Complete specification on
Climate and Energy Assessment of Museums (CEAM) software

v. 1.2.1 published on 01.07.2025

Marcin Zygmunt, PhD

KU Leuven, Department of Civil Engineering, Building Physics and Sustainable Design,
Kasteelpark Arenberg 40, BE-3001 Leuven, Belgium

This document was delivered as an objective of the Climate2Preserv research project, founded by the Belgian Science Policy Office (BELSPO).

2025

TABLE OF CONTENTS

1. About the CEAM

You can only start to save energy when you see how you are using it. It is essential for buildings housing cultural heritages, where standard approaches used to assess buildings' energy performance cannot be applied. These types of buildings are also uniquely managed, focusing mostly on heritage preservation. Energy-efficiency improvement is possible only when understanding the examined building. The selection of the actions considered for potential improvements is highly dependent on the examined case study and should be executed only after the approval of the delegated building managers. The knowledge of the examined building and its systems is a key factor for the appropriate building assessment, no matter the quality and comprehensiveness of the used software.

The **C**limate and **E**nergy **A**ssessment for **M**useums (**CEAM**) was developed to allow its users a valuable and constructive energy- and climate-related overview of the examined heritage buildings. CEAM aims to help Federal Institutions preserve heritage. The developed decision-making support software provides information to building/collection owners on how to adapt sustainable solutions to their buildings, utilizing indoor climate management. CEAM is an open-source, data-driven tool scripted using Python and consists of a modular structure. This user-friendly tool allows for a comprehensive multicriteria analysis considering heritage preservation, indoor climate management, and energy efficiency improvement. The tool uses raw input data (records of the measured energy consumption and indoor/outdoor climate conditions) to provide useful outputs. The obtained outputs allow for a better understanding of the current building management and provide a helpful guide on how to, and to what extent, improve the energy efficiency of the examined case. Within the generated outputs, the following main aspects can be addressed:

- a comprehensive overview of input data,
- outline of the current building management,
- overview of potential energy savings.

The software is available online[1] as a stand-alone Windows app[2], as well as raw scripts[3]. In this document, a complex overview of the CEAM software is presented. User guidelines, methodology overview, as well as examples are addressed in the following sections.

CEAM was developed by Marcin Zygmunt[4] as a part of the Climate2Preserv (C2P) project[5], founded by the Belgian Science Policy Office[6] (BELSPO). The C2P project is the combined effort of project partners: KU Leuven[7], KIK-IRPA[8], and ULiège[9].

Please address all the comments related to the CEAM software via the [contact email](mailto:climate2preserv@gmail.com): **climate2preserv@gmail.com**.

[1] available here: https://drive.google.com/drive/folders/1wIT0tG_Lo9iJ8fh8vHU-SB-zgOoksMFa?usp=sharing
[2] developed on Windows 11; no tests run on previous versions
[3] available here: https://github.com/Climate2Preserv
[4] see more: https://www.linkedin.com/in/marcin-zygmunt-06596115a/
[5] see more: https://www.kikirpa.be/en/projects/climate2preserv
[6] see more: https://www.belspo.be/
[7] see more: https://www.kuleuven.be/kuleuven/
[8] see more: https://www.kikirpa.be/en/
[9] see more: https://www.uliege.be/cms/c_8699436/en/uliege

2. INSTALLATION

This section goes through the installation process. Based on the needs, CEAM can be used as a standalone program or as a set of scripts (the so-called developer version), allowing for individual and unique modifications.

2.1. STAND-ALONE VERSION

By default, CEAM is a standalone software, working on a Windows operating system. After downloading[1], the tool is ready to use by the user. There is no need to install any additional software if the updated version of Windows is used. It is recommended to create a new folder and put all the downloaded file(s) there. It is also suggested to run the software via '*run as administrator*' option.

The **CEAM.exe** file should be used to run the tool. CEAM operates using the Graphical User Interface (GUI) presented in the next section (see more in section **4**). Additionally, specific input files are required for a successful application of CEAM software (see more in section **3**). Depending on the given scope of analysis and system performance, computing time can take from a few minutes up to a few hours. Users are informed about the computation in progress with a visible pop-up window and an updated simulation status. While computing, the numerous supporting files are generated in the given working directory; all the files should remain intact throughout calculations (file management is performed by CEAM). Users are informed about the computing time, as well as the execution time, when calculations are finished with the status textbox, as well as the status pop-up window, required to accept when the analysis is over (followed by the system sound).

2.2. DEVELOPER VERSION

CEAM is an open-source tool, written in Python[10]: a high-level, multi-purpose, and easy-to-read programming language. Python has a straightforward structure and syntax with quite simple to maintain structural code. This programming language provides very high-level dynamic data types and supports dynamic type checking with both object-oriented and procedural programming styles. Finally, it has a sizeable standard library collection, applied for data collection, data exploration, and data visualization, among others.

All the CEAM scripts are available[3] for further individual use and modification by the users. Keep in mind that it is the user's responsibility to verify the results obtained from the modified version of the CEAM software or any of its scripts. All the suggestions for further improvements of the tool are highly appreciated and can be provided via the **contact email**.

In this part, a brief description of how to use CEAM in the so-called developer version is shown. It is assumed that those who decide to use this version of the tool have a basic knowledge of programming, as well as are familiar with the Python environment; for more information, visit the GitHub repository[3].

First of all, you should have installed Python (version no older than 3.12) on your computer. CEAM was developed in the PyCharm[11] Integrated Development Environment (IDE), but users

[10] see more: https://www.python.org/
[11] available here: https://www.jetbrains.com/pycharm/

can use any IDE based on their preferences (with potential adjustments required). Downloaded scripts should be placed as a new Python project, preferably in the working directory. For a complete application of the developer version, the following Python libraries are necessary to be installed:

- *csv*[12]: provides tools for reading and writing csv files,
- *customtkinter*[13]: an extension of *tkinter* library,
- *datetime*[12]: functionalities for manipulating dates and times,
- *gc*[12]: module provides optional garbage collector,
- *glob*[12]: library enables file pattern matching and pathname expansion,
- *holidays*[14]: handles holidays for predefined countries and regions,
- *importlib*[12]: allowing to impost functionalities from different scripts,
- *keras*[15]: a high-level neural networks Application Programming Interface (API) used for building and training machine learning models,
- *matplotlib*[16]: library for visualizations, graphs in particular,
- *numpy*[17]: support for large, multi-dimensional arrays and numerical computations,
- *os*[12]: library offering functions to interact with the operating system,
- *pandas*[18]: library for various data manipulation and analysis,
- *platform*[12]: access to the underlying platform's identifying data,
- *psychrolib*[19]: allowing to generate and modify psychrometric charts,
- *re*[12]: library facilitates operations involving regular expressions for pattern matching,
- *scipy*[20]: provides scientific and technical computing tools,
- *shutil*[12]: offers many high-level operations on files and collections of files,
- *shuttle*[21]: to handle file operations (e.g. copying and moving),
- *sklearn*[22]: a machine learning library for data mining and analysis,
- *subprocess*[12]: enables running and managing additional processes,
- *sys*[12]: library provides access to system-specific parameters and functions,
- *tensorflow*[23]: a platform for machine learning and deep learning,
- *threading*[12]: library to run the subprocess (parallel tasks) in a separate thread,
- *time*[12]: functions for working with time-related tasks,
- *tkinter*[12]: library for building GUIs,
- *webbrowser*[12]: used to open Uniform Resource Locator (URL) in a web browser,
- *winsound*[12]: library for playing Windows-specific sound effects.

CEAM (in the current 1.2.1 version) consists of 6 modules, contains 26 scripts, which are executed in a specific order regarding the requested scope. 1 additional script (named *_00_time_index_generator.py*) is included as part of the CEAM software, even though it is

[12] see more: https://docs.python.org/3/
[13] see more: https://pypi.org/project/customtkinter/0.3/
[14] see more: https://pypi.org/project/holidays/
[15] see more: https://pypi.org/project/keras/
[16] see more: https://pypi.org/project/matplotlib/
[17] see more: https://pypi.org/project/numpy/
[18] see more: https://pypi.org/project/pandas/
[19] see more: https://github.com/psychrometrics/psychrolib
[20] see more: https://pypi.org/project/scipy/
[21] see more: https://pypi.org/project/shuttle/
[22] see more: https://pypi.org/project/scikit-learn/
[23] see more: https://pypi.org/project/tensorflow/

considered a small individual sub-software, not being a part of any modules, as well as not used directly for computing. The initial operation of the tool is performed in the so-called mother script named *_00_CEAM_run.py*. Users can run the simulations from this script alone, selecting the preferred scope of analysis (by adjusting the list of executed modules or scripts). When a stand-alone version is used, users will operate on the GUI defined in the *CEAM_GUI.py* file. Multiple predefined links between different modules are present; the correct operation of the CEAM software is possible only if that order is maintained. The user-dependent adjustments can be made in each of the provided scripts; a compact description of each script is shown in **Table 1**. Each script is presented more in detail in section 5, distinguished by each module. Scripts are presented in the recommended application order, following the modular structure of the tool. Moreover, all the scripts are rich in comments to help users navigate the codes.

Table 1. Developer version of the CEAM software: scripts overview.

#	MODULE	NAME	DESCRIPTION
1	*	*CEAM_GUI.py*	This script generates the software GUI with all the functionalities.
2	**	*_00_CEAM_run.py*	The so-called *mother script* predefined the scope of analysis employing the computed scripts.
3		*_11_initial_csv_check.py*	This script initially checks the provided input files and performs folder and file management.
4		*_01_convert_units.py*	Script used to convert IU units to SI units.
5		*_02_pv_calc.py*	Script used to calculate vapor pressure.
6		*_03_degree_days.py*	Script used to calculate heating and cooling degree days[24].
7		*_12_data_overview_stats.py*	This script verifies the potential incongruities of the provided inputs, as well as generates daily, monthly, and weekly averaged files.
8	0	*_04_extreme_weeks.py*	Script is used to define the so-called extreme weeks and save them for later assessments.
9		*_05_1_combi_files.py*	Script combining the input files into the so-called *combi_files*, with all input variables in one file, based on the frequencies of records.
10		*_05_2_adjust_combi_files.py*	Script adjusting the pre-generated *combi_file* with the desired date and time formats.
11		*_06_1_Td_Twb_calc.py*	Script used to calculate the dew point[25] and wet bulb temperature[26].
12		*_06_2_factors.py*	Script used to calculate other factors for climate and energy assessments.
13		*_06_3_abs_humi.py*	Script used to provide outputs of absolute humidity[27].

[24] see more: https://www.degreedays.net/
[25] see more: https://en.wikipedia.org/wiki/Dew_point
[26] see more: https://en.wikipedia.org/wiki/Wet-bulb_temperature
[27] see more: https://en.wikipedia.org/wiki/Humidity#Absolute_humidity

14	1	*_13_data_overview_plots.py*	This script provides a graphical overview of the input data.
15	2	*_14_correlation.py*	This script examines correlations between the provided input variables.
16	3	*_15_energy_signature.py*	This script performs an energy signature assessment.
17	4	*_07_CC_file.py*	Script used to generate the input file for the Climate Class assessment.
18		*_08_psychrometric.py*	Script used to generate Climate Class assessment by means of psychrometric charts.
19		*_16_CC_overview.py*	This script performs a Climate Class assessment.
20	5	*_09_prediction_inputs.py*	Script used to generate the input files required for the assessment of potential energy savings.
21		*_17_predictions.py*	This script provides the energy predictions by means of black box modeling[28].
22		*_18_1_best_predictioons.py*	This script allows for an overview of the performed energy-saving assessment and to select the best approximators.
23		*_18_2_best_predictions_graphs.py*	This script provides a graphical overview of the selected, optimal models for energy optimization.
24		*_18_3_best_predictions_sorting.py*	This script provides the sorting of the files obtained during the optimization process and the selection of the best approximators.
25		*_19_cc_reeval.py*	This script allows for Climate Class reassessment (in a similar manner as in module 4) for the outputs obtained from the energy optimization process.
26	***	*_00_time_index_generator.py*	Script generates a selected data frame using the provided data format for the selected period. The output is provided as a csv file with a *Time* column.

