model also considers building envelope retrofits to reduce space heating and cooling needs, thereby partially offsetting the energy consumption for heating and cooling.

The overall structure for existing low-income apartments is illustrated in Figure 6. The same structure applies for all 16 demand segments.

1.6.2. Architecture

Figure 6. Architecture of the residential sector for existing low-income apartments



1.6.3. Data sources

If a user selects the simplified calibration method, the RECS data source [4] is applied to calibrate energy usage by end use. These data help allocate energy consumption across the different end uses (such as space heating, cooling, cooking, etc.) when the GHG inventory calibration method is selected. Some technical parameters for existing technologies (such as efficiency or utilization factors) are extracted from the EPAUS9rT model database [7].

The residential demand module of NEMS [8] is used to identify new technologies available to replace existing ones, as well as their techno-economic parameters. This source also provides the parameters for building envelope retrofit technologies, allowing the model to simulate energy savings from upgrades.

1.7. Commercial building sector

1.7.1. Energy service demand

In the commercial sector (COM), buildings are divided into 16 different demand segments, based on two key parameters:

- Building use: Hotel (HOT), Retail (RET), Office (OFF), Hospital (HOP), Educational (EDU), Warehouse (WAR), Municipal (MUN) and Other (OTH)
- Vintage: Existing (Old) and New (New)

As for the residential sector, demand for each commercial segment is named by combining the building's use and vintage, with COM (for commercial) at the beginning and DMD (for demand) at the end. For example, the demand for new offices is named COMOFFNewDMD. Demand is expressed in millions of square meters (Mm^2). The surface area of existing buildings remains constant over time, while the surface area of new buildings grows with the forecasted GDP growth of the city.

The model considers eight different end uses:

- Heating and cooling: space heating (SHEUE), space cooling (SCEUE) and water heating (WHEUE)
- Food management: cooking (CKEUE) and refrigeration (REEUE)

- Building management: lighting (LTEUE) and ventilation (VTEUE)
- Other end uses (OEEUE)

Energy consumption for these end uses follows the same naming convention as the demand segments, with DMD replaced by the corresponding end use. For example, the energy consumption for space heating in existing hotels is named COMHOTOldSHEUE.

All lighting requirement (LTEUE) values are expressed in millions of lumens per millions of square meters (Mlm/Mm²). All other end uses are expressed in TJ/Mm².

For space heating and cooling, end-use consumption is calculated based on 2019 data and projected to 2050 and 2100, taking into account the city's heating/cooling degree days. For all other end uses, consumption per square meter is assumed to be constant over time.

There are six fuels available for the commercial sector: electricity, natural gas, fuel oil, liquified petroleum gas (LPG), steam and renewable natural gas.

Options for retrofitting the envelope of commercial buildings are currently not included in the model due to the lack of reliable data on potential savings. However, they can be added by the user, per the approach indicated for residential buildings.

The overall structure is represented in Figure 7 for existing hospitals. The same structure applies to all 16 demand segments.

1.7.2. Architecture

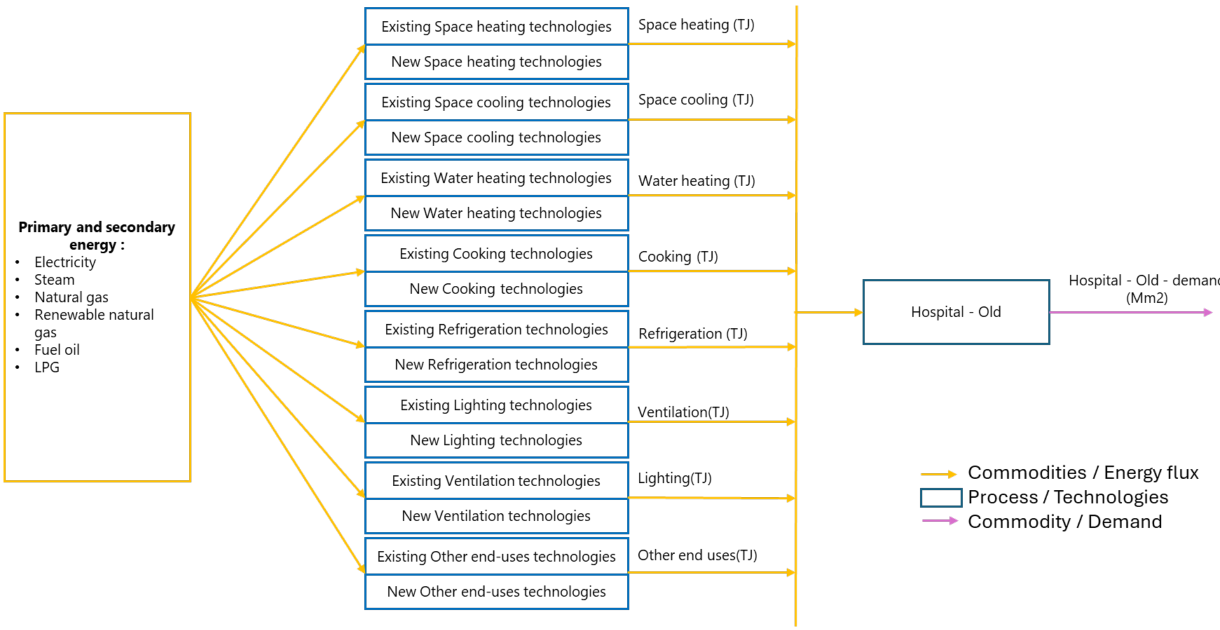


Figure 7. Architecture of the commercial sector for existing hospitals

1.7.3. Data sources

If the user selects the simplified calibration method, the CBECS data source [2] is used to calibrate energy usage by end use. This data helps allocate energy consumption across the different end uses when the GHG inventory calibration method is selected. Techno-economic parameters for existing technologies (as efficiency or utilization factors) are extracted from the EPAUS9rT model database [7].

The commercial demand module of NEMS [8] is used to define the technologies available to replace existing ones, along with their techno-economic parameters.

1.8. Industrial sector

1.8.1. Energy service demand

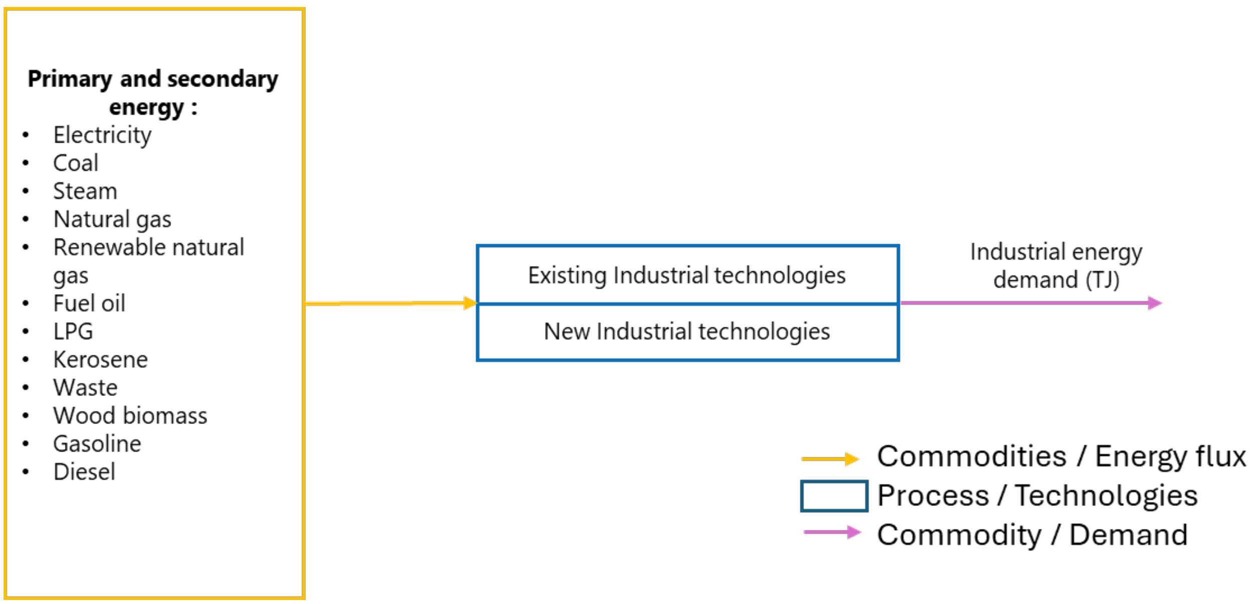
In the industry sector, demand is handled via a single aggregated demand segment named INDAINDMD, which represents the overall energy consumption of the city's industrial sector

(Figure 6). Fuels that can be consumed in the industry sector include: coal, petroleum coke, liquified petroleum gas (LPG), fuel oil, natural gas, renewable natural gas (RNG), kerosene, waste, wood biomass, gasoline, diesel, steam, and electricity. Industrial demand (in TJ) is projected to increase with the city's GDP growth.

A retrofit technology is implemented to represent potential efficiency gains in industry, since the demand is modeled directly in energy units. This technology can partially meet industrial energy demand replacing fuel usage. The model includes the maximum potential for efficiency improvements.

1.8.2. Architecture

Figure 8. Architecture for the industrial sector



1.8.3. Data sources

If the user selects the simplified calibration method, SEDS [5] is used to calibrate industrial existing energy consumption, with city adjustment factors calculated based on the state's manufacturing GDP. Otherwise, consumption is calibrated by applying the user-defined value.

Since the total industrial demand is expressed directly in energy consumption units, new technologies are just similar copies of the existing ones. The industrial sector is represented in a very aggregated manner; the different industrial sub-sectors are not modelled specifically. Fuel switching is driven by energy prices rather than by technology costs and technical attributes. The model does not capture process emissions related to industrial sector.

1.9. Transportation sector

1.9.1. Energy service demand

The transportation sector (TRA) is divided into 20 demand segments:

- Passenger transport (PAS), for which the available technologies are:
  - Personal vehicles: car (CAR), truck (LDT) and motorcycle (MOT)
  - Public transport: transit bus (TBU), school bus (SBU), other buses (OBU), rail (RAI), tram (TRM) and subway (SUB)
  - Active transport: bicycle (BIC) and walking (WAL)
- Freight transport (FRE), for which the available technologies are:
  - light commercial trucks (CLT), refuse truck (RFT), single unit long-haul truck (SLH), single unit short-haul truck (SSH), combination long-haul truck (CLH) and combination short-haul truck (CSH)
- Other transport (OTH) for which the modeled technologies are:
  - off-road (OFR), waterborne (WAL) and aviation (AIR).

Each technology in the model is referred to using the abbreviation of the transport type (passenger, freight or other) and the vehicle abbreviation, preceded by TRA for transportation. For example, passenger motorcycle is referred to by the acronym TRAPASMOT.

All the transportation demands are expressed in millions of vehicle kilometers traveled (MVkm) in the model but converted in millions of vehicle miles traveled (VMT) in the result dashboard, such as in MOVES [3].

For each vehicle type, different fuels are available: gasoline, ethanol, diesel, renewable diesel, biodiesel, compressed natural gas, renewable natural gas, electricity, hydrogen, LPG, fuel oil, aviation gasoline, jet fuel for aviation and sustainable aviation fuels.

Different propulsion technologies are available depending on the vehicle type:

- For on-roads vehicles: conventional internal combustion engine, hybrid (conventional plus electric motor with small battery that cannot be plugged in for recharge), plug-in hybrid

(conventional plus electric motor with a larger battery that can be plugged in for recharge), battery electric (electric motor only)
- For off-roads vehicles: mainly conventional propulsion, with some biofuel-based and hydrogen-based propulsion systems available in certain cases.

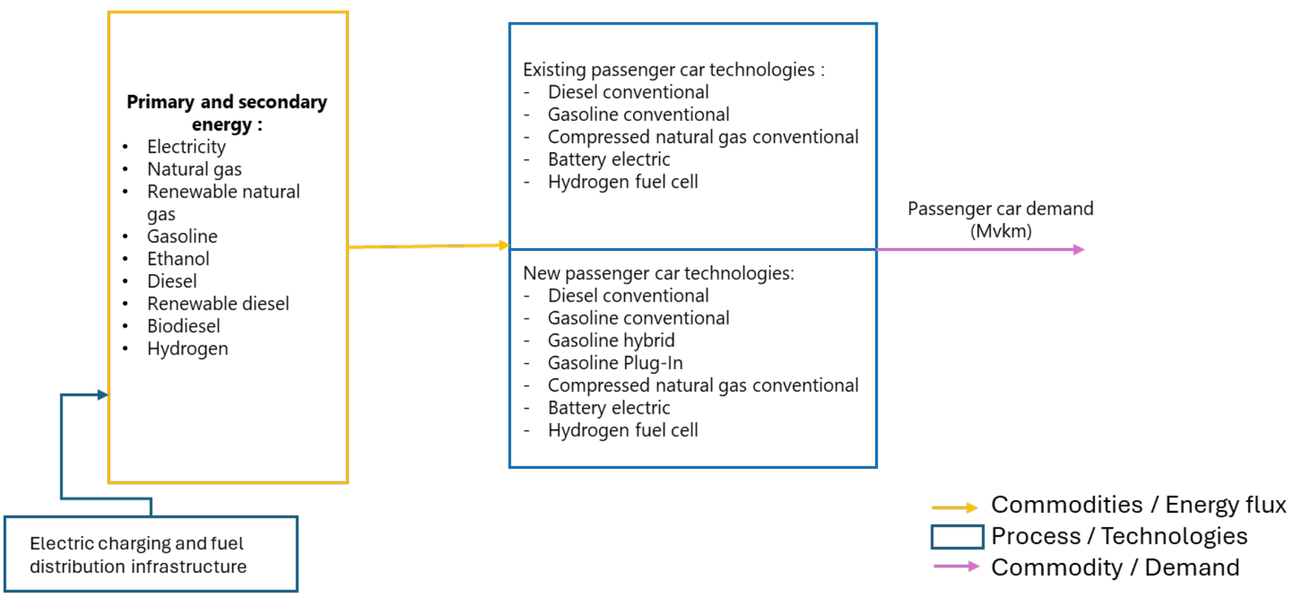
Different types of charging infrastructure are modeled with their specific techno-economic parameters:

- For passenger vehicles: 120V and 240V charging for both residential and commercial use.
- For freight vehicles: slow and fast charging options.

1.9.2. Architecture

The overall structure is illustrated in Figure 7 for passenger car demand. The same structure applies to all transportation segments, with available technologies adapted for each segment.

Figure 9. Architecture for passenger car demand segment



1.9.3. Data sources

For existing technologies, lifetime and operational (OPEX) costs are extracted from the 2022 Transportation Annual Technology Baseline (ATB) [9], and vehicle efficiency data comes from

MOVES [3] and the CURB tool [10]. If the user selects the simplified calibration method, energy consumption is calibrated using data from MOVES [3]. Otherwise, it is calibrated by applying the user-defined value.

The techno-economic parameters for new technologies are also taken from the 2022 Transportation ATB [9].

1.10. eGRID Power sector

Electricity production in the model is represented by the eGRID2019 subregion to which the city belongs. In some cases, the eGRID subregion aligns almost perfectly to city boundaries, while in other cases the city is only a small part of the subregion. Both situations have their advantages and disadvantages. For subregions that closely match the city, there is no uncertainty about which percentage of electricity production is consumed by the city. However, this limits the potential for new electricity production plants located within the city. In cases such as these, the advanced model is recommended. On the other hand, when the subregion is much larger than the city, there is more space for locating electricity generation facilities but greater uncertainty about how much of the electricity produced in the subregion is allocated to the city. More information is provided below.

1.10.1. Electricity production

1.10.1.1. Centralized generation

The model includes 14 different energy sources for electricity production, each associated with a specific technology:

- Natural gas (NGA)
- Renewable natural gas (RNG)
- Fuel oil (FOI)
- Coal (COA)
- Biomass (BIO)
- Other bioenergy (OTB)
- Waste (WAS)
- Waste heat (HET)
- Steam (STM)
- Nuclear (URN)

- Geothermal (GEO)
- Wind (WND)
- Hydro (HYD)
- Solar (SOL)
- Storage (STG)

Electricity production technologies are named using a specific format: "ELC" followed by the abbreviation of the energy source, then "GRD" for centralized or "DCE" for distributed, the technology abbreviation, and the first available year. For existing technologies, the default first year of availability is set to 2019. For example, an existing natural gas centralized combustion turbine is named: ELCNGAGRDCTU19. The set of technologies is too extensive to list here, but all technology abbreviations are provided in the *Dict* sheets in both the *GC_Existing* and *GC_NewTech* working files.

If electricity production within the eGRID subregion is insufficient to meet the city's electricity consumption, the model imports electricity from another eGRID subregion using a defined import technology. For this imported electricity, the applied emissions factors are the average of the emissions factors within the city's eGRID subregion, which are kept constant through the modeling horizon (the user can change these values to account for different generation mix pathways).

At all times, the eGRID subregion must match not only the electricity consumption of the city but also the consumption of the rest of the subregion. The demand for this area is assumed to remain constant over time and is defined as the difference between electricity production in the subregion and the city's electricity consumption in 2019. If this demand is negative (i.e., the subregion needs to import electricity to meet the city's 2019 demand), the demand is set to zero. This means that the growth in electricity generation represents the growth in electricity consumption for the city only.

There is an endogenous trade setup between the EGRID region and the CITY region, representing the transmissions lines for the city. The existing capacity is calibrated to the 2019 electricity consumption level, and the model can invest in new capacity if the existing capacity is insufficient to satisfy the consumption growth. Transmission grid losses are accounted for using specific TIMES parameters and default values. Default electric grid parameters are:

- Transmission grid losses of 5% (applied to the centralized electricity commodity in the EGRID region)
- Peak reserve margin of 20% for centralized electricity in both CITY and EGRID regions.

1.10.1.2. Distributed generation

In addition to centralized electricity generation, the model also incorporates distributed electricity generation from wind and solar. This is calibrated using electricity production data from commercial, residential, municipal, and industrial photovoltaic sources. Combined heat and power (CHP) options are not included in this version of the model.

The new electricity production technology database (located in the *SubRES_TechELC* file within the *SubRES_TMPL* folder) contains a broad range of options for all the input fuels mentioned previously. Each new technology includes a maximum capacity limit, which represents the maximum technical potential for that type of electricity production. For instance, for solar PV applications, we relied on Google Environmental Insights Explorer to come with these bounds.

1.10.2. Architecture

The architecture for centralized and distributed electricity production are presented in Figure 8 and Figure 9, respectively.

Figure 10. Architecture of centralized electricity production

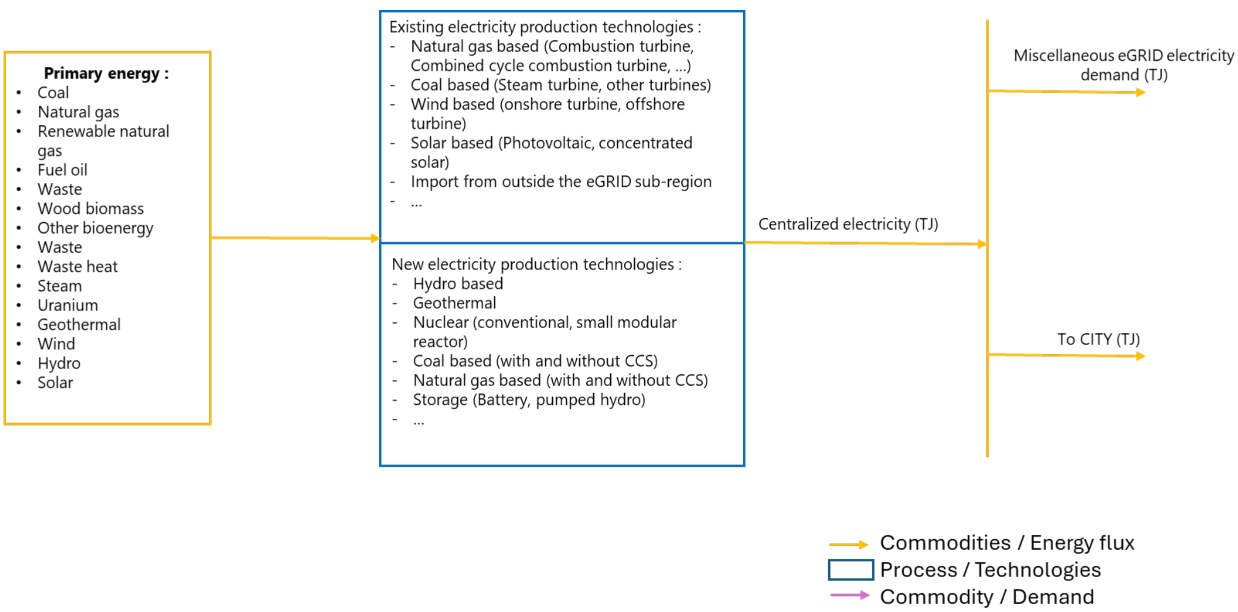
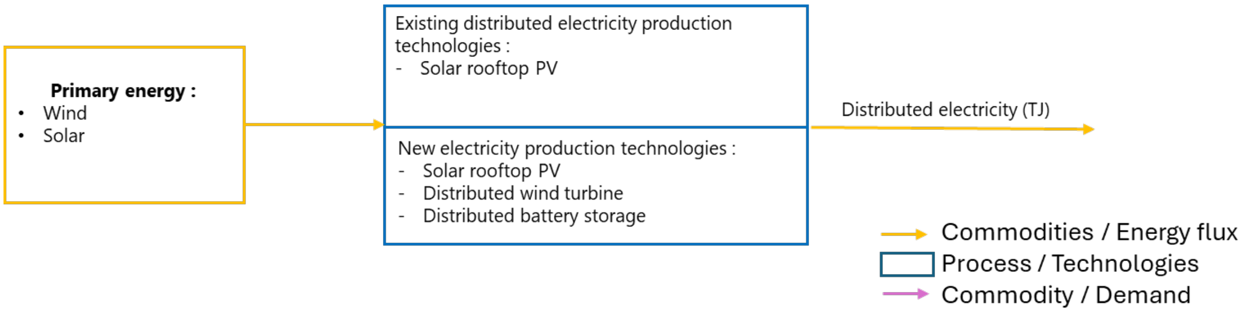


Figure 11. Architecture of distributed electricity production



1.10.3. Data sources

For the base year calibration, different types of centralized electricity production plants are modeled and calibrated using eGRID2019 [11] for production, capacity and efficiencies values. Distributed electricity production is calibrated with production data from SEDS [5].