* script providing a GUI for the CEAM software
** script defining the scope of analysis performed using the CEAM software
*** supporting script, which is not a part of the modular structure of the CEAM software

The developer version, thus all the Python scripts, as well as exemplary input files and user manuals, are available at the GitHub repository[3].

[28] Black-Box Modeling is one of the approaches applicable for simulations and modeling. An overview of the available modelling techniques and their comparison can be found in: https://doi.org/10.3390/buildings12081284. See more in subsection 5.6 for the techniques used in CEAM software.

3. INPUT DATA

CEAM is a data-driven analytical software. Based on the provided data, a comprehensive overview of indoor climate control, as well as corresponding energy consumption, can be performed. Furthermore, the assessment of the indoor climate can be obtained, as well as optimization of energy consumption is possible. In this section, a detailed overview of the necessary structure and content of the input data is discussed.

3.1. INPUT FILES

CEAM software works based on the provided input files; only 3 comma-separated values (csv) input files are required for successful computing. Yet, a certain input data structure is required. A detailed overview of the structure of the required input files, as well as some guidelines on how to prepare correct entries, are shown in the subsection 3.2. Records in the input files can be in either international[29] (SI) or imperial[30] (IU) system of units. A precise naming for the used input files is required, as follows:

- ***ECD.csv***: Exterior Climate Data; a file containing information about the weather conditions measured preferably on-site or at the closest distance,
- ***ICD.csv***: Indoor Climate Data; a file containing information about the inner conditions measured for the selected zone(s),
- ***ED.csv***: Energy Data; a file containing historical records of the consumption used for specific needs (e.g., heating), distinguished for the preselected zone(s).

Exterior climate should be measured preferably on-site, or as close to the examined case study as possible. It is possible to use a weather file from a nearby meteorological station if measurements on-site are not possible/available. It is also possible to use (after adjustments) the Typical Meteorological Year[31] (TMY) files, yet it is not recommended. Finally, users should keep in mind that their weather data needs to be adjusted to follow the ECD file template. The required variables are exterior temperature (T_e) and exterior relative humidity (RH_e), where total solar radiation (I_s) is recommended, but not obligatory for computing.

Indoor climate should be measured for the selected room/zone. The measurements are done using specialized sensors. It is up to the user to decide how to address the examined case study, in particular, what assumptions to use during the preparation of the input files. The records used in input files should follow the assumptions made. The recommended approach is to select the most representative room/zone of the building and adjust the indoor climate measurement accordingly (i.e., use data from sensors located in the selected area). Data should be averaged if multiple records are available for the selected room/zone, or the records from the most representative sensor (e.g., located in the middle of the area) should be used. Finally, users should keep in mind that their indoor climate data needs to be adjusted to follow the ICD file template: the required variables are interior temperature (T_i) and interior relative humidity (RH_i).

[29] see more: https://en.wikipedia.org/wiki/International_System_of_Units
[30] see more: https://en.wikipedia.org/wiki/Imperial_units
[31] A collation of historical weather data derived from a multi-year time series selected to present the unique weather phenomena with annual averages that are consistent with long-term averages. Some TMY files can be found and downloaded from here: https://energyplus.net/weather

Energy consumption can be measured via various approaches and for various levels of detail for the indoor climate management system, typically called HVAC (heating, ventilation, and air conditioning). Moreover, the consumption can be measured at different stages of usage: building-, system-, unit-, zone-, or room-level. The available measured energy consumption should be adjusted in accordance with the examined room/zone. Modern HVAC systems are operated via a building management system (BMS), allowing to access more valuable and distinguished information regarding energy consumption. All the above-mentioned information should be considered while preparing the ED file. It is recommended to work with a system specialist to make a correct input file. Finally, users should keep in mind that their energy consumption data needs to be adjusted to follow the ED file template. The provided ED file can consist of any combination of the following consumptions: heating (*D1*), cooling (*D2*), ventilation (*D3*), humidification (*D4*), and total (*D5*).

The collected and used input data is crucial for a successful CEAM application. The granularity of the data (its frequency) affects the available assessment and its accuracy. Instinctively, the more detailed the data, the better and more comprehensive the examination; yet in practice, it is more compound and case-dependent. Detailed interval data (hourly frequencies for CEAM software) allows unveiling of patterns/profiles of energy demand, while daily or monthly inputs are accurate only for simpler methods and data overviews, based on the performed assumptions.

Some exemplary sets of input files are shared with the software[1,3]. These files can be considered as templates to prepare inputs for future use of CEAM. The description of each set, their application, possible scope of analyses, as well as the obtained outputs, are reviewed in the subsections 7.2-7.5.

3.2. INPUT CONTENT AND HOW TO ADJUST IT

For every csv file, we can distinguish headers (present in the first line of the file), separator (symbol separating each record/variable), as well as list(s) of records (measured data) with corresponding time format. A separator should be the same throughout the file, avoiding confusion with the used decimal point; a typical separator is a comma or semicolon[32]. The headers used should clearly describe the collected records in the given column. CSV files can be opened by MS Excel, but it is recommended to open them as text files, via e.g., Notepad[33].

CSV input files required by CEAM software have a predefined, simple structure that needs to be respected. All the provided input files (the so-called *raw files*) will be safely copied and saved before any actions are performed by the CEAM software. All the headers should be in the very first row of the file with proper spelling (see more in **Table 2**), yet the tool handles various fonts and placements. Header verifications will be performed; CEAM will detect any incompatibility (font, placement) and reorganize the header to follow the predefined template (sub- and super-scripts are skipped in csv files). Each file must contain the *Time* column, with the same known structure of the formatting[34] throughout each file. It is essential to know the positioning of the basic components of the date, day, month, and year, in particular. The used

[32] the naming of the csv (comma-separated value) file is nonspecific because the used separator can be selected by the user, software used, or the operating system

[33] available here: https://notepad-plus-plus.org/

[34] it is impossible to understand correctly the dates without knowing the context of its formatting

separator of the components, as well as the style for time formatting, will be automatically evaluated by CEAM. The *Time* format will be adjusted before computing: the default format (*dd.mm.yyyy HH:MM*) will be used with a dot as a separator for all input files. Finally, the data frame frequency (the time between each record) should remain the same throughout the file. The software will detect the most common frequency used and adjust the input files accordingly. CEAM supports hourly, daily, or monthly frequencies. The used unit system (SI or IU) should be the same for all the input files. The software will automatically check the used separators in the provided csv files, and then modify them if necessary (the comma is used as default). The user needs to verify that the separator and decimal point do not match. It should also be checked if the provided records are aligned with the corresponding headers. The zeroth values should be input as 0, while no record should be left as empty evidence.

Table 2. Information on the headers of the input files for computing using CEAM.

INPUT FILE	HEADER	DESCRIPTION
ECD.csv	*Time*	Indexing for the given data frame, using the predefined format; the default format is: *dd.mm.yyyy HH:MM*
	T_e	Exterior air temperature in [C] (SI) or [F] (IU)
	RH_e	Exterior relative humidity in [%]
	I_s	Total solar radiation in [Wh/m^2]
ICD.csv	*Time*	Indexing for the given data frame, using the predefined format; the default format is: *dd.mm.yyyy HH:MM*
	T_i	Interior air temperature in [C] (SI) or [F] (IU)
	RH_i	Interior relative humidity in [%]
ED.csv	*Time*	Indexing for the given data frame, using the predefined format; the default format is: *dd.mm.yyyy HH:MM*
	D1	Demand group no.1: heating, in [Wh] (SI) or [BTU] (IU)
	D2	Demand group no.2: cooling, in [Wh] (SI) or [BTU] (IU)
	D3	Demand group no.3: ventilation, in [Wh] (SI) or [BTU] (IU)
	D4	Demand group no.4: humidification, in [Wh] (SI) or [BTU] (IU)
	D5	Demand group no.5: total consumption, in [Wh] (SI) or [BTU] (IU)

Considering all the above-mentioned, exemplary file adjustments of the input file that CEAM software will perform before computing are shown in examples #1 and #2 (subsection 7.6).

The available measurement data can be provided to CEAM in various ways. The most universal approach to generate the necessary input files is to create the csv file using MS Excel. User should place the required variables in the obligatory order and put all the records in a single spreadsheet. The proper positioning and length of each column should be verified with the *Time* index. When using MS Excel, it should be verified if the time format was not automatically adjusted by the default setting of the software (the constant and uniform time format should be used in a single csv file). If all the above-mentioned is provided, the file should be saved as a csv file, using the option '*CSV (Comma delimiter) (*.csv)*' with proper naming (e.g., *ECD.csv*).

Additional caution is required when the ED file is generated. The consumption records should be distinguished due to the fuel used. In practice, a separate file should be made for electricity consumption, and another one(s) if other fuel(s) are used (e.g., gas). Separate analyses by means of CEAM are required if the examined building uses more than one fuel, rotating the ED files.

Moreover, the total consumption (*D5*) should be assumed with care: awareness of the collected historical data is required. By default, the *D5* is a sum of all preceding demands, but difficulties in distinguishing individual consumptions are highly possible. For instance, the user might have information on heating demand (*D1*), as well as a total consumption *D5* (from e.g., energy invoice), which combines all the electricity consumption (used for heating, cooling, ventilation, and humidification). Therefore, the user does not know the distinguished consumptions for *D2*, *D3*, or *D4*, but still can perform individual assessments for *D1* and *D5* demands. The above-mentioned is presented in examples #3 and #4 in subsection 7.6. The selection of the demands to be included in the analysis is done from the GUI level (see more in section 4).

The analysis by means of CEAM is performed for hourly inputs, yet the provided tool allows the use of daily or monthly energy consumption. If daily or monthly records are provided, the tool will first adjust the given records for hourly archives during preprocessing in Module 0; this process is explained in the subsection 5.1.

The user has full flexibility in choosing the span of the examined data frame; nevertheless, it is recommended to adjust all the input files to the same length. Using a full calendar year is also recommended, but not strictly from 1st January to 31st December. The shortest recommended period of analysis is one month.

Within the CEAM software documentation, four sets of exemplary input files are provided. Within the provided arrays, various combinations of stored records are included and examined in the exemplary application of CEAM software (see more in subsections 7.2-7.5). Users should use the shared files (shared with the developed tool[1,3]) to get familiar with the required template and to help them adjust their input files. The regularity of the inputs is a key to a successful computation with CEAM software.

Finally, the user can use the provided *_00_time_index_generator.py* script to generate any data frame with the given date and time formats, as well as frequency. The generated file will contain the *Time* column with all the records for the predefined data frame, which can be further used to generate input files for CEAM assessment. This script is available via the CEAM GUI as a simple sub-software, operable in the pop-up window (see more in subsection 5.7). The proper and precise adjustment of the given historical records with the generated data frame is essential, in particular, to respect all the gaps in the examined archives.

4. How to use CEAM software

The CEAM software interface is overviewed in this section. The explanation of the GUI is shown, the required actions are discussed, and the computing is introduced.

4.1. GUI overview

The GUI of the tool has 3 main parts (see **Figure 1**): the information header (assigned as 1, in yellow), the modules (assigned as 2, in blue), and the additional information (assigned as 3, in green) sections.