Techno-economic parameters for new technologies are extracted from 2024 Electricity ATB Technologies [12], while maximum capacity parameters specific to the city's eGRID subregion are taken from supply curve data from Regional Energy Deployment System (ReEDS) [13].

1.11. Energy supply and prices

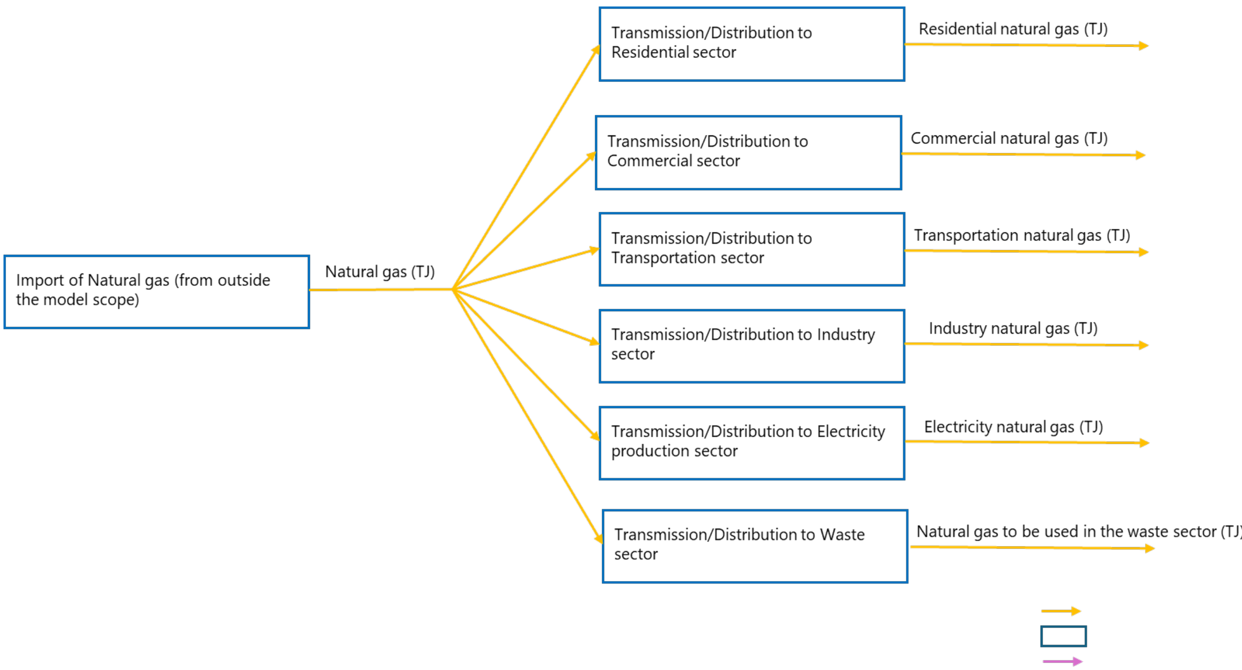
1.11.1. Energy type and prices

The distribution systems for each energy commodity used in the city are modeled through a naming process that transforms the general commodity (e.g., diesel = DSL) into a sector-specific commodity that can only be used within its corresponding sector (e.g., diesel for transportation = TRADSL). This process involves an activity cost, which represents the variable cost in $/GJ transformed, corresponding to the sector-specific energy price.

Except for electricity, which is traded within the EGRID region, all other commodities are outputs of a mining and/or extraction process. The architecture of the distribution system for natural gas is presented in Figure 10. The architecture is similar for all fuels.

1.11.2. Architecture

Figure 12. Architecture of the energy supply sector (example of natural gas)



1.11.3. Data sources

Historical sectoral energy prices are sourced from SEDS [5] , while projected values are extracted from the Annual Energy Outlook 2023 (AEO2023) [14]. Users can update this information using more recent version of the Annual Energy Outlook as they become available.

1.12. Waste sector

1.12.1. Waste generation

The waste module includes solid waste and wastewater treatment:

- For water (WAT): fresh water (FWA), annual precipitation (APR) and wastewater treatment (TRT).
- For solid waste (SOL): paper (PAP), glass (GLA), metals (MET), plastics (PLA), food (FOO), yard trimming (YAR) and all other wastes (AOT).

The demand for each type of waste is named by using the sector abbreviation (WAS), followed by the category of waste (solid or water), the type of waste, and ending with "DMD" for demand. For instance, the demand for glass waste is named WASSOLGLADMD. The projections for waste demand are based on population growth.

There is currently no process related to fresh water and annual precipitation. Wastewater treatment has four modeled pathways using placeholder data that can be updated by the user:

- Treatment plant without de-nitrification and without anaerobic digesters (NOT).
- Treatment plant without de-nitrification and with anaerobic digesters (WAD).
- Treatment plant with de-nitrification and without anaerobic digesters (WDN).
- Treatment plant with de-nitrification and with anaerobic digesters (WAN).

All plants consume electricity and natural gas for their operation, and the associated combustion emissions are accounted for. Each treatment pathway also has N_2O process emissions, originating from the treatment plant and the effluents. The emission levels depend on whether a de-nitrification process is in place. Additionally, treatment plants equipped with anaerobic digesters will produce biogas. The overall process architecture is illustrated in Figure 11.

For solid waste, there are five different treatment pathways included (the amount of waste generated annually is fix exogenously by the user and the model endogenously optimises the selection of treatment options).

- Solid waste incineration (INC): available for all type of solid waste. It produces electricity and steam, with process CO_2, CH_4 and N_2O emissions accounted for.
- Solid waste anaerobic digestion (ADG): available for paper, food, yard trimming and all other wastes. It consumes electricity and diesel to treat the waste, leading to combustion (from energy use) emissions and process emissions. It will also produce biogas.
- Solid waste compost (CMP): available for paper, food, and yard trimming. It consumes electricity and diesel to treat the waste, with process emissions also accounted for.
- Solid waste recycling (REC): available for paper, glass, metals, and plastics. It consumes diesel to treat waste, and there is no process emissions considered.
- Solid waste landfill (LDF): available for all type of waste. This pathway is separated into two main processes:
    - The waste added to the landfill every year. This process consumes electricity and diesel and accounts for the emissions of this energy use. It produces two outputs:

the biogas produced from the waste during the first year, and the amount of waste remaining at the end of the year that enters the storage process.

- The storage process that represents accumulated waste over the years, indicating that waste remains in the landfill for multiple years and produces biogas over time. This process is modeled as an inter-period storage, with a storage loss parameter that represents the waste decay factor each year.
- Biogas is composed of mainly methane, carbon dioxide and trace elements

When biogas is produced, either from wastewater or solid waste treatments, there are five different pathways available, generating process emissions (the carbon neutrality assumption is used for combustion emissions for bio-based technologies):

- Venting (VEN): The biogas is released into the atmosphere, releasing mostly methane, accounted as process emissions.
- Flaring (FLA): The biogas is combusted, carbon dioxide is released into the atmosphere. A relatively small amounts of methane is produced and accounted as process emissions.
- Used in cogeneration facilities (COG): The biogas is used to produce both electricity and steam, with relatively small amounts of methane emitted, accounted for as process emissions.
- Used in electricity production facilities (POW): The biogas is used for electricity generation, resulting in relatively small amounts of methane being accounted for as process emissions.
- Upgraded into renewable natural gas (PUR).

The architecture of the overall solid waste treatment process is represented in Figure 12 while Figure 13 detailed the landfill portion.

The waste treatment processes are named using the abbreviation of the sector (WAS), the category of waste (solid or water), the type of waste and the treatment pathway. For example, the technology recycling glass is named WASSOLGLAREC.

1.12.2. Architecture

Figure 13. Wastewater module architecture

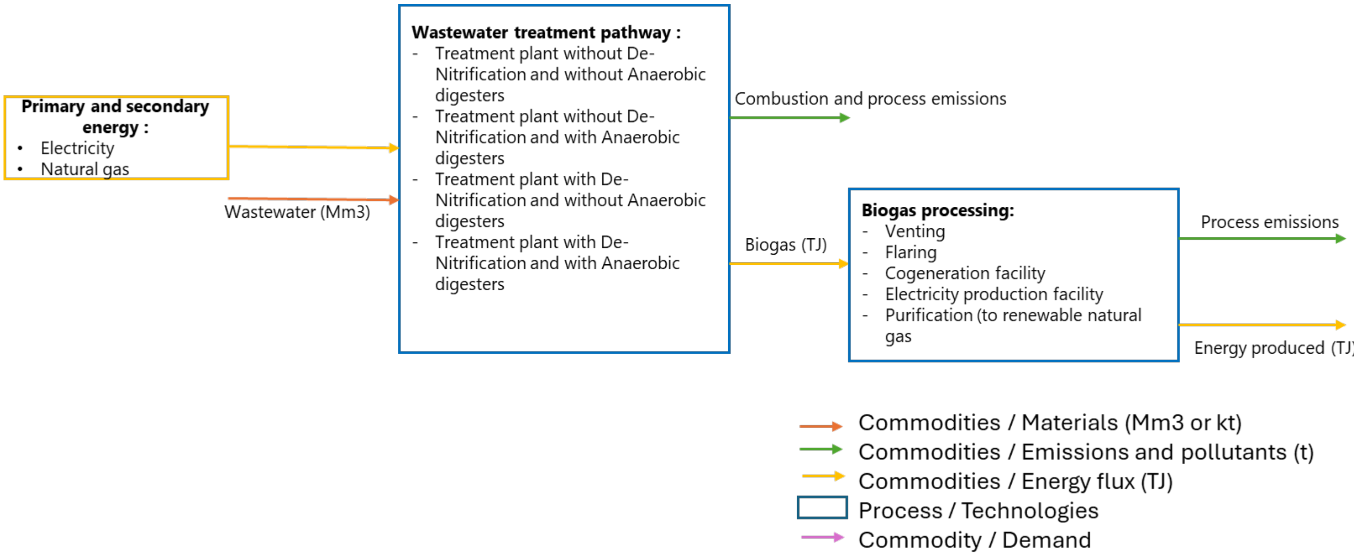


Figure 14. Solid waste module architecture

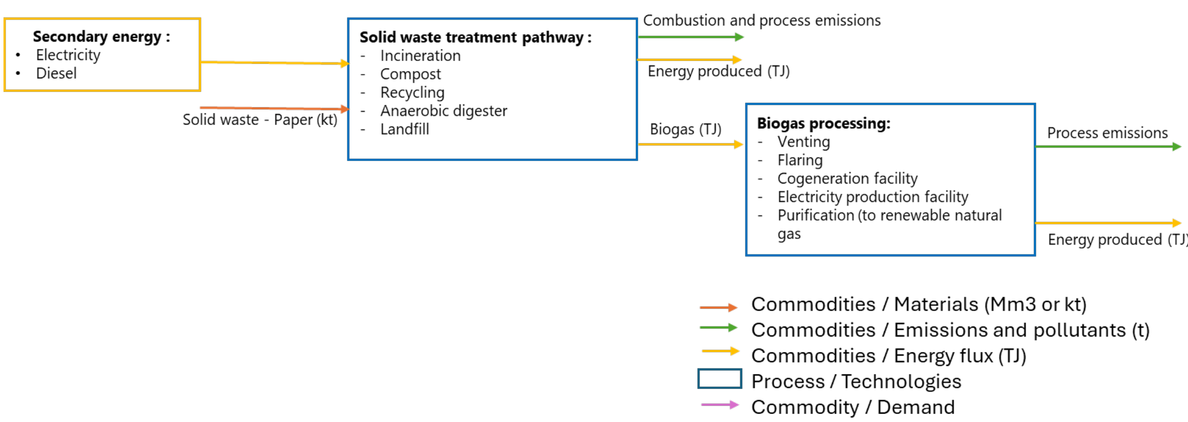
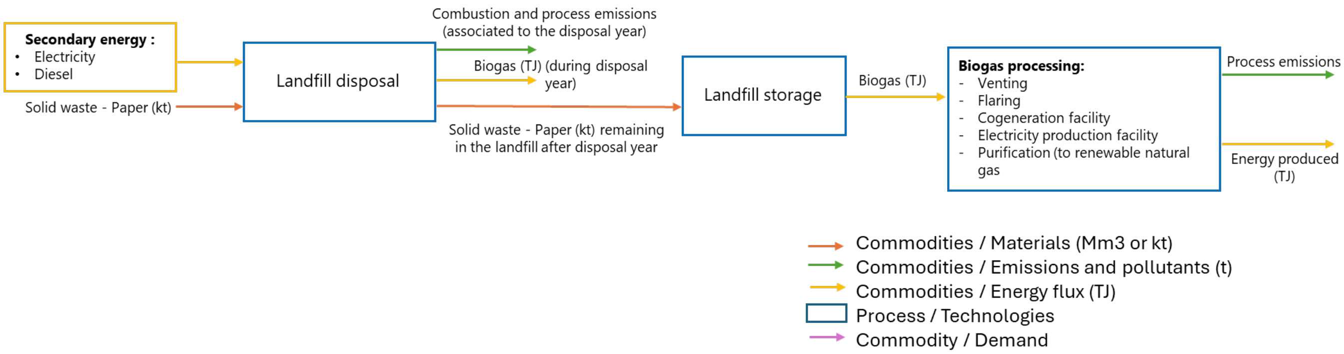


Figure 15. Landfill architecture



1.12.3. Data sources

All data are extracted from the Solid Waste Optimization Life Cycle Framework [15], except for combustion emission factors, which are sourced from the GHG Emission Factors Hub [16].

The data describing the biogas processing pathways are based on default values documented in the calibration sheets of the chosen method.

1.13. GHG emissions and air pollutants

1.13.1. Reported GHG emissions

GHG emissions (i.e., CO_2, CH_4 and N_2O) from fuel combustion in the city are accounted for. Bio-based fuels are considered to result in zero net fossil-CO2 emissions, and are excluded from the GHG accounting. Hydrogen consumption is considered GHG-free, as it is assumed to be imported from outside the city's boundary, where GHG emissions occur during hydrogen production (no specific assumptions are made regarding the production mix). Electricity consumption is also considered GHG-free, as emissions are accounted for during electricity production in the eGRID region or in the CITY region for in-city generation. In cases where electricity is produced using carbon capture and storage (CCS), only emissions occurring after the capture process are considered; emissions from carbon sequestration and transportation are likewise not tracked. If electricity needs to be imported from outside the eGRID subregion, average emissions from the eGRID in 2019 are used for all three GHGs (CO_2, CH_4 and N_2O).

Combustion GHG emissions are categorized by the sector in which the fuel is used (COM, RES, ELC, TRA, IND and WAS) and by the type of GHG (CO_2, CH_4 and N_2O). Additionally, total GHG emissions are quantified in tonnes of CO_2 equivalent (tCO_2e) for each sector. It is assumed that 1 tonne of CH_4 is equivalent to 25 tonnes of CO_2, and that 1 tonne of N_2O is equivalent to 298 tonnes of CO_2 following the IPCC Fourth Assessment Report [17]. All GHG emissions are reported in metric tonnes.

For the waste sector, the model also tracks process emissions associated with waste processing: WASCO2P, WASCH4P, WASN2OP, and WASCO2e (in CO_2e).

Non-GHG air emissions (NO_x, SO_2, $PM_{2.5}$, PM_{10}) are also quantified in each sector - COM, RES, IND, WAS, TRA and ELC - depending on the fuels combusted. All air pollutants are reported in tonnes.

1.13.2. Data sources

GHG emissions factors are taken from the GHG Emission Factors Hub [16], pollutants factors are sourced from the EPAUS9rT model database [7], where necessary, and waste process emissions factors are primarily adapted from the Solid Waste Optimization Life Cycle Framework [15].

1.14. Templates for user-defined constraints

For each sector, default market share constraints have been implemented. These are compiled in one file per sector, named *Scen_UC_**. These constraints regulate the pace at which technology transitions occur, ensuring that the replacement of existing technologies takes place at a consistent and coherent rate.

Residential: For each building type, end-use service type (e.g., space heating, space cooling, cooking), and fuel type category, both lower and upper market share constraints are implemented for each combination. Future fuel shares are derived from the 2019 base year and relaxed over time to provide the model more flexibility for the optimisation. The fuel share for each category, excluding electricity, must fall between 80% and 120% of its 2019 share by 2025, and between 20% and 150% by 2030. By 2050, the share can range from zero to three times the 2019 share. For electricity, the share must not fall below 75% of the 2019 share, with the upper bound gradually increasing to a 100% market share by 2050. These constraints only apply to fuels that were consumed in 2019 and for which a "New technology" option exists.

Commercial: The same set of market share assumptions and constraints applied to the residential sector are implemented in the commercial sector.

Industrial: Market share constraints for fuels used to meet industrial demand follow a similar trend as in the residential sector. Specifically, the fuel share must remain between 80% and 120% of the 2019 levels by 2025, between 20% and 150% by 2030, and between 0 and 3 times the 2019 levels by 2050. For electricity, the share must stay at least at the 2019 level and can gradually increase to reach a 100% market share by 2050.

Transportation: Four market share constraints are implemented to both existing and new vehicles in the transportation sector for calibration purposes.. Market share values can be adjusted by the users.

- Conventional vehicles: For all vehicles that are neither 100% electric nor 100% hydrogen, their share can decrease to reach zero by 2050. Intermediate bounds are set at 80% of the 2019 share by 2025 and 20% by 2030. Simultaneously, these vehicle shares may increase to reach three times its 2019 share by 2050, with intermediate bounds of 120% of the 2019 share by 2025 and 150% by 2030).
- Battery electric vehicles: The share of electric vehicles cannot go below the 2019 share but can increase up to 100% in 2050.
- Internal combustion engine (ICE) vehicles: The share of ICE vehicles cannot reach zero before 2040.
- Hybrid and plug-in vehicles: The combined share of hybrid and plug-in vehicles cannot exceed 40% of the number of ICE vehicles.

Electric grid (file *Scen_Growth_ELC*): instead of implementing a market share constraint, a growth constraint has been applied. This constraint limits the capacity growth of a certain type of electricity production (i.e., hydro-based generation capacity) to a maximum of 20% per year.

The following templates for user-defined constraints were implemented:

- **Carbon tax constraint:**
    - File name: *Scen_CarbonTax*
    - Description: This file outlines the implementation of a carbon tax on emissions. The emissions subject to the tax are identified in the *Cset_CN* column (this column includes the commodity set to which the constraint applies and filters the commodity names). The tax rate can be specified in the *Data* sheet. Currently, the Durbin carbon tax proposal is utilized [18].
- **Technology subsidies:**
    - File name: *Scen_subsidies*

  - Description: This file provides an example of how subsidies for technologies can be implemented. It allows users to specify both annual and cumulative budgets. In the *UC_SUB* sheet, subsidies can be defined in two ways: as a percentage of the investment cost (using the attribute NCAP_ISPCT) or as a fixed amount deducted from the investment cost (using the attribute NCAP_ISUB). The *Pset_XX* columns select the processes eligible for these subsidies. The *Max Sub* sheet controls whether a maximum budget for subsidies can be applied, either cumulatively over the entire time horizon or on an annual basis. All relevant data are summarized in the *Data* sheet.
- **Zeros emissions vehicle sales mandate:**
  - File name: *Scen_ZEVMandate*
  - Description: This file implements an example of a zero-emissions sales mandate, which sets policies requiring a minimum share of new vehicles to be a specific type. For passenger cars and light-duty trucks, the mandate specifies that by 2030, at least 10% of new vehicles must be battery electric or hydrogen-based. This share increases to 50% by 2050.
- **GHG targets:**
  - File name: *Scen_GHG_Target*
  - Description: This file exemplifies an emission constraint that requires total emissions to remain below a specified limit. The emission levels are configured in the *Pre-processing* sheet, with the *REF_data* sheet containing the emission levels from a reference scenario. The *UC sheet* implements the constraint, and the *Cset_CN* column lists the emissions covered by the constraint.