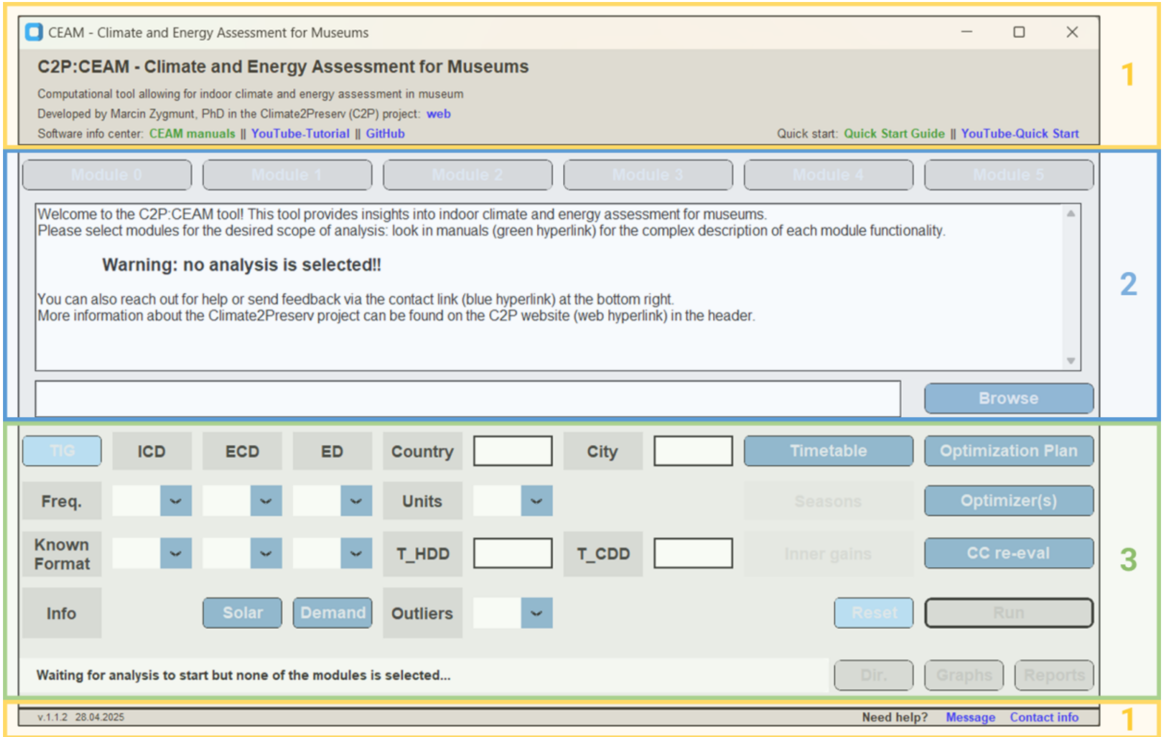


Figure 1. The GUI of the CEAM software with the main sections highlighted.

The information header consists of some contact information (C2P project website[5]), as well as provides active hyperlinks to the software manuals[35] (this document in particular), quick start guide[36] (included in the subsection 7.1), YouTube tutorials[37], as well as GitHub[3] repository. In the footnote, the information on the current version is shown, with the release date (left alignment), as well as a direct contact link to the developer[4] (right alignment).

[35] available here: https://drive.google.com/file/d/14pi2nRSRpZwOy8r0y87SqgOQ7U6qvplu/view
[36] available here: https://drive.google.com/file/d/1HdxbHnI0BT-e0AznzwNTVAQfMDlq8crr/view
[37] available here: https://www.youtube.com/@climate2preserv

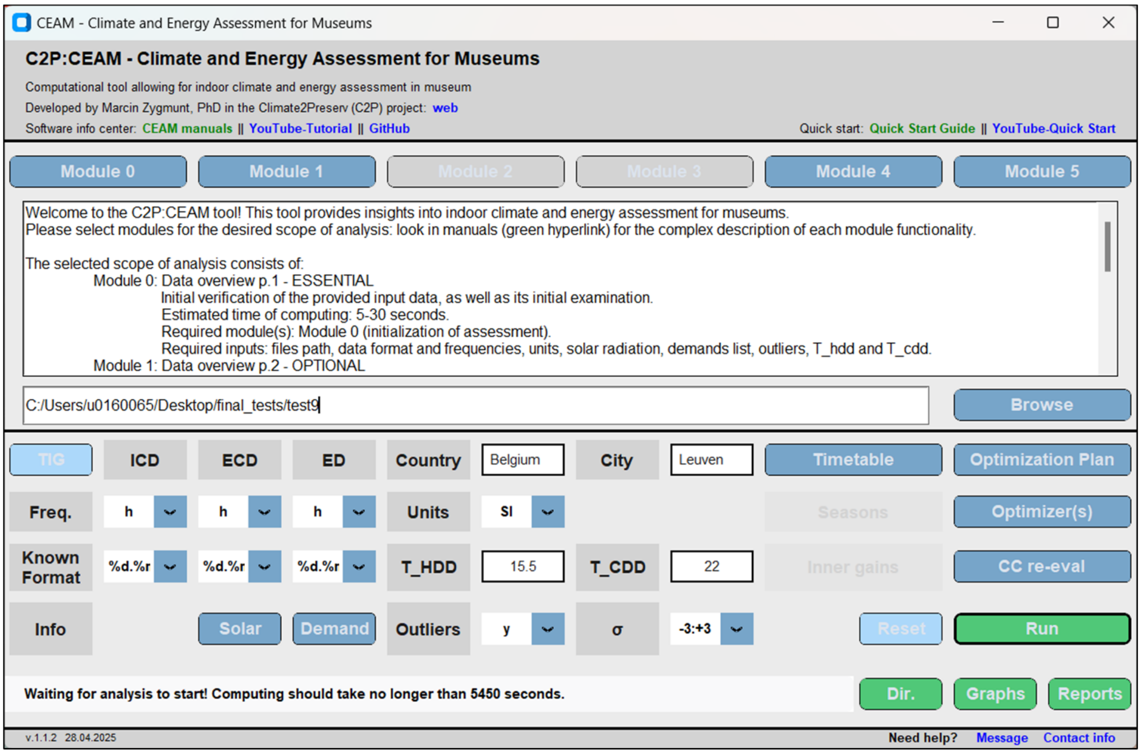


Figure 2. The GUI of the CEAM software with some made choices.

The modules section consists of 6 functional buttons, a scrollable textbox, as well as a browse box. The text displayed in the textbox corresponds to the selected scope of analysis, which is defined by activating the module buttons. By default, all the module buttons are off (in gray), and you can activate them by simply clicking on them (the button will turn blue). Users can freely define the scope of the assessment, yet during the very first run, it is obligatory to start with *Module 0* included (see more in the subsection 5.1). The displayed text in the textbox will update with the performed selection, showing a brief description of each module (see more in the section 5), with the estimated computing time. For the selected module, a list of required inputs is printed, as well as information on which previous modules are needed to be executed for a successful analysis. Finally, the browse button should be used to show the location of the working folder with the input files. By clicking the button, the typical system file explorer pop-up window will open to select the chosen localization (path). The chosen path will appear in the adjacent textbox when selected. When the path is selected, the run button (in the additional information section) will become active (it turns green from gray). The CEAM GUI with exemplary selections made in the module section is shown in **Figure 2**.

The information part of the GUI requires the most attention from the user. In this section, all the necessary additional information regarding the provided inputs and the examined case is required (see **Figure 3**). Firstly, the record frequencies of data (assigned as 1) in all the input files should be selected from the expandable option menu: the user can choose between monthly (*m*), daily (*d*), and hourly (*h*) granularities. The hourly frequency should also be selected if the provided records have a lower frequency (e.g., 15min). Next, the format of the *Time* records in all the input files should be defined (assigned as 2). The choice in expandable option menus is not active till the selection of the given frequencies is made, as well as the available selection is adjusted by the previous selection. The list of available Time formats for given frequencies is shown in **Table 3**.

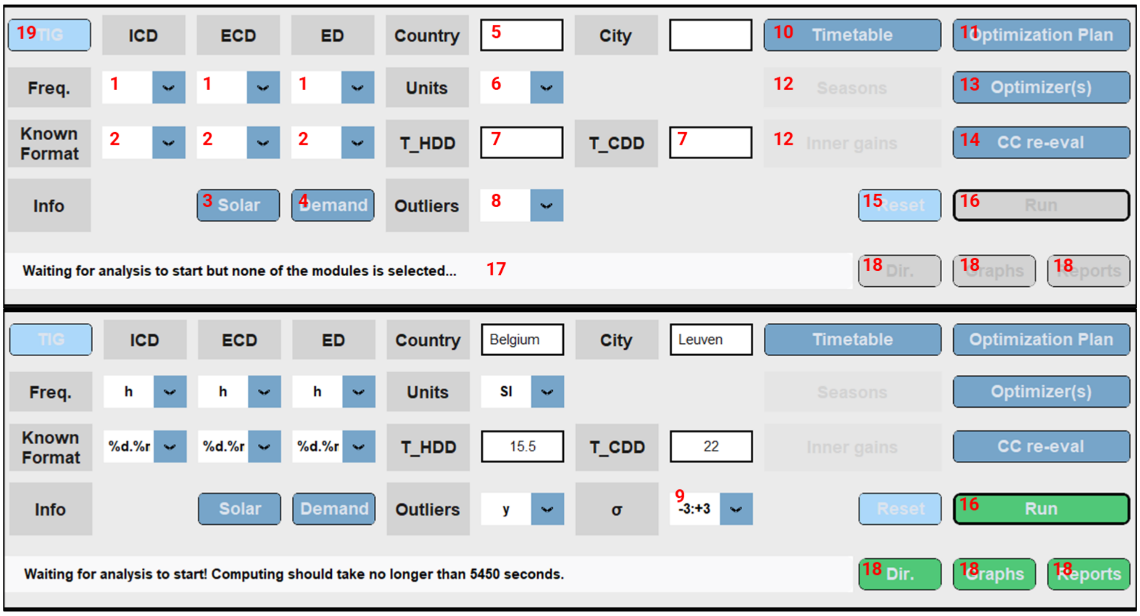


Figure 3. The additional information part of the CEAM GUI: default and empty stage (on top) and the fulfilled stage (on bottom) with assignments (in red).

The ECD file can contain solar radiation data; thus, it must be specified (assigned as 3). After clicking on the *Solar* button, a new pop-up window will open, with a *y/n* (yes/no) expandable option available. The selection should be confirmed with the green *Accept* button (see in **Figure 4** on the left). The ED file can contain various demands available for the assessment; thus, it is required to specify (assigned as 4). After clicking on the *Demand* button, a new pop-up window will open, with a checkbox list to be adjusted. By default, none of the demands is selected. After the appropriate selection, the choice should be confirmed with the green *Accept* button (see in **Figure 4** on the right).

Table 3. Overview of the Time format definition in CEAM.

FREQUENCY	AVAILABLE FORMATS	COMMENTS
(***h***)ourly	*%Y.%m.%d* *%Y.%d.%m* *%m.%d.%Y* *%d.%m.%Y*	**1.** The selection should be made without concerning the separation of the components. CEAM software will verify the used separator and adjust it to the default (dot). **2.** For hourly files the time formatting will be handled by the tool. *
(***d***)aily	*%Y.%m.%d* *%Y.%d.%m* *%m.%d.%Y* *%d.%m.%Y*	
(***m***)onthly	*%Y.%m* *%m.%Y*	

* as long as the provided format for time records is either only hourly (*HH*), hours and minutes (*HH:MM*), or hours, minutes, and seconds (*HH:MM:SS*).