1.15. Result dashboard

A dashboard is available to display the most important information about the city. This dashboard is in the *Exported_files* folder. The process for updating the dashboard is explained in the User Guide section of this document.

1.16. How to build a marginal abatement cost curve

Marginal abatement cost curves (MACCs) are valuable tools for analyzing the cost-effectiveness of various GHG reduction measures across different sectors of the economy. These curves illustrate the cost of the last tonne of carbon dioxide equivalent (in US$/tCO2e) reduced under different target or tax levels. Each point on the curve represents different abatement measures, with the corresponding cost. By comparing the costs and effectiveness of these strategies, MACCs help policymakers prioritize interventions based on their economic efficiency.

To build a MACC, several model runs are conducted with varying emissions targets or carbon taxes. Each emissions target or carbon tax level corresponds to a step in the curve. In the context of TIMES and VEDA, the MACC generation is built using the parametric scenario functionality which allows the definition of several runs using a single scenario file. This scenario file defines a constraint or update that is applied for each step of the curve. A parametric scenario called *Scen_Par-MCC* has been added to show how to establish a MACC between two runs: GC_REF and GC_NZ. The curve spans 30 steps, with the first step being the GC_REF scenario and the 29th step being the GC_NZ scenario, slightly extending beyond the GC_NZ endpoint to explore potential solutions after it.

In VEDA, once the summary parametric scenario file is synchronized, files are created for each level of the curve in a *ParScenFiles* folder. A parametric group additionally needs to be created, containing all these scenario files. A single run can that is using this parametric group can then be executed. After this run is complete, it will launch multiple runs, each corresponding to a different step in the curve. An example from the *Scen_Par-MCC* scenario file is defined under the name GC_REF_NZ_MCC.

After executing the runs, the results are analyzed and formatted for a specific year to produce the MACC. The total and sector-specific MACCs are derived by extracting the marginal cost of carbon (for each step of the curve) and the total emissions reduction (between each step of the curve), or by sector emissions reduction (between each step of the curve).

SECTION 2

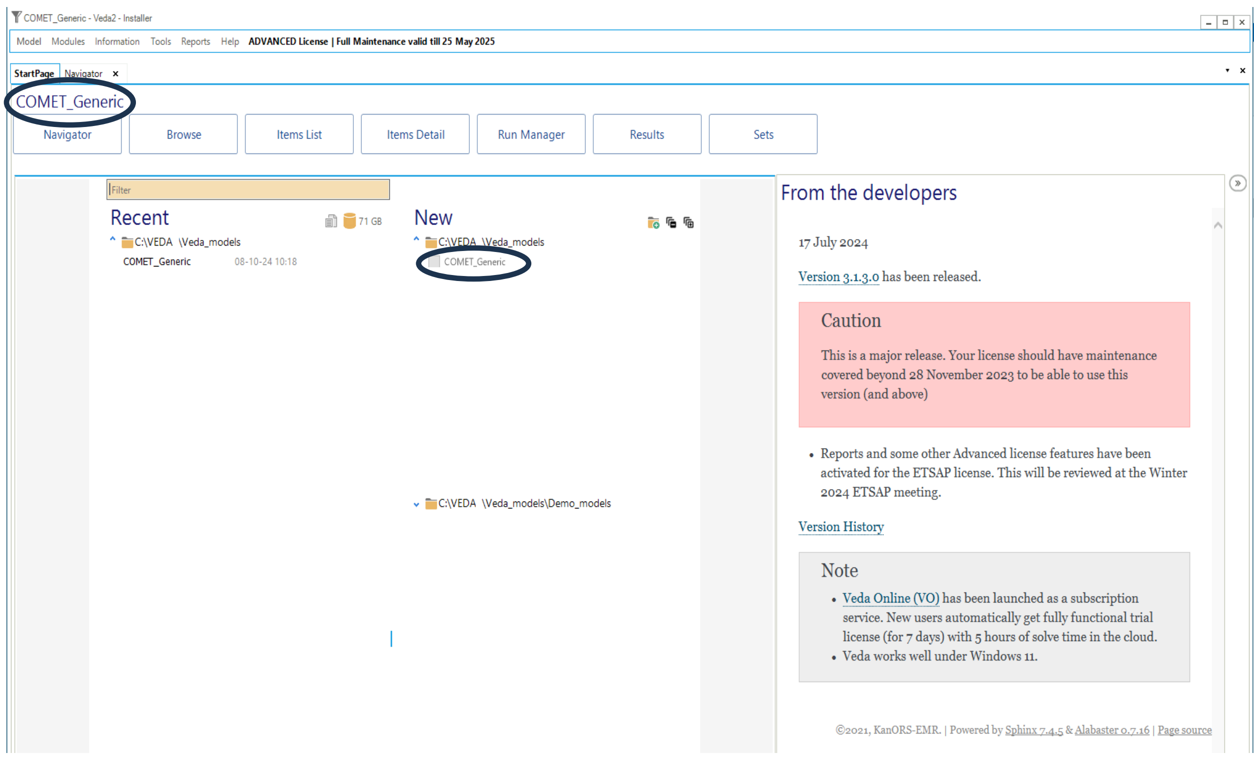
2. User guide

2.1. Model installation

To initiate model set-up, the first step is installing the VEDA and GAMS software. A tutorial can be found here: https://veda-documentation.readthedocs.io/en/latest/pages/Getting%20started.html

After installing VEDA, users can copy the full model folder to their *VEDA/Veda_models* folder. Then, to open the model, the following steps are taken:

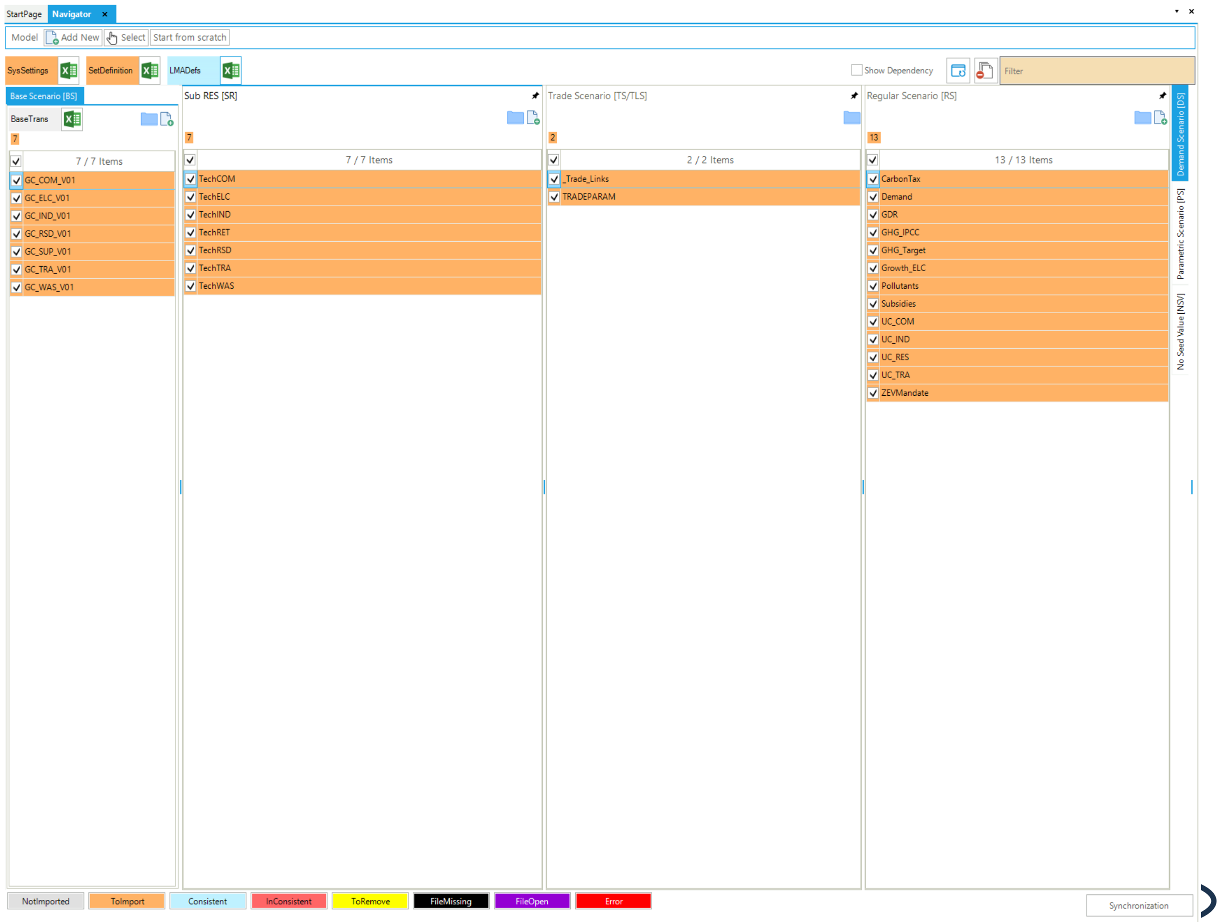
1. Open VEDA
2. On the StartPage, double click on the model of interest. Users will then see the relevant model name displayed at the top (above the horizontal buttons)



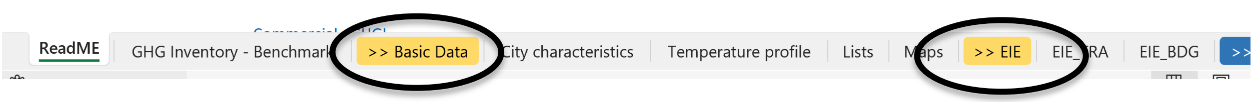
2.2. How to calibrate the model: Three approaches

To calibrate the model to a specific city, follow these steps:

1. The model's base year calibration is done in the *GC_Existing* working file, located in *COMET_Generative_v1/Working_files*. City-specific data must be entered into this file to complete the calibration. There are three different calibration methods: simplified approach (Section 2.2.1), GHG inventory approach (Section 2.2.2), and the advanced approach (Section 2.2.3).
2. Input city-specific data into the *COMET_Generative_v1/Working_files/GC_NewTech* working file (following the approach described in section 1.5.1) to calibrate the new technology database.
3. Either run the *Update_Excel.py* Python script located in *COMET_Generative_v1/Working_files/* or manually copy and paste each >> *For TIMES* sheet from the two previous working file into the model workbook/s, following the mappings listed in Table 4 and Table 5.
4. Synchronize the full model in the VEDA Navigator.