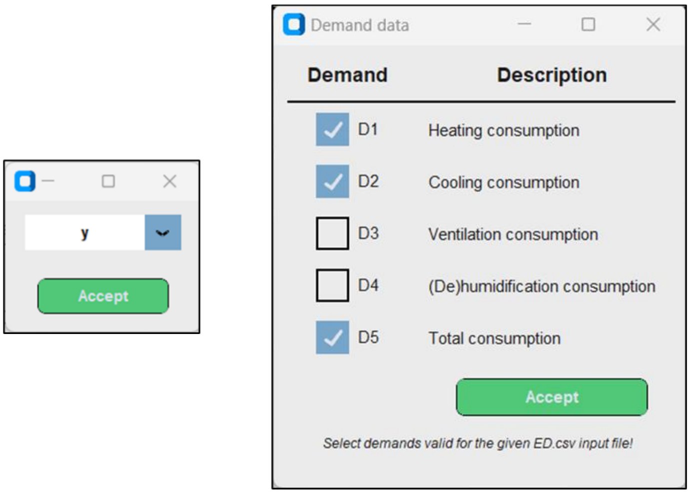


Figure 4. The additional pop-up windows for Solar (on the left) and Demand (on the right) definitions.

The *Country* textbox (assigned as 5) should be filled in with the examined localization. Users can input the full country name (e.g., Belgium), as well as the corresponding abbreviation (e.g., BE). CEAM will handle various fonts, yet the spelling must be correct (see example #5 in the subsection 7.6). A full list of the available countries can be found in **Table 5** in the subsection 7.8. If no input is provided in the *Country* textbox, Belgium is assumed as the default. The input in the *City* textbox has no further functionality in the current version of the CEAM software.

Units used in the prepared input files should be defined. The *Units* expandable option is available (assigned as 6), with SI and IU choices. All the records in the input files will be adjusted based on the selection made, in particular, converted into SI if IU is selected. If no selection is made, the SI system is assumed as the default.

Two textboxes (assigned as 7) are given to provide information about heating degree day (HDD) and cooling degree day (CDD) temperatures[24], named accordingly as *T_HDD* and *T_CDD*. The input values should follow the unit system defined in the *Units* section. If no value is provided, or if the provided input is incorrect (not a number), then the typical values (in °C) for Belgium are assumed[38]; 16°C and 21°C, for *T_HDD* and *T_CDD* respectively. The provided values will notably impact the received outputs, especially for the analyses based on inputs of daily/monthly energy demands.

CEAM is capable of examining the outliers in the provided records: it should be selected in the *Outliers* expandable option (assigned as 8). If the *y* (yes) option is selected, the range of the considered outliers should be defined in the σ section (assigned as 9). The outliers examination is performed based on the standard deviation[39] series, selected from the given expandable option. The outliers assessment is further explained in the 5.1 subsection.

The operation hours of the examined building should be defined in the *Timetable* section (assigned as 10). After clicking on the *Timetable* button, a new pop-up window will open, with a checkbox list and corresponding textboxes to be adjusted (see **Figure 5** on the left). Users should select the days when the examined building is open (i.e., the default setpoint temperature and humidity are fixed) and define the corresponding opening hours (in *From* and

[38] according to: https://doi.org/10.1016/j.enbuild.2020.109935
[39] see more: https://en.wikipedia.org/wiki/Standard_deviation

Till textboxes). Only the complete inputs are considered, i.e., a day with working hours (a full hour should be input, using 24-hour timing). The definition should be confirmed with the green *Accept* button. An exemplary fulfillment of the *Operational Schedule* window is explained in subsection 7.6 (example #6).

The energy optimization is performed based on the short-term modification of the given setpoints for temperature and humidity management (see more in the subsection 5.6). In the *Optimization Plan* section (assigned as 11), the predefined scenarios for the optimization should be selected. Users should choose the desired to-be-examined *Schedule*, based on the provided description in the *Details* column (see **Figure 5** on the right). The selection should be confirmed with the green *Accept* button.

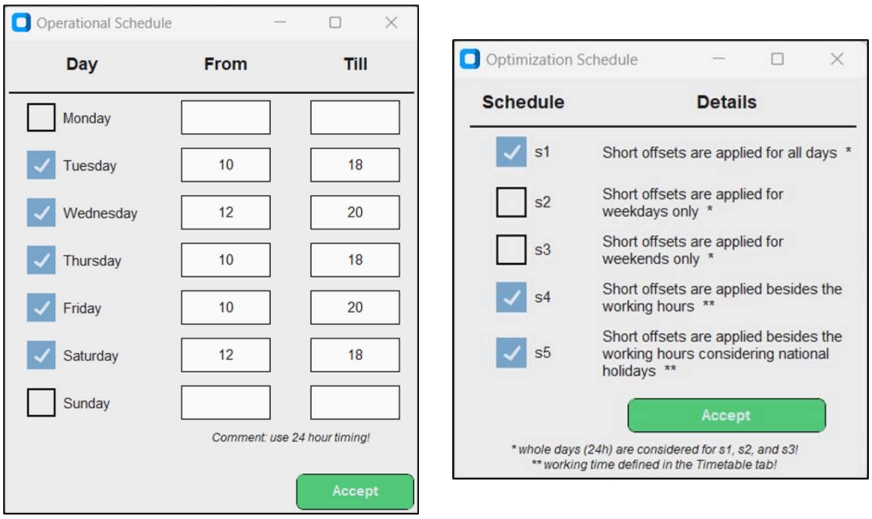


Figure 5. The additional pop-up windows for Operational Schedule (on the left) and Optimization Schedule (on the right) definitions.

The part of the GUI containing the *Seasons* and *Inner gains* labels (assigned as 12) is not yet developed. It is planned to include these additional features in the future version of CEAM.

The *Optimizer(s)* button (assigned as 13) will open the pop-up windows allowing to selection of the optimization algorithms used in module 5 to perform energy optimization (see more in the subsection 5.6). Users should choose the desired models to be used for optimization purposes (see **Figure 6** on the right). The selection should be confirmed with the green *Accept* button.

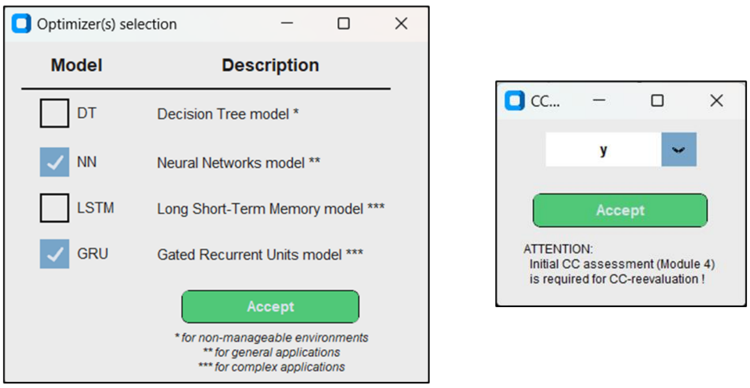


Figure 6. The additional pop-up windows for Optimizer(s) selection (on the left) and Climate Class re-evaluation (on the right) definitions.

CEAM software allows for the Climate Class (CC) assessment in accordance with the ASHRAE 55 requirements[40] (see more in the subsection 5.5). After the assessment of energy optimization, users will receive the modified temperature and relative humidity distributions due to the applied short-term strategies regarding indoor climate management. Therefore, the indoor climate re-assessment following the ASHRAE CC requirements is available, as shown in the *CC re-eval* button assigned as 14. After clicking the button, the pop-up window will open with a yes/no expandable option available. The selection should be confirmed with the green Accept button (see in **Figure 6** on the left). The received CC overview is performed analogous to Module 4, for the selected, optimal scenarios.

The information part of the GUI has a *Reset* button (assigned as 15). All the choices made in the information part can be erased after clicking this button (it does not affect selections made in the module selection).

The *Run* button (assigned as 16), which initially is inactive (in gray), turns green after providing the working directory. After starting the calculation, an *Analysis* pop-up window will show (see **Figure 7** on the left) and the *comment textbox* (assigned as 17) will present the status of computing. When computing is finished, the pop-up window will update (see **Figure 7** on the right), followed by the system sound. Users are required to confirm the end of the calculation by pressing the *OK* green button.

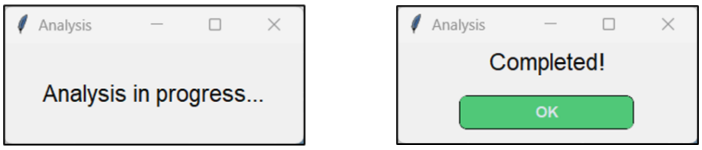


Figure 7. The computing pop-up Analysis window: calculations in progress (on the left) and finished computing (on the right).

Next to the comment textbox, the last three buttons are located. The *Dir.*, *Graphs*, and *Reports* buttons (assigned as 18) are initially inactive (in gray), and will become active (in green) if folders with button names are available in the selected working directory (the given path in the modules section). If the corresponding folder is present, it will open the typical system explorer window after clicking the matching button. Therefore, the *Dir.* button will become active after choosing the working directory, while the remaining two buttons will become active whenever the calculations of *Module 0* are completed (either in the given run of the CEAM or during any of the previous usage).

The last feature is assigned with number 19, as a *TIG* (Time Index Generator) button. It provides the functionality of the sub-software for generating the time index. The sub-software is generating a csv file named *time_index_{min}*, where the *min* is defined by the user as a time frequency. This feature works based on the *_00_time_index_generator.py* script, and a definition in the pop-up window (see **Figure 8** on the left) is required for a successful execution. The path where the outputted file should be saved is required; it is selected via the *Browse* button, in a similar way as for the working directory definition (in the modules section). Next, the starting and ending dates are required for the considered time frame, using the *YYYY-MM-DD HH:MM* format. The output date format is the following required input: the Python

[40] as described in Chapter 24: Museums, Galleries, Archives, and Libraries; shared here: https://drive.google.com/drive/folders/1AOUCIsjCwcblJIMxYSYoOU5_2tLTE3Yg?usp=sharing

programming language format is required, as shown in **Table 3** (e.g., *%d.%m.%Y %H:%M*). Finally, the frequency (i.e., time step) in minutes is to be provided; the value is required (e.g., 5). The exemplary definition is shown in **Figure 8** on the right, and the exemplary outcome is shown in example #7 in the subsection 7.6.

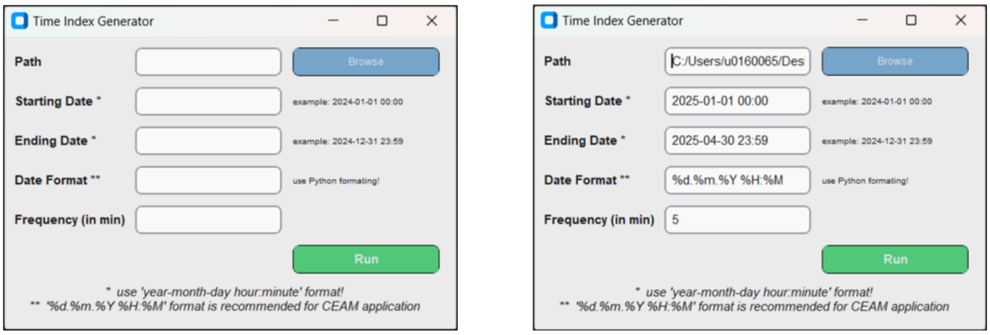


Figure 8. The TIG pop-up window: the default (on the left) and filled-in views (on the right).

Some sets of the CEAM GUI inputs are presented in subsections 7.2-7.5 where the exemplary applications are discussed.

4.2. Computing

The computing for the selected scope of the analysis (i.e., modules selection) is performed upon clicking on the *Run* button. The simulation being in progress is indicated with the lively updating *Comment* textbox, as well as the *Analysis* pop-up window (see **Figure 7** on the left). The scope of the analysis is presented in detail in the section 5 distinguishing each module in the consecutive subsections. The computing time varies mostly depending on the selected scope of assessment, as well as the specs of the used computer. For each module selection, the estimated time of computing is provided in the main textbox (module description), as well as in the *Comment* textbox. When computing is finished, the user is informed by the sound as well as the updated Analysis pop-up window (see **Figure 7** on the right); following, the user is asked to close the window. It is recommended to keep the used machine awake during the computing (i.e., turning off the power sawing options).