2.2.1. Simplified approach



This approach only requires generic city data, and the calibration is performed using external data sources. This approach requires to complete the >> *Basic Data* section the *GC_Existing* working file as follows:

- *City characteristics* sheet: All white cells with a black border need to be filled. Maps defining the region categories can be found in the *Maps* sheet. The list and description of the different regions and scenarios than can be selected can be found in the *Lists* sheet.
- *Temperature profile* sheet: Data from U.S. Climate Normals [17], should be downloaded and entered into this sheet.

Then, the >> *EIE* section needs to be updated with the city data from Google EIE [6]. These data are used to update the city transportation and building files.

2.2.2. GHG inventory approach

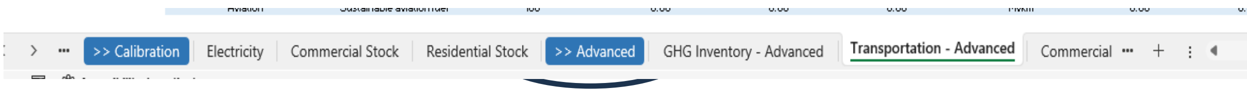
The GHG Inventory approach is designed to align the base year of the model with the city's emissions inventory. Examples of GHG inventory sources include New York City [18] or Chicago [19]. The >> *Basic Data* and >> *EIE* sections of the *GC_Existing* working file also need to be completed for this approach. Additionally, the >> *GHG Inventory* section needs to be filled in. As before, only white cells with black border need to be completed, while the cells in light blue contain calculations.

In the >> *GHG Inventory* section, data is required in several sheets:

- *GHG Inventory – GHGI* sheet: Energy consumption data by type of fuel for residential, commercial, municipal buildings and industry.
- *Transportation – GHGI* sheet: Energy consumption by type of fuel and type of vehicle.
- *Commercial – GHGI* sheet: Commercial building area per capita for each building type.
- *Residential – GHGI* sheet: Residential share of units by building type, or a city-specific correction factors based on state data.
- *Other building – GHGI* sheet: Energy consumption for street lightnings and distributed electricity production (Solar PV production is estimated from external sources but can be overwritten if more accurate data is available).
- *Waste – GHGI* sheet: Various waste parameters can be adjusted, such as the amount of solid waste or the waste type shares. If not specified, default values will be used.
- *Water -GHG* sheet: Default and average data are used but can be overwritten if more specific data is available.

In this approach, the *GHG Inventory – GHGI* sheet summarizes energy consumption by fuel for each sector and corresponding emissions. The model's base year should align with these figures.

2.2.3. Advanced approach

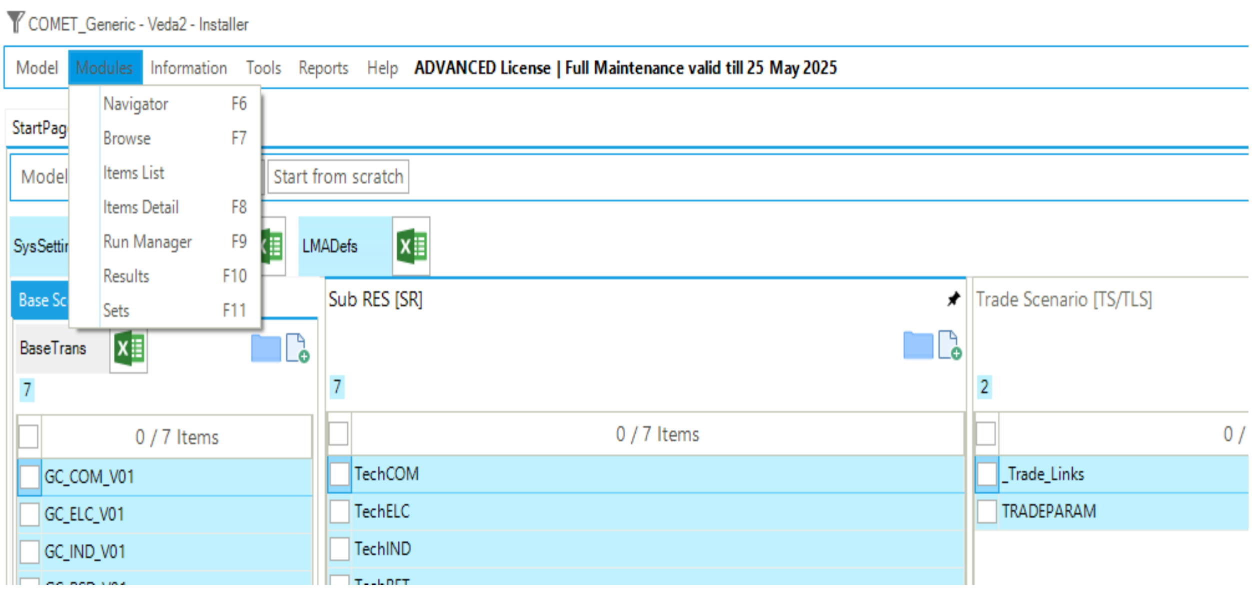


For the advanced calibration method, any data in the >> *Advanced section* can be adjusted from the default computed values. This method should only be selected if the user has access to very detailed energy and emissions data for the city and enough expertise to calibrate the model using these data.

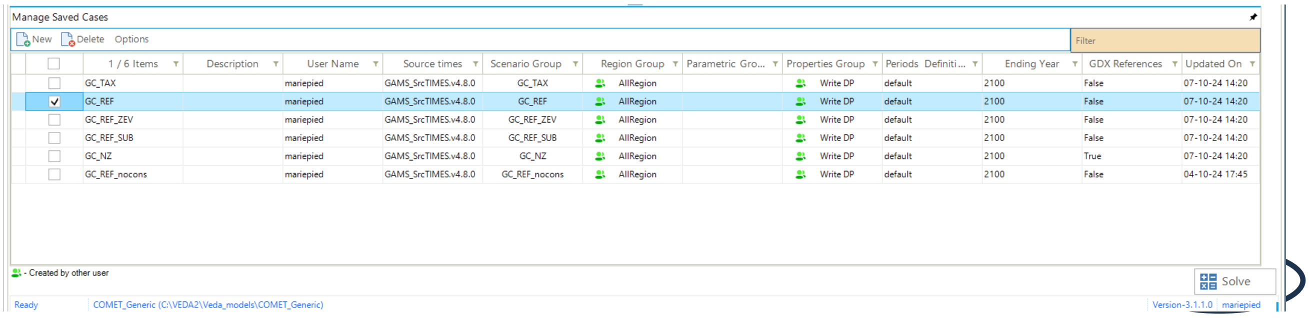
2.3. How to run the model

After the model is fully synchronized, follow these steps to run the model:

1. Open the Run Manager



2. Select a scenario and click the Run Scenario button.



It takes a minimum of approximately three minutes for a scenario to solve. Six scenarios are defined and included by default in the model:

- GC_REF, which is a reference scenario.
- GC_REF_SUB, which is the GC_REF scenario plus the *Scen_Subsidies* file.
- GC_REF_ZEV, which is the GC_REF scenario plus the *Scen_ZEVMandate* file.
- GC_REF_nocons, which is the GC_REF scenario without any market shares constraint.
- GC_TAX, which is the GC_REF scenario plus the *Scen_CarbonTax* file.
- GC_NZ which is the GC_REF scenario plus the *Scen_GHG_Target* file.

3. After the scenario has finished running, you can now look at the results.

2.4. How to compute results in the dashboard

2.4.1. Steps to update the dashboard

To update the dashboard, the first step is to choose the scenarios that will be displayed:

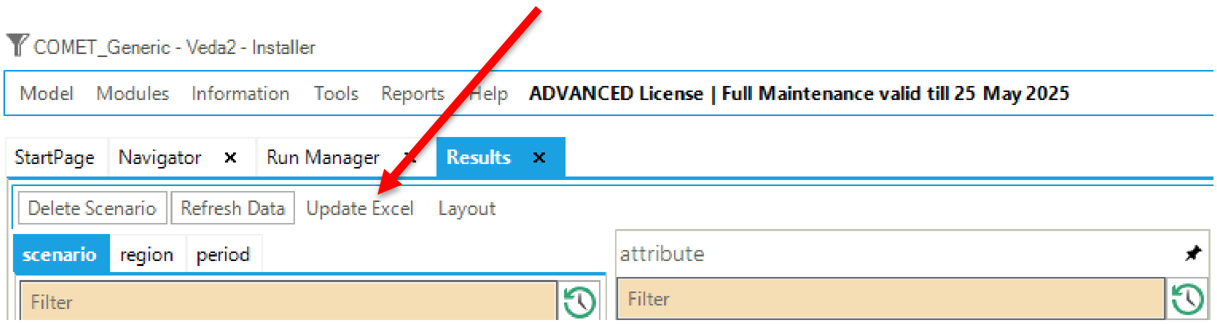
- Open the file *Update_excel.xlsx*, in the *Exported_files* folder, and go to the *Scenarios* sheet.

1. In the first column, indicate the names of the scenarios to analyze, as they are named in VEDA. There is a maximum of 10 scenarios that can be analyzed at the same time.
2. In the second column, indicate the name you want displayed in the tables and charts of the dashboard.
3. Save and close the *Update_excel.xlsx* file.

Note: The other sheets of this file should not be modified as they are specifically formatted to be updated by VEDA.

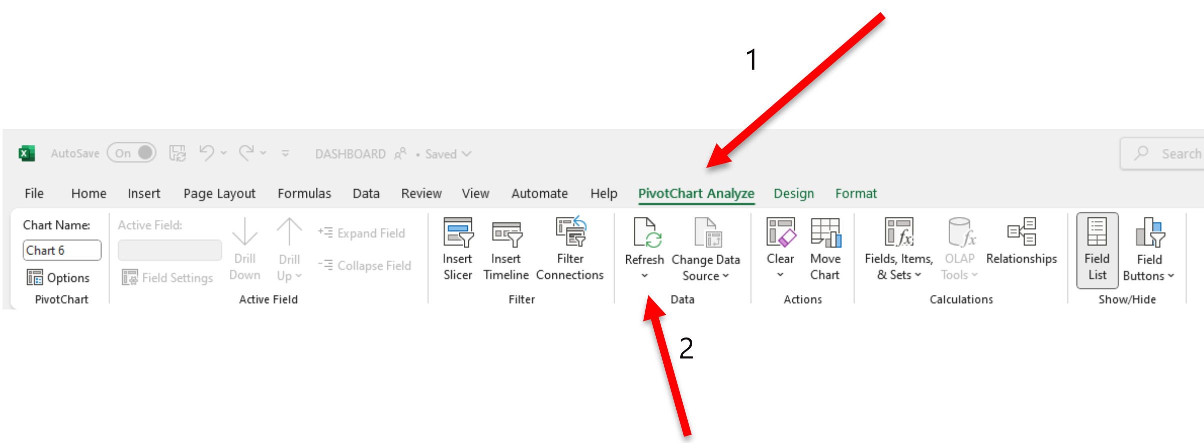
Once the scenarios are selected, the results can be updated from VEDA:

1. In VEDA, go to the "Results" module.
2. Select the scenarios you want to analyze – those that were chosen earlier.
3. Click on Update Excel and then All Tables.
4. Choose the *Update_excel.xlsx* file you set earlier. This will automatically fill the file with the data necessary to populate the dashboard tables and charts. If the file already contains data, it will be overwritten.