All the generated outputs are saved in the given working directory: the same directory where the input files were originally located. The files and outputs management is performed automatically by the software during computing. All the archives are sorted in a way presented in **Table 4**. During the computing, the dedicated folders are generated, allowing for efficient file management. The csv files mentioned in **Table 4** are the main files for the successful simulation by means of CEAM software.

It is important to mention that whenever the CEAM is run again (re-run of the analysis), the generated outputs will overwrite the existing ones (even if the input variables are different). Thus, if derivative analysis is required, users should copy the initially obtained results if they are needed. Also, users should not intervene with the files while CEAM is computing.

When the desired assessment is completed, users can freely close the software without any further actions. The input definition process is required to be repeated every time the software is opened.

Table 4. Files and outputs management by means of CEAM software.

FOLDER/FILE	DESCRIPTION
graphs	The *graphs* folder contains all the graphical outputs of the performed assessment. The folder includes 5 subfolders that store the outputs of different modules. Each graph generated is named to help users navigate: a complete list of possible generated graphs is shown in subsection 7.8.
others	The *others* folder contains all the temporary files generated by the CEAM software for a successful analysis.
raw_data	The *raw_data* folder stores the original input files, without any changes.
reports	The *reports* folder consists of text files with summaries and comments on the performed assessments. The generated reports will be used in the future version of the CEAM to provide one main report as an extra output of the performed analysis.
d_comb.csv	One of the key data files used during computing, with daily records. The file was generated during *Module 0* from the input files.
h_comb.csv	One of the key data files used during computing with hourly records. The file was generated during *Module 0* from the input files. This file is considered the default input file for the assessments by means of CEAM (*Modules 1-5*).
h_comb_CC.csv	It is an extended version of the *h_comb.csv* file required for CC assessment (the file is generated only if *Module 4* is examined).
m_comb.csv	One of the key data files used during computing with monthly records. The file was generated during *Module 0* from the input files.
w_comb.csv	One of the key data files used during computing with weekly records. The file was generated during *Module 0* from the input files.
terminal_records.txt	The text file contains all the information which typically shown in the terminal (code-level). In case of facing errors while computing, users can share this file with the developers for help. The file is constantly updating while the simulation is on. There is no practical application of this file for users.
var.txt	The text file is generated every time the Run button is pressed, collecting all the provided inputs in the CEAM GUI. The *var.txt* file is used by the tool for successful computing. The content of the file is shown in subsection 7.8.

5. MODULAR STRUCTURE: OVERVIEW

CEAM has a modular structure consisting of Python codes. Each of the modules provides a specific functionality, delivered by primary and supplementary scripts. The naming of primary codes is initiated with *1*, while the supplementary scripts start with *0* (see **Table 1**). The numbering of the modules starts with 0 because the first one is essential for suitable computing performed by means of the CEAM software. A description of each module is given in the following subsections; each script and its functionality are mentioned, as well as included methods are presented. Moreover, a simple text-based file report is generated for each module executed.

5.1. MODULE 0: DATA OVERVIEW PART 1

Module 0 is the first of the available six. The main purpose of this module is to verify the provided inputs, as well as to perform all the necessary adjustments required for further assessments, performed within the following modules of the CEAM software. Thus, *Module 0: Data Overview Part 1* is obligatory to be included in each run of the CEAM software. The result of completing Module 0 is a set of key data files, in particular *d_comb.csv*, *h_comb.csv*, *m_comb.csv*, and *w_comb.csv* (see **Table 4**). This module should be run only once for each assessment, repeated every time the new inputs are provided, using the raw files. It consists of 10 scripts (2 primary and 8 supplementary codes) presented below (in the execution order).

INITIAL CVS CHECK

The initial csv check is performed by the *_11_initial_csv_check.py* script, and the main objective of this part is to verify if the provided input files are correct (following the software requirements). This script generates all the supplementary folders in the working directory to support transparent file management, which includes renaming, moving, copying, merging, as well as deleting the files. The *01_missings_report.txt* report file is generated after executing the script. There are several main functionalities performed by this script, listed as follows:

- check and (if needed) change the separator,

A function providing a separator (a symbol used to discrete each record) verification, and adjusting it to the default one if needed (see example #2 in the subsection 7.6).

- process the input files,

A function is used to adjust the provided structure of the input files to follow the default one. It focuses on the provided headers: their font and placement. If the restructuring is necessary, the whole column is reordered; see example #1 in the subsection 7.6.

- check and (if needed) change the date format,

This part of the module is to check and fix the date format in the *Time* column of csv input files. It verifies the given format based on the information provided by users in the GUI and adjusts it to the default one (see example #2 in the subsection 7.6).

- adjustment of the ED file with lower (daily or monthly) frequency data,

This is a function that adjusts the monthly or daily energy demand file based on the heating or cooling degree days[24]. Based on the provided input file with low frequency or records, it is possible to generate an hourly-based ED file with the corresponding demands. The adjustments of the energy consumption are made based on the HDD and CDD, calculated as defined in the *_03_degree_days.py* script (see below). Only heating (*D1*) or cooling (*D2*) assessments are possible whenever the user provides the ED file with daily or monthly records. *D1* records are adjusted based on the HDD distribution, while *D2* on the CDD.

- examine outliers,

A supplementary function that is executed only upon selection in the GUI. If an outliers assessment is selected (see the subsection 4.1) the script will detect the records outside the given distribution range and consider them as outliers. All the selected outliers will not be included in the key data files, keeping empty records of data for the specified periods. The range of the considered data points is defined by the standard deviation, calculated with **eq.1** (see the subsection 7.7). The visualization of the considered outliers' shares in a normal distribution[41] is shown in **Figure 9**[42]. The mentioned ranges are based on the mean value (μ) and the scopes defined by subtraction or addition of the σ (i.e., how many standard deviations we are away from the mean). The range of $\mu\pm\sigma$ contains approx. 68.3% of the data, range of $\mu\pm2\sigma$ approx. 95.4%, while range of $\mu\pm3\sigma$ approx. 99.7%. Thus, the selection made in the GUI should be thoughtful: if the data is corrupted, with numerous irregularities, the narrower range is recommended, while in most cases, where the collected data is accurate, the broader range (or even assessment without outliers examination) should be selected.

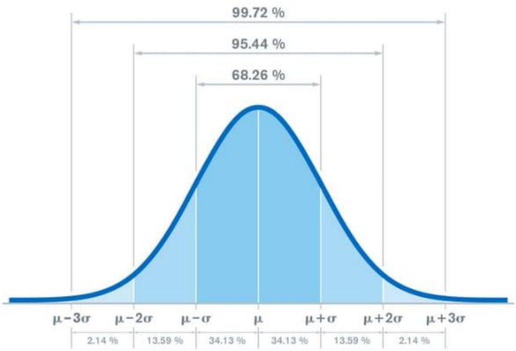


Figure 9. Standard deviations in normal distribution.

- check for the missing values,

A function used to verify if records are missing, thus no data is present for a given *Time* log. The function also reports on the length and appearance of shortages in data.

- most common frequency,

Functionality used to check the granularity of the record (e.g., 15min).

- fill missing timestamps,

A function used to detect the missing gaps in data, as well as to fill up with empty records (rows); see example #8 in the subsection 7.6. The added records follow the prerecorded frequency, as well as the *Time* format.

- averaging data,

Script performing records averaging to the hourly values (only if frequencies in the provided input files are shorter, e.g., 15min); see example #8 in the subsection 7.6.

[41] see more: https://en.wikipedia.org/wiki/Normal_distribution
[42] source: https://www.sixsigmadaily.com/standard-deviation-6-steps-to-calculation/

- convert units.

Functionality is provided in the *_01_convert_units.py* script to substitute the records given in the IU unit system to the default SI system. The parameters considered are temperature (both T_e and T_i in the corresponding ECD and ICD files) and all the energy demands in the ED file. For temperature, the [C] to [F] conversion[43] is used, while for energy demand, the [Wh] to [BTU] translation[44] is executed.

CALCULATE THE VAPOR PRESSURE

The functionality provided in the *_02_pv_calc.py* script is to receive the vapor pressure (p_v). The calculation is based on the *Magnus-Tetens* equation (see **eq.2** in the subsection 7.7), as well as the function defining the relation of the saturation vapor pressure and the relative humidity (see **eq.3** in the subsection 7.7). The script provides the p_v values for indoor ($p_{v,i}$) and outdoor ($p_{v,e}$) climates, and saves them in the corresponding key files. A similar calculation can be performed using external tools[45], or assumed manually from the psychrometric charts as shown in **Figure 10**[46].

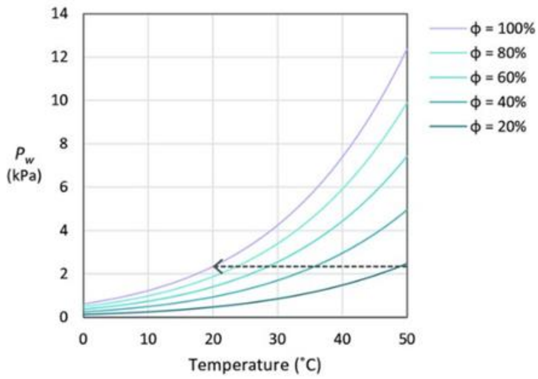


Figure 10. Vapor pressure as a function of temperature and relative humidities.

CALCULATE THE DEGREE DAYS

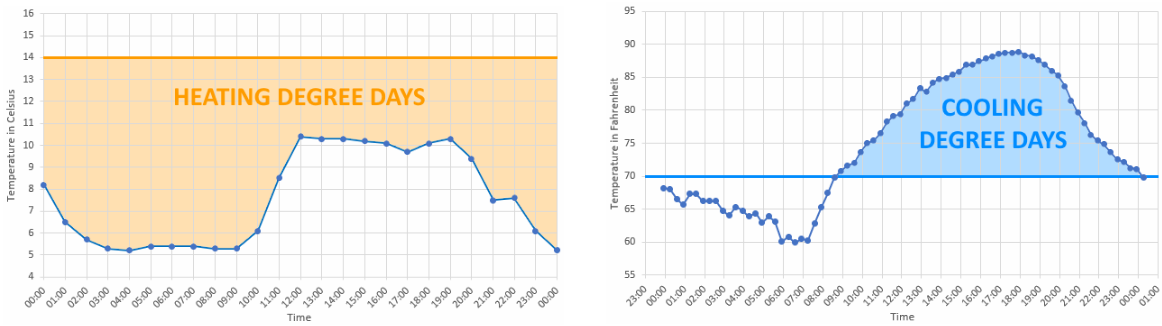


Figure 11. Heating (on the left, using [ºC] and hourly frequency) and Cooling (on the right, using [ºF] and 20min frequency) Degree Days assessment with the base temperatures (horizontal lines).

The functionality is defined in the *_03_degree_days.py* script, allowing for the calculation of heating and cooling degree days. Both HDD and CDD are calculated based on the corresponding base temperatures (T_{HDD}, T_{CDD}) provided as input (see more in the subsection 4.1). The usage of HDD/CDD is to assess the impact of weather (the most important boundary condition impacting building performance, in both energy demand and indoor climate[47,48]). The degree days are calculated based on the T_e records (see **eq.4** and **eq.5** in the subsection

[43] using formula: $T\ (\text{in } ^{\circ}\text{C}) \cdot {}^{9}/_{5} + 32$
[44] using formula: 1Wh = 3.412BTU
[45] e.g., available here: https://www.weather.gov/epz/wxcalc_vaporpressure
[46] source: https://doi.org/10.3390/geosciences6030038
[47] source: https://doi.org/10.3390/su7079207
[48] see more: https://doi.org/10.1016/j.apenergy.2019.03.167

7.7) for each frequency and saved in the corresponding key data files. The amount of HDD and CDD can also be calculated using various web tools[24]. An exemplary graphical assessment of daily HDD and CDD is shown in **Figure 11**[49].

The calculation method of HDD and CDD is used for demand records assessment into hourly granularity if the provided ED input file uses daily or monthly frequencies; the process is explained in the subsections 7.4 and 7.5.

STATISTICAL OVERVIEW OF THE INPUT DATA

The functionality is provided in the *_12_data_overview_stats.py* script. The main objective is to examine the temporary data files, which in the end will be changed into key data files, for their statistical overview. Each variable is checked for validity (e.g., whether the recorded humidity is between 0 and 100), and any exemptions are recorded if present. The statistical measures are calculated for the base variables, in particular, max, min, average, and median values, as well as 1st (lower) and 3rd (upper) quartiles, for the summary statistics[50] of the inputs; data assessment based on the boxplots is explained in **Figure 12**[51]. Also, daily, weekly, and monthly averages are computed based on the hourly records. The *02_stats_report.txt* report file is generated after executing the script.

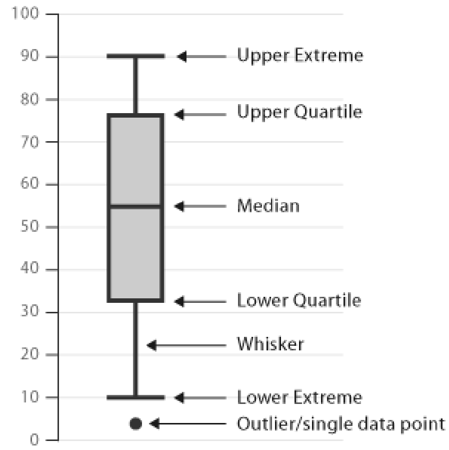


Figure 12. Introduction to data analysis using boxplots.

EXTREME WEEKS ASSESSMENT

The *_04_extreme_weeks.py* script defines 6 extreme weeks (Extreme Dry Week[52] – *EDW*, Extreme Humid Week[53] – *EHW*, Extreme Summer Week[54] – *ESW*, Extreme Winter Week[55] – *EWW*, Extreme Sunny Week, negative[56] – *ESunW_n*, Extreme Sunny Week, positive[57] – *ESunW_p*) that are based on the exterior climate condition (provided in the *ECD.csv* file). The *EDW* week is the one with the lowest RH_e for the examined data frame, and the *EHW* is the one with the highest. The *ESW* is the one with the highest T_i, and the *EWW* with the lowest. Finally, the *ESunW_n* is the week with the lowest sum of total radiation, and the *ESunW_p* with the highest. The associated datasets for the main input variables are generated. The purpose of this script is to provide an overview of the extreme weeks of the examined period. Assessments performed based on the extreme conditions are particularly important for various actions, e.g., HVAC design or building performance overview[58].

[49] source: https://www.degreedays.net/introduction
[50] see more: https://en.wikipedia.org/wiki/Summary_statistics
[51] source: https://datavizcatalogue.com/methods/box_plot.html
[52] a week (7 days) characterized by the lowest average exterior humidity from the examined period (input)
[53] a week (7 days) characterized by the highest average exterior humidity from the examined period (input)
[54] a week (7 days) characterized by the highest average exterior temperature from the examined period (input)
[55] a week (7 days) characterized by the lowest average exterior temperature from the examined period (input)
[56] a week (7 days) characterized by the lowest sum of solar radiation from the examined period (input)
[57] a week (7 days) characterized by the highest sum of solar radiation from the examined period (input)
[58] source: https://doi.org/10.1016/j.buildenv.2023.110124

GENERATE AND ADJUST COMBINED FILES

The functionality provided in the *_05_1_combi_files.py* combines the content of temporary files into the key data files, with hourly, daily, weekly, and monthly records. The necessary adjustments (e.g., correct allocation based on the *Time* indexes) are provided, and the past files are secured (archived). In the *_05_2_adjust_combi_files.py* script, some additional verification of the already generated key data files is performed (e.g., *Time* format check). All the actions provided by the above-mentioned scripts secure the functionality of the following analyses.

DEW POINT AND WET-BULB TEMPERATURES

Script *_06_1_Td_Twb_calc.py* is used to calculate the dew point (T_d)[25] and wet-bulb (T_{wb})[26] temperatures. The calculated values are added to the already existing files. The calculation method is shown in **eq.6** and **eq.7** (see the subsection 7.7). The calculated variables are used for energy optimization, discussed in the subsection 5.6.

ADDITIONAL FACTORS FOR THE CLIMATE AND ENERGY ASSESSMENTS

Script *_06_2_factors.py* is used to calculate some additional factors used for further evaluation in CEAM software. Differences between the pre-calculated factors are received, in particular for air, dew point, and wet-bulb temperatures, designated as *deltaT*, *deltaT*$_d$, and *deltaT*$_{wb}$ (see **eq.8**-**eq.10** in the subsection 7.7). The variation of relative humidity (*deltaRH*) is also provided (see **eq.11** in the subsection 7.7).

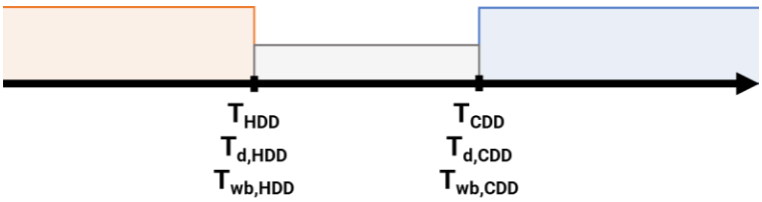


Figure 13. Schema of the TIS, TDIS, and TWBIS factors definition.

Moreover, unique parameters used for energy optimization (see more in the subsection 5.6) are defined and evaluated. The above-mentioned parameters consider solar radiation's impact on the overall building energy performance, as well as indoor climate behavior. Overall, solar radiation has a positive impact on heating demand and a negative impact on cooling demand[59]. Solar radiation also impacts the indoor climate through thermal and visual comforts[60]. The evaluated factors are designated as *TIS* (based on air temperature), *TDIS* (based on dew point temperature), and *TWBIS* (based on wet-bulb temperature). The graphical overview of how the above-mentioned factors are evaluated is shown in **Figure 13**. *TIS* factor is calculated for every record using the formula shown in **eq.12**; the impact of solar radiation is included only if the T_e is lower than T_{HDD} or higher than T_{CDD}. *TDIS* factor is calculated in the same way, yet using the supplementary values of $T_{d,HDD}$ and $T_{d,CDD}$; those specific values of dew point temperatures are calculated using **eq.6** with T_{HDD} and $RH_{e,HDD,ave}$[61] or T_{CDD} and $RH_{e,CDD,ave}$[62] as corresponding inputs. *TWBIS* is obtained in a similar approach, using a wet-bulb

[59] impact of solar heat gains on energy consumption; see more: https://doi.org/10.1016/j.enbuild.2012.10.050
[60] see more: https://doi.org/10.1007/s00484-006-0050-y
[61] average value from the RH records which are allied with the presence of HDD (during the heating season)
[62] average value from the RH records which are allied with the presence of CDD (during the cooling season)

temperature formula (see **eq.7**) with T_{HDD} and $RH_{e,HDD,ave}$ or T_{CDD} and $RH_{e,CDD,ave}$ as corresponding inputs. The aforesaid parameters are considered only if solar radiation (I_s) is provided in the ECD input file; otherwise, those factors are skipped in the performed assessment.

ABSOLUTE HUMIDITY

Script *_06_3_abs_humi.py* is used to calculate the absolute humidity (*AbsH*)[27] based on exterior ($AbsH_e$) and interior ($AbsH_i$) parameters. The script also provides the difference of absolute humidity (*deltaAbsH*), showing the variation of this parameter. The calculated values are added to the already existing files. The calculation method is shown in **eq.13** in the subsection 7.7. This script provides the difference for absolute humidity as well.

5.2. MODULE 1: DATA OVERVIEW PART 2

This module provides a graphical overview of the provided inputs, as well as all the corresponding variables calculated from the given records. The visualization is provided based on the data in the generated files, in particular, *d_comb.csv*, *h_comb.csv*, *m_comb.csv*, and *w_comb.csv*. The functionality of *Module 1: Data Overview Part 2* is provided by only one script named *_13_data_overview_plots.py*. The script contains the formula to generate scatter and bar plots (see **Figure 14**), adjust the axis and labels, as well as provide an accurate sorting of the generated figures (i.e., file management). Minimal and maximal values, as well as the average line, are plotted on each graph. Moreover, the empty records are highlighted if present in the given dataset.

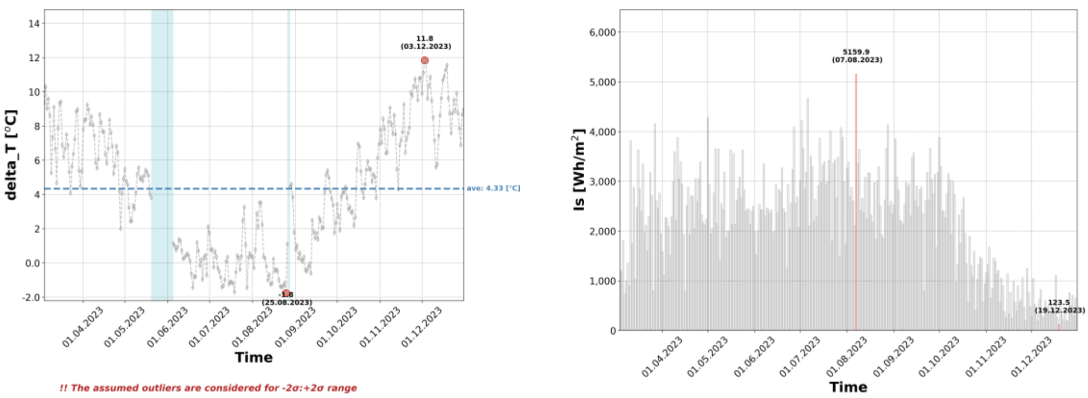


Figure 14. The exemplary presentation of the graphical overview of the examined data using a scatter plot (on the left) and a bar plot (on the right).

More examples of the performed graphical overview of the examined data are shown in the subsections 7.2-7.5.

5.3. MODULE 2: CORRELATION ASSESSMENT

This module allows users to perform correlation[63] assessments based on the predefined sets of variables in the examined data files. Correlation assessments[64] are performed for both indoor and outdoor climate conditions, often used to evaluate the influence of various

[63] i.e., any statistical relationship, whether causal or not, between two random variables or bivariate data; see more: https://en.wikipedia.org/wiki/Correlation
[64] see more: https://doi.org/10.1016/j.enbuild.2012.08.037

variables on energy use[65]. The functionality of this module is provided in one script named *_14_correlation.py*. The script includes a standard procedure used to fit a linear regression using the least squares method, with the evaluation performed using the residual sum of squares (RSS)[66]. Whenever the regression fitting is achieved, the graphs are printed for the examined sets of variables. A complete set of the available combinations is shown in **Table 6** in the subsection 7.8. Only the linked variables are examined; e.g., a correlation between indoor relative humidity (RH_i) is not analyzed with heating consumption (*D1*). The correlation assessment is performed for hourly, daily, and monthly data granularities (see **Figure 15**). The accuracy of the performed matching is evaluated and shown in the generated report file named *04_correlation_report.txt*. The reporting contains information about the examined variables, the validity metrics with the coefficient of determination (R^2)[67] and Pearson correlation coefficient (r)[68], as well as the obtained regression line equation. The perfect r value can be either 1 or -1, where the negative value indicates an inverse relationship between two examined variables. The selection of the best-performing correlations (mentioned in the *04_correlation_report_selected.txt* file) is also provided within the functionalities of this module. Finally, the file management after computing is also provided.