Once the update is complete, the *Update_excel.xlsx* and the *DASHBOARD.xlsx* files need to be synchronized to display the correct scenarios and data:

1. Open first the *Update_excel.xlsx* file and then the *DASHBOARD.xlsx* file. This will automatically copy the data from the *Update_excel.xlsx* file to the designated worksheets in the dashboard. The scenarios will also be copied in the *Scenario* sheet of the dashboard.
2. Go into any of the *Charts* sheets and make a selection. Then navigate to "PivotChart Analyze," click "Refresh," and select "Refresh all". This will ensure that the latest data are displayed in each of the charts.



2.4.2. Dashboard structure

The dashboard is built using four types of sheets:

- *Info* sheets (blue): these sheets group relevant information.
- *Charts* sheets (green): These sheets present the charts created in this dashboard.
- *Tables* sheets (orange): These sheets contain the tables used to create the charts presented throughout the document.
- *Translated data* and *RAW data* sheets (grey): These sheets contain raw data extracted from the *Update_excel.xlsx* file.

2.4.2.1. The info tabs (blue)

DOC. This sheet serves as reminder of the links between the *DASHBOARD* file and the *Update_excel.xlsx* file. It also contains links to the different sections of the dashboard.

SCEN. This sheet is a copy from the *Scen* tab in *Update_excel.xlsx*, listing the scenario names in VEDA alongside the names to be printed in the charts. If these names need to be changed, they should be modified in the *Update_excel.xlsx* file. If both files are open simultaneously, the update should be made immediately in the dashboard, while maintaining the link between the files for future use.

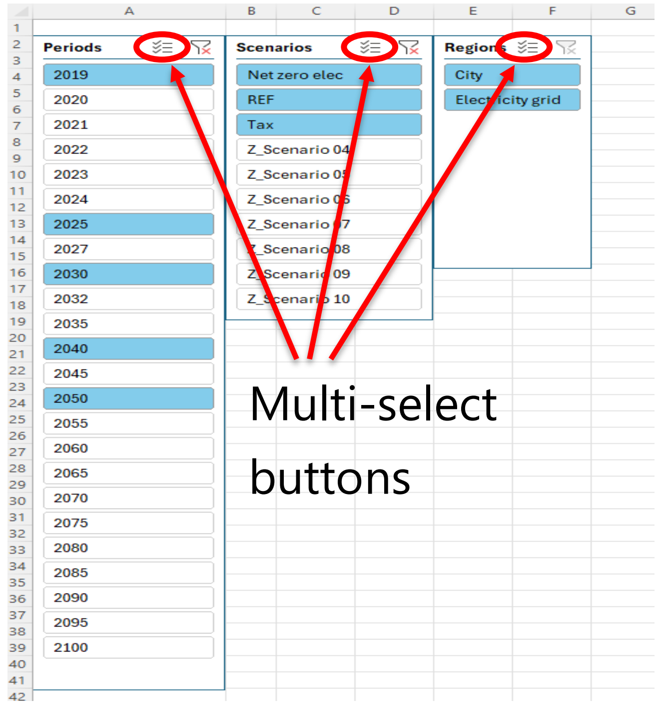
Dict. This sheet functions as a dictionary to translate the VEDA terms into more understandable descriptions. All the regions, commodity sets, and process sets used to build the charts in the dashboard are included in this sheet. For each of them, the first column (with grey cells) contains the VEDA terms, which should not be modified. The second column (with green cells) includes the translated terms. This column can be modified to reflect the preferred terminology in the charts.

2.4.2.2. The *Charts* sheets (green)

There are nine sheets containing relevant charts:

- *Emissions*: This sheet displays total GHG emissions over the entire timeline, as well as GHG emissions by sector and pollutant. Here, emissions from electricity production of the entire eGRID subregion are displayed.
- *EMS_Elec_correction*: This sheet presents total GHG emissions over the entire timeline, as well as GHG emissions by sector and pollutant. The emissions from electricity production have been adjusted to better reflect the city's actual emissions. The ratio of electricity production dedicated to the city has been calculated and this ratio has been applied to the emissions from electricity production.
- *NRG*: This sheet shows final energy consumption by sector and fuel type.
- *NRG_Sect*: This sheet details the final energy consumption in each sector, broken down by fuel type.
- *Electricity*: This sheet illustrates electricity production by energy source, considering both the entire eGRID and the ratio of electricity consumed by the city.
- *Buildings*: This sheet presents thermal energy production for heating in buildings, covering both water and space heating for commercial and residential buildings.
- *COM_Buildings*: This sheet focuses on thermal energy production for space and water heating in commercial buildings, categorized by technology.
- *RES_Buildings*: This sheet highlights thermal energy production for space and water heating in residential buildings, categorized by technology.
- *Transport*: This sheet shows total road transport data, categorized by transportation type, as well as transportation technology for public transport, personal vehicles, and freight transport.

All sheets contain three slicers on the left side, representing the years defined in the model, the scenarios analyzed, and the different regions considered. These slicers control what is displayed on all charts in the dashboard. When the multi-select button is activated, it allows for the selection of multiple values; otherwise, only the selected value will be displayed.



Therefore, if one wants to display a single scenario, the multi-select button should be deactivated so that only one scenario can be chosen.

The slicers are all synchronized, meaning that whenever a slicer is modified, all the slicers will update accordingly, and the charts across all tabs will adjust their display.

The only two charts not fully connected to these slicers are one displaying total of GHG emissions over time. These two charts are linked solely to the scenario slicer, which means that the years represented will always be the same, and they will account for both the city and eGRID-determined emissions.

It is also possible to select the unit in which data will be displayed in the first row of the *Charts* sheets:

- *Emissions* sheet: You can choose the emissions unit among t.CO_2e, kt.CO_2e and Mt.CO_2e in the C1 cell of the tab, in the drop-down menu. This choice impacts the three charts on this sheet as well as the three charts in the *EMS_Elec_correction* sheet.
    - The pollutant chart is synchronized with the unit choice even if this graph displays data in t, kt, or Mt of pollutants.
- *Energy* sheets: For the six tabs that display energy charts (*NRG*, *NRG_Sect*, *Electricity*, *Buildings*, *COM_Buildings,* and *RES_Buildings*), you can choose among TJ, PJ, GWh, and TWh.
    - Caution: the energy unit can only be modified in the NRG tab, and this will affect all charts. In this tab, cell C1 is a drop-down menu for selecting the unit. The other tabs indicate the chosen unit as a reminder, but it is not possible to modify the unit from any of these tabs.
- *Transport* sheet: You can choose the overall distance unit between millions of vehicle kilometers (M v.km) and millions of vehicle miles (M v.miles)

For the unit change to take effect, it is necessary to refresh the pivot charts, as done when synchronizing the dashboard file with the *Update_excel.xlsx* file.

2.4.2.3. The *Tables* charts (orange)

There is one table tab for each chart represented in the *Charts* sheets, totaling 23 sheets. In each of these sheets, there is a pivot table containing the data used to create the corresponding chart. This pivot table is influenced by the slicers and unit selections made in the *Charts* sheets. These sheets provide better access to numerical data; however, the tables should not be modified (outside using the slicers), as doing so would alter the charts displayed in the *Charts* sheets.

2.4.2.4. The *Translated Data* and *RAW Data* sheets (grey)

These sheets are used to import the data from the *Update_excel.xlsx* file, translate the VEDA nomenclature using the *Dict* sheet, and, in some cases, process the data for better understanding. These tabs are not to be modified.

2.5. How to understand model results

COMET, as an optimization model, provides the least-cost system solution under a given set of constraints. The solutions are not necessarily the most "likely" outcomes as they represent perfect foresight into the future; rather, they represent "ideal" solutions based on a least-cost pathway.

As such, results from this model should not be interpreted as a forecast. Instead, they can be interpreted as projections of what may happen in each scenario, under certain conditions. With the use of market share constraints, there is (limited) consideration of market barriers or economically irrational decisions. The modeling results represent the least-cost solution at a societal level, not the behavior of a specific economic agent.

3. Limitations and perspectives

3.1. Difference from COMET NYC

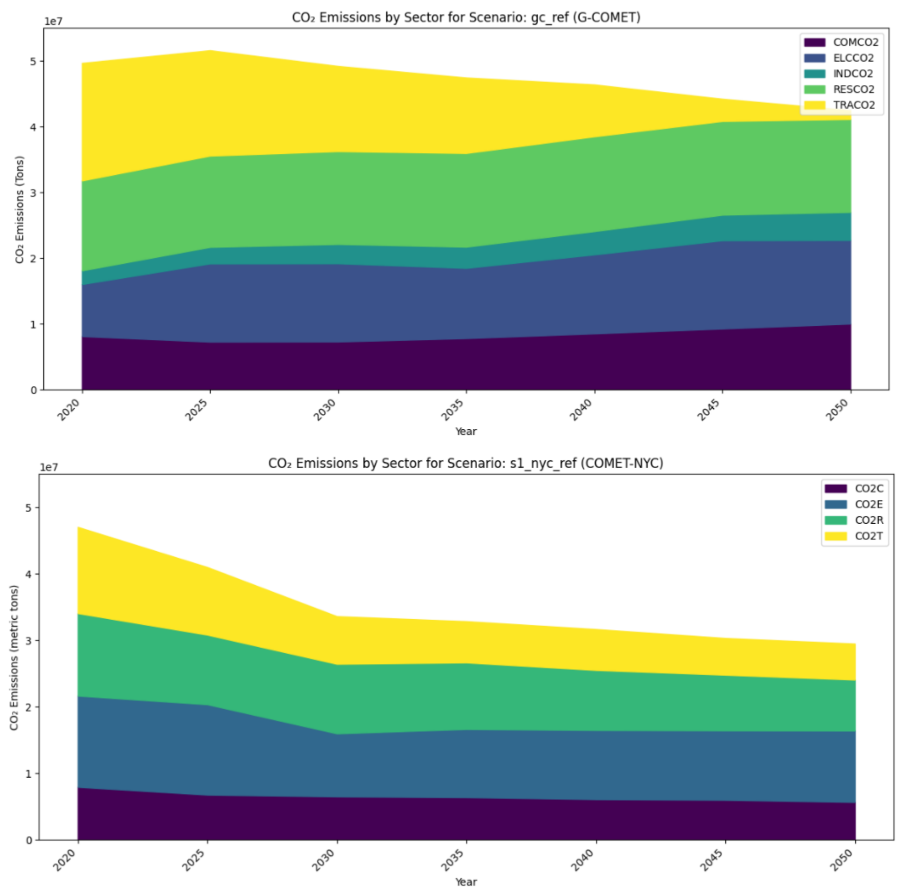
There are several modeling differences between the COMET_NYC model and the Generative COMET model applied to New York City that account for the differences in their results. While the COMET_NYC model is specifically tailored for New York City, using detailed local data, assumptions, and parameters, the Generative COMET model is designed as a versatile framework that can be applied to any city. It offers a flexible approach for analyzing energy and GHG emission scenarios, with innovative capabilities that allow users to calibrate energy usage and GHG emissions based on the level of data available. Consequently, the Generative COMET model employs a more generalized set of assumptions and default parameters in some areas to accommodate different levels of data granularity and availability. The Generative COMET model can be further customized with detailed local data for New York City, which would help align its results more closely with those of the COMET_NYC model.