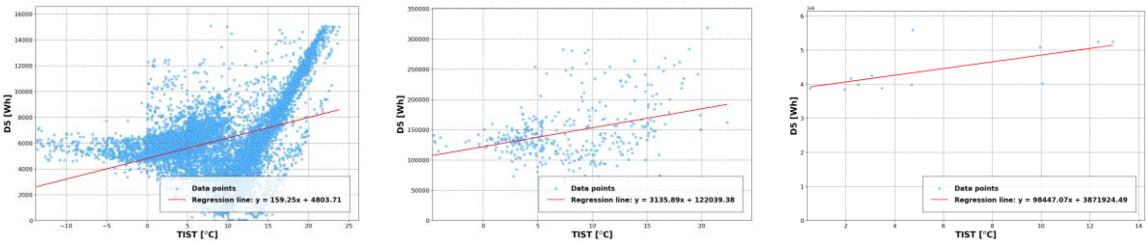


Figure 15. The exemplary presentation of the correlation assessment of the examined data using hourly (on the left), daily (in the middle), and monthly (on the right) records.

More examples of the performed correlation assessment of the examined data, along with guidance on interpretation, are shown in the subsections 7.2-7.5.

5.4. Module 3: Energy Signature

The Energy Signature (ES) method, introduced by Hammarsten[69], is an evaluation method in which energy consumption is correlated with climatic variables aimed at representing the actual energy behavior of the building, as well as estimating baseload (not-weather-dependent energy consumption) and base temperature[70]. This approach requires minimal input and involves plotting energy use against outdoor temperature over time. It is widely used in energy auditing, baseline modeling, and performance verification for various purposes, particularly to identify the malfunctions of a building, to evaluate energy refurbishment, to raise users' awareness of consumption, and to assess the performance of the examined building[71].

ES assessment in CEAM software is provided by the empirical evaluation with the change-point models. The functionality of module 3 is provided within one script named

[65] see more: https://doi.org/10.1016/j.engappai.2022.105287
[66] see more: https://en.wikipedia.org/wiki/Residual_sum_of_squares
[67] see more: https://en.wikipedia.org/wiki/Coefficient_of_determination
[68] see more: https://en.wikipedia.org/wiki/Pearson_correlation_coefficient
[69] see more: https://doi.org/10.1016/0306-2619(87)90012-2
[70] see more: https://en.wikipedia.org/wiki/Energy_signature
[71] see more: https://doi.org/10.1016/j.enbuild.2015.10.038

_15_energy_signature.py. The script includes the assessment of the energy consumption with temperature, using hourly and daily frequencies of data (see **Figure 16**). Typically, daily records provide sufficient detail to characterize the building's energy profile concerning climatic variables. In contrast, hourly records are more challenging to assess, as they are influenced not only by environmental factors but also by the building's intrinsic properties (e.g., thermal mass). Therefore, a constructive approach involves applying the ES method to available data and comparing results across different temporal resolutions[72]. The used change-point ES method can be described by the general equation for the overall consumption (5-parameters model) as shown in **eq.14** in the subsection 7.7. In **Figure 28** in the subsection 7.8, all considered models for the ES assessment in the CEAM software are presented. A graphical overview of the performed ES assessment, as well as reporting (in *05_ES_report.txt* file), is generated within this module of CEAM. The selection of the best-performing ES models (mentioned in the *05_ES_report_selected.txt* file) is also provided within the functionalities of this module. Finally, the file management after computing is also provided.

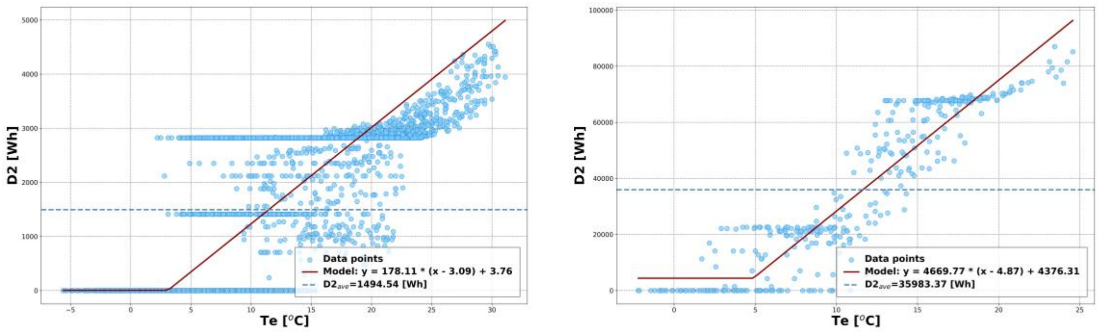


Figure 16. The exemplary presentation of the ES assessment of the examined data using hourly (on the left) and daily (on the right) records.

More examples of the performed ES assessment of the examined data, along with guidance on interpretation, are shown in the subsections 7.2-7.5.

5.5. MODULE 4: CLIMATE CLASSES OVERVIEW

Module 4 allows for assessing the recorded indoor climate conditions (as provided in the IEC input file) for the examined building/zone in the context of heritage preservation, particularly evaluating temperature and relative humidity records.

ASHRAE Handbook chapter 24 presents best practices and advice on planning, designing, and implementing environmental strategies for long-term preservation of cultural heritage that also support access in an economically and environmentally responsible way. In this document, CC are introduced as a conceptual technique to help define and manage environmental conditions in spaces housing sensitive cultural heritage, such as museums, libraries, archives, and galleries. This technique serves as guidelines for establishing indoor environmental setpoints and tolerances for temperature and relative humidity, tailored to the preservation needs via different levels of environmental stability and control. These guidelines account for permissible short-term fluctuations, seasonal adjustments, as well as acceptable upper and lower bounds for T and RH (see **Table 7** in the subsection 7.8). Therefore, CC allow

[72] see more: https://doi.org/10.1515/oere-2015-0008

for management of the indoor environment more efficiently without a negative impact on housed heritage. A more comprehensive overview of the CC, as well as the effectiveness of their application to preserve heritage, will be discussed in the article in the publishing[73].

CC assessment via CEAM provides intuitive graphical outputs and corresponding descriptive reporting. The reporting is given in the *06_CC_report.txt* file, showing the % fulfilment time for each CC, providing a score for a given class, as well as distinguishing several requirements (i.e., time-dependent offsets) defined in the ASHRAE standard (see example #9 in the subsection 7.6). Graphical results are shown in two ways, as psychrometric charts[74] or typical scatter plots (see **Figure 17**). Each of the generated graphs shows the limitations for the predefined filters for each class. Psychrometric charts are well known for specialists keen on heritage preservation, allowing for the visual presentation of the acceptable range of environmental conditions contrary to the recorded data. The time evolution scatter plots (presented individually for T and RH records) allow addressing the presence and duration of fluctuations, malfunctions, or issues (e.g., system malfunctioning) of the indoor climate. Evaluation of the historical records of the indoor conditions using all 3 outputs allows for a throughout understanding of the examined environment and rooms where it can be either improved or modified.

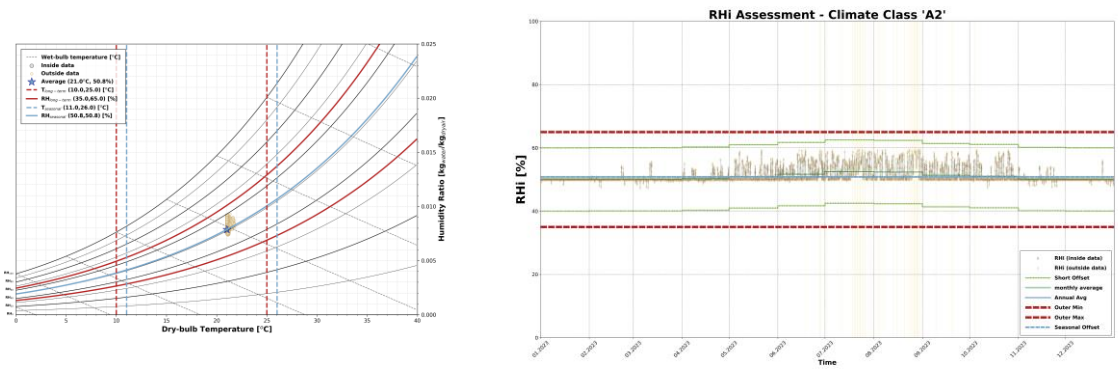


Figure 17. The exemplary presentation of the CC assessment of the examined data with a psychrometric chart (on the left) and a scatter plot (on the right).

More examples of the performed CC evaluation for the examined data, along with guidance on interpretation, are shown in the subsections 7.2-7.5.

5.6. MODULE 5: ENERGY OPTIMIZATION

Module 5 is the most complex in the CEAM software. It provides the energy savings predictions obtained with the short-term adjustments of the indoor climate conditions. Module 5 consists of 6 scripts (see **Table 1**), and for its successful computing, the post-processed files (generated by CEAM software from the input files) are required. Thus, to run energy optimization, the initial finalization of Module 0 is required, while outputs from Module 4 are also needed if we consider the CC-reevaluation. Due to the complexity of the performed calculations, Module 5 is time-consuming.

The first step is to generate the working file: *prediction_inputs_all.csv* is generated based on the user inputs provided in the *Timetable* and *Optimization Plan* (see more in the subsection

[73] the hyperlink to the article will be shared once it is published in print
[74] see more: https://en.wikipedia.org/wiki/Psychrometrics

4.1). The proposed offset for the temperature and relative humidity will be applied to the periods following the selected scenarios for optimization. The generated dataset also includes the remaining inputs provided, in particular, available demands and the selected country. This functionality is provided with the *_09_prediction_inputs.py* script.

Next, the predictions are calculated with the functionality provided by the *_17_predictions.py* script. Energy optimization is performed using Black-Box Modeling (BBM)[28]. The computing is performed by the optimization algorithms, which are trained to predict potential energy savings based on the feed-in data, in particular, historical records provided within the input files. The predicted energy savings resulting from the adaptation of short-term (i.e., offsets) indoor climate management strategies following the ASHRAE recommendations (see more in the subsection 5.5).

The available predictions in the current version of the software are linked with the examined demand; *D1* (heating only), *D2* (cooling only), *D4* (humidification only), as well as *D5a* and *D5b* (total consumption) are available (the received outputs are provided automatically, based on the inputs and prior definition by user). The *D1*, *D2*, and *D5a* analyses are performed based on temperature-based offset strategies, the *D4* on humidity-based, and *D5b* examines the simultaneous offsets for temperature and humidity strategies. CEAM allows for loosening the indoor climate management to climate class B, and thus, the available offsets are ±2°C or ±5°C for temperature and ±5% or ±10% for relative humidity. The visualization of the above-mentioned offsets can be found in the outputted scatter plots for climate class assessment (one of the outputs in Module 4; see more in the subsection 5.5).

Depending on the specific energy demand, various models are used, incorporating air, dry bulb, wet bulb temperatures, solar radiation, as well as relative and absolute humidities as key predictive variables (a short overview of each parameter can be found in the subsection 5.1). The variables' numerosity used for energy optimization is employed to obtain the best possible optimization model. The models employed include specialized algorithms, in particular Neural Network (NN)[75], Recurrent Neural Network[76] using Long Short Term Memory (LSTM)[77] or Gated Recurrent Units (GRU)[78], and Decision Tree Regression (DTR)[79]. The process of assessing the prediction model requires the training input used for the model to learn about the examined case. The training process is validated, as well as the selected version of the model is tested to check its accuracy. During the training process, the input dataset is split into 70% for training, 15% for validation, and the remaining 15% for testing. User can choose which optimizers to use (see more in the subsection 4.1): the NN model is considered universal for all kinds of applications, the LSTM and GRU are recommended for more complex analysis (the first one predicts the time-response better, while the second one handles peaks more efficiently), and the DT model works best with non-manageable environments and to predict the baseload. The valid model can provide accurate predictions based on the considered offset strategies. Yet, it is important to remember that these models

[75] see more: https://en.wikipedia.org/wiki/Neural_network_(machine_learning)
[76] see more: https://en.wikipedia.org/wiki/Recurrent_neural_network
[77] see more: https://en.wikipedia.org/wiki/Long_short-term_memory.
[78] see more: https://en.wikipedia.org/wiki/Gated_recurrent_unit
[79] see more: https://en.wikipedia.org/wiki/Decision_tree_model

are obtained based on the provided input dataset within some ranges: the trained model might be unreliable for predictions outside of the original array[80].