The major differences are:

- **Electricity and natural gas prices:** The ratio between electricity and natural gas prices in the residential sector is significantly higher in the COMET_Generative model than in the COMET_NYC model. As a result, one may observe a much smaller increase in electricity consumption in the residential sector in a reference scenario of the COMET_Generative model.
- **District heating availability:** There are several differences across both residential and commercial sectors between COMET models. These differences occur due to the specified technology options in each model. Many of these differences occur in the space heating and cooling technologies. For example, district heating is not available in the Generative COMET model. Consequently, the adoption of district heating options is not specified for the commercial sector, unlike in the COMET_NYC model. Additionally, the Generative COMET model expands the types of specific appliances and technologies that the build sector specifically considers (e.g. refrigeration, freezing, clothes washing and drying, etc.).
- **Electricity production modeling:** Electricity production is modeled quite differently in the two models. In COMET_NYC, it is based on all electricity generation in New York state, along with interconnection to neighboring jurisdictions. In contrast, the COMET_Generative model relies on the eGRID2019 sub-region the city is part of, and builds out existing energy generating units (EGUs) from the region if current production importsis insufficient. Additionally, a fixed amount of imported energy is available for when existing production is insufficient or new EGUs cannot be reasonably expanded. For New York City, the eGRID sub-region primarily represents only the city itself. Due to the limited

scope of the New York City eGRID sub-region, a substantial portion of the city's consumption is sourced from outside this area, which is modeled with a single import process with a constant emissions factor. Additionally, production expansion is very restricted because the capacity potential for various types of electricity generation within New York City limits is quite limited. Specifically, the model prioritises expanding existing electricity generating technologies.

Figure 16. Comparison of CO2 emissions between the two models



3.2. Possible future addition of technology options

Some technology options are missing in this version of the model and could be added in future updates. Amongst others, there are:

- **Residential consumption data:** The RECS [4] does not contain steam, wood biomass, or kerosene consumption. Since this source is used to disaggregate residential consumption into end uses, the current model version does not include any existing steam, wood biomass, or kerosene-based technologies in the residential sector. Generic assumptions or hypotheses from other sources could be added in an updated version to address this gap.

- **Commercial consumption data:** The CBECS [2] does not include information on liquified petroleum gas (LPG), gasoline, diesel, kerosene, waste, or wood biomass consumption. As this source is used to disaggregate commercial consumption into end uses, the model currently lacks existing technologies based on LPG, gasoline, diesel, kerosene, waste, or wood biomass in the commercial sector. Generic assumptions or hypotheses from other sources could be used to fill this gap in an updated version.
- **Space cooling options:** Currently, only electricity-based options are implemented for space cooling in both the existing and new technology database. Natural gas, steam, and fuel oil-based options could be added in a future update, as CBECS [2] considers these alternatives for some states.
- **Charger efficiency:** Charger efficiency is not accounted for when calibrating electricity consumption for transportation. This oversight may lead to small discrepancies in the base year between the calibration sheets and the model results. These calibration differences could be easily addressed in a future update.
- **Biofuel consumption:** For the base year, biofuel consumption is only accounted for in the transportation sector, as CBECS [2] and RECS [4] do not include biofuel consumption and are used to disaggregate building consumption into end uses. Generic assumptions or hypotheses from other sources could fill this gap in a future update.
- **NEMS database limitations:** The new technology database for commercial and residential sectors is based on NEMS databases [8]. However, some relevant technology options, such as natural gas cooking in the commercial sector or LPG-based technologies, are not included in this source but may be important to add.
- **Industrial sector technologies:** The new technology database does not include steam-based technology for the industrial sector. This could be easily added when incorporating a steam network.
- **District heating options:** As mentioned in the previous section regarding differences with COMET_NY, the new technology database lacks district heating options, as well as all end-use technologies utilizing steam in buildings. These options are highly relevant for cities and could be added in a future update of the model.
- **Electricity production calibration:** As noted earlier, using eGRID sub-regions to calibrate electricity production for the city can impact expansion and emissions accounting. A better representation of the electricity production mix consumed by the city can be achieved by

using eGRID data at the state level. This change can be incorporated into an updated version of the model that enables users to conduct state-level analysis.

- **Hydrogen production technologies:** Hydrogen, like electricity, has a GHG impact that occurs not at the combustion point but at the production stage. Including hydrogen production technologies in the model may be beneficial, as it would allow for a more comprehensive assessment of emissions. This approach would mirror how electricity production is currently represented in the model. Furthermore, considering hydrogen emissions for hydrogen consumed within city limits could align with the methodology used for accounting for electricity emissions.
- **Data center energy consumption:** Given the projected growth in data center energy use in the US, due to applications such as Generative AI, future versions of the model may need to incorporate energy consumption from the data centers located inside the city with a new end-use demand.

References

[1] BTS - Bureau of Transportation Statistics, "Freight Analysis Framework Version 4 (FAF4)," 2024. [Online]. Available: https://www.bts.gov/faf/faf4.

[2] EIA - U.S. Energy Information Administration, "Commercial Buildings Energy Consumption Survey (CBECS)," 2018. [Online]. Available: https://www.eia.gov/consumption/commercial/data/2018/index.php?view=microdata.

[3] USEPA - United States Environmental Protection Agency, "MOtor Vehicle Emissions Simulator (MOVES) Version 3," 2022. [Online]. Available: Provided by EPA.

[4] EIA - U.S. Energy Information Administration, "Residential Energy Consumption Survey (RECS)," 2020. [Online]. Available: https://www.eia.gov/consumption/residential/data/2020/index.php?view=microdata.

[5] EIA - U.S. Energy Information Administration, "State Energy Data System (SEDS): 1960-2022 (complete)," 2024. [Online]. Available: https://www.eia.gov/state/seds/seds-data-complete.php?sid=US#CompleteDataFile.

[6] Google, "Environment Insights Explorer," 2024. [Online]. Available: https://insights.sustainability.google/.

[7] USEPA - United States Environmental Protection Agency, "EPAUS9rT database for use with the TIMES modeling platform," 2019. [Online]. Available: https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NRMRL&dirEntryId=346478.

[8] EIA - U.S. Energy Information Administration, "Annual Energy Outlook 2023, National Energy Modeling System (NEMS)," 2023. [Online]. Available: https://www.eia.gov/outlooks/aeo/info_nems_archive.php.

[9] NREL - The National Renewable Energy Laboratory, "2022 Transportation annual Technology Baseline," 2022. [Online]. Available: https://atb.nrel.gov/transportation/2022/index.

[10] World Bank Group, "The CURB Tool: Climate Action for Urban Sustainability," 2016. [Online]. Available: https://www.worldbank.org/en/topic/urbandevelopment/brief/the-curb-tool-climate-action-for-urban-sustainability.

[11] USEPA - United States Environmental Protection Agency, "Emissions & Generation Resource Integrated Database (eGRID)," 2019. [Online]. Available: https://www.epa.gov/egrid.

[12] NREL - The National Renewable Energy Laboratory, "2024 Electricity ATB Technologies," 2024. [Online]. Available: https://atb.nrel.gov/electricity/2024/technologies.

[13] NREL - The National Renewable Energy Laboratory, "Regional Energy Deployment System, ReEDS," 2024. [Online]. Available: https://www.nrel.gov/analysis/reeds/.

[14] EIA - U.S. Energy Information Administration, "Annual Energy Outlook 2023, Projections Tables for Side Cases," 2023. [Online]. Available: https://www.eia.gov/outlooks/aeo/tables_side_xls.php.

[15] USEPA - United States Environment Protection Agency, "Solid Waste Optimization Life-cycle Framework," 2021. [Online]. Available: Provided by EPA.

[16] USEPA - United States Environmental Protection Agency, "GHG Emission Factors Hub," 2022. [Online]. Available: https://www.epa.gov/climateleadership/ghg-emission-factors-hub.

[17] IPCC, "Fourth Assessment Report," 2007.

[18] D. Durbin, "Durbin Introduces Bill to Invest in a Clean Climate Future and Spur Job Creation," 2021. [Online]. Available: https://www.durbin.senate.gov/newsroom/press-releases/durbin-introduces-bill-to-invest-in-a-clean-climate-future-and-spur-job-creation.

[19] National Centers for Environmental Information, "U.S. Climate Normals Quick Access," 2021. [Online]. Available: https://www.ncei.noaa.gov/access/us-climate-normals/#dataset=normals-hourly&timeframe=15.

[20] NYC - Mayor's Office of Climate & Environmental Justice, "NYC Greenhouse Gas Inventories," 2024. [Online]. Available: https://climate.cityofnewyork.us/initiatives/nyc-greenhouse-gas-inventories/.

[21] Chicago Metropolitan Agency for Planning, "Environmental Data, Greenhouse Gas Emissions," 2024. [Online]. Available: https://datahub.cmap.illinois.gov/search?categories=%252Fcategories%252Fenvironment&q=greenhouse%20gas.

[22] United Nations Climate Change, "Global Warming Potentials (IPCC Fourth Assessment Report)," [Online]. Available: https://unfccc.int/process-and-meetings/transparency-and-reporting/greenhouse-gas-data/frequently-asked-questions/global-warming-potentials-ipcc-fourth-assessment-report. [Accessed 2024].

United States
Environmental Protection
Agency

PRESORTED STANDARD
POSTAGE & FEES PAID
EPA
PERMIT NO. G-35

Office of Research and Development (8101R)
Washington, DC 20460

Official Business
Penalty for Private Use
$300

Recycled/Recyclable Printed on paper that contains a minimum of
50% postconsumer fiber content processed chlorine free