The optimization process includes the training, the validation, as well as the predictions. The first step is to train the model using the selected algorithm, as introduced above. The trained models are saved and used in validation. The validation process can be considered as a 2-step process. The first step is the testing based on the selected part of the input records; this process is done automatically, during the learning process. The second step runs the initial predictions based on the original inputs of indoor climate conditions. Moreover, for each model considered, the prediction outputs are saved in the temporary csv files stored in the *predictions_outputs* folder in the main *others* folder.

The base consumption is compared with the predicted counterpart obtained via the trained model. The validation (using the functionality provided in the *_18_1_best_predictions.py* script) is recorded with numerous factors, particularly the difference in total energy use and peak demand, as well as using statistical factors: coefficient of determination (R^2)[67], Root Mean Square Error (*RMSE*)[81], Mean Absolute Error (*MAE*)[82], and Concordance Correlation Coefficient (*CCC*)[83]. All the above-mentioned performance indicators are presented in the subsection **7.7** (**eqs.15**-**18**). Based on the recorded results of the training process, the best-performing models can be selected (see below), allowing for accurate optimization with the short-term offsets. The validation process is shown in the generated report file called *07_predictions_overview.txt* (see example #10 in the subsection **7.6**). Having the trained model, the selection of the best ones is performed. The best-performing models are selected for each of the examined demands and summarized in the *07_best_predictions.txt* file. The selection is performed based on a multi-metric ranking[84] that prioritizes the highest R^2, the highest *CCC*, and the lowest peak absolute difference (in the given order). Therefore, a model for each of the examined demand and prediction variables used is selected for energy optimization (see example #11 in the subsection **7.6**). Moreover, the selected prediction files are stored in the *predictions_selected* folder in the main *others* folder.

The *_18_2_best_predictions_graphs.py* scripts present the energy optimization result as graphs. All the generated visuals are saved in the dedicated folder (named *05_predictions*) in the main *graphs* folder. The graphs are furtherly sorted into the corresponding scenarios folder, as well as the folder for best outcomes. The exemplary graph with energy prediction obtained from a trained model is shown in **Figure 18**. The output consists of three graph with different time horizon, always showing the base demand (historical record in grey), the base (obtained on the inputted indoor environment records) prediction (in blue), as well as two optimization scenarios considering lower (±2°C or ±5%, in light green) or higher (±5°C or ±10%, in dark green) offsets. The upper half of the graph shows the whole analyzed data frame, while the ones below are zoomed in for a week period (with the highest base consumption) and a day with the peak demand recorded. For the day graph, the peak demand is shown with

[80] see more: https://doi.org/10.1111/2041-210X.13650
[81] see more: https://en.wikipedia.org/wiki/Root_mean_square_deviation
[82] see more: https://en.wikipedia.org/wiki/Mean_absolute_error
[83] see more: https://en.wikipedia.org/wiki/Concordance_correlation_coefficient
[84] additional methods on different and potentially more balanced selection are under investigation, possibly to be include in the future version of the software

a vertical dashed red line; here, it is not visible because the peak demand was recorded on the edge of the plot. Finally, the empty records are highlighted in pale yellow, informing about a lack of input data or records omitted due to the outliers assessment.

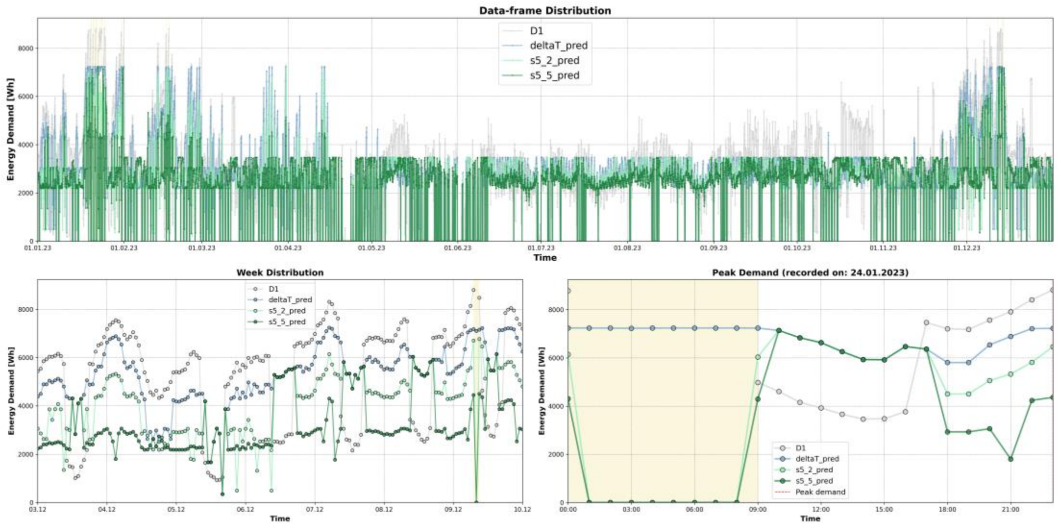


Figure 18. Exemplary graph with energy optimization with GRU model for *D1* demand: whole data frame (the upper part), demand for week distribution (on the lower left), and daily overview for the peak demand (on the lower right).

Lastly, all the outputs for the best-performing models are organized with the functionality provided by the *_18_3_best_predictions_sorting.py* script. It generates the *bests* folder in *05_predictions* in the *graphs* main folder (to store only figures), as well as the working folder of *predictions_best* in the *others* main folder. The *predictions_best* folder includes all the generated graphs (similar to those shown in **Figure 18**), the report, as well as the corresponding csv files with all the outputs. The same report named *07_prediction_summary.txt* is stored in the reports folder; it provides the final comparison of the performed energy optimization (see example #12 in the subsection 7.6).

Finally, the CC-reevaluation is available via the functionalities provided in the *_19_cc_reeval.py* script. The process is performed in the same manner as presented in Module 4 (see more in the subsection 5.5). Outputs are different because this time the indoor climate data is taken from the best-performing optimization models: the examined records include the assumed T_i and RH_i offsets. Running the CC-reevaluation (see subsection 4.1) is a recommended part of the energy optimization: it provides a confirmation that application of the proposed offset strategy will maintain the indoor climate under given CC rigor. It provides the graphical (psychrometric and scatter plots)and report assessment, as well as corresponding csv data file for each of the examined variants (e.g., for the given demand model, different operation schedules and temperature offsets are saved in individual folders). All the subfolders are stored in the *08_CC_ReEval* folder in the *reports* main folder.

5.7. Additional scripts

For the current version of the CEAM software, the only additional script is included with the shared tool, as the sub-software called TIG (Time Index Generator). It is a simple tool allowing for time index generation in a csv format. The generated file has a header *Time*, with a single column only. The consecutive rows consist of the time indexes generated with the pre-defined format, with the selected frequency in minutes. The functionality is provided in the *_00_time_index_generator.py* script. A short overview of the sub-software operation within the CEAM is shown in the subsection 4.1, the definition process is presented in **Figure 8**, and the exemplary output is shown in example #7 in the subsection 7.6. This functionality assumes the default inputs if they are not provided:

- the directory is set as the desktop path,
- the frequency is set to 60 minutes,
- starting date is set as *2025-01-01 00:00*, ending date as *2025-04-30 25:59*,
- the data format is set as the default for CEAM (i.e., '*%d.%m.%Y %H:%M*').

6. RESULTS: A BRIEF OVERVIEW

The CEAM software enables indoor climate assessment and energy consumption optimization through short-term strategies involving different offset setpoints. It quantifies potential energy savings using a BBM approach: an approximator that predicts energy demand based on ASHRAE's recommendations.

In this section, a closer look at the generated outputs of the assessment performed by means of the CAEM software is shown. The exemplary overview of the performed analyses, along with guidance on interpretation, is shown in the subsections 7.2-7.5. This section presents what outputs are generated, how they are stored, and what their applications are. This brief overview is made based on the complex analysis, assuming all the possible parameters are included and all the software features are used.

A correct set of input files is obligatory for the computing (see section 3). To obtain the analysis outputs, the correct fulfillment of all the required inputs via the software GUI is required, as explained in the section 4. With minimal input data, CEAM provides an in-depth assessment of the museum's existing control approach, revealing opportunities for optimization.

CEAM generates three types of outputs: csv files, text reports, and graphics. The csv files typically serve as temporary input/output repositories (used for the specific calculations). We can distinguish csv files used as supplementary inputs (the basis for the performed assessment), as well as the ones used only for temporary support for calculations. All the generated files are saved and can be accessed by users. Detailed reporting for each side analysis is stored in the generated text files, while key results are visualized through various types of graphs. The generated reporting is linked to the module selection: the dedicated files are created for each type of analysis. The number of generated plots is based on the provided input and features selected. Again, the type of generated graphs is linked to the module selection. This structure ensures free access to all generated resources, enabling users to utilize the outputs according to their specific needs outside the CEAM software.

CEAM provides organized file management for all outputs. After finishing complete computing, the main folder contains:

- 4 main folders: *graphs*, *others*, *raw_data*, and *reports*,
- 5 key csv data files: *h_comb.csv*, *h_comb_CC.csv*, *d_comb.csv*, *w_comb.csv*, and *m_comb.csv*,
- 2 text files: *terminal_records.txt* and *var.txt*.

The key data files are used throughout the computing in all the modules of the CEAM software. These files contain the key variables for the overview of the historical records. The available text files contain supplementary information regarding the performed computing: the variables input, as well as terminal outputs (with a valuable source of information for debugging). These files have rather limited utility for the users of CEAM software, yet might be very helpful for further development and improvement of the software.

5 folders are generated in the *graphs* folder, each linked to the module executed: *01_input overview*, *02_correlation*, *03_energy signature*, *04_climate class overview*, and *05_predictions*.

The first one (*01_input overview*), generated after successful computing of Module 1, contains 4 subfolders, managing the graphs based on the frequency of data: *01_hourly*, *02_daily*, *03_monthly*, and *04_extreme* (the selection of the extreme weeks is explained in the subsection 5.1). For the first 3 subfolders, up to 27 graphs can be generated (example shown in **Figure 14**). The naming of each graph contains the frequency prefix, as well as the variable symbol (see **Table 8** in the subsection 7.8); e.g., *h_AbsHe.png* presents the hourly evolution (*h*) of external absolute humidity (*AbsHe*). The *04_extreme* folder should contain 78 plots, using the symbol of the examined type of the extreme week as a prefix. The utility of the graphs stored in the folder is purely informative: it provides an overview of the given parameter. It can be used to pinpoint moments of increased energy consumption or system malfunctioning.

The second folder (*02_correlation*) contains 4 subfolders with the graphs generated after the successful execution of Module 2. The first 3 subfolders store the outputs based on the frequency of data: *01_hourly*, *02_daily*, and *03_monthly*, while the 4th one (*04_selected*) contains the selected graphs with the best correlation. These outputs indicate that the two presented parameters are correlated, e.g., the *Cor_d_AbsHe_D4.png* graph indicates a high correlation between daily records of exterior absolute humidity and energy consumption for humidification. The best performing correlations are selected with the R^2 [67] and r[68] values. The outputs stored can help the user identify the key parameters affecting energy consumption. An exemplary correlation graph is shown in **Figure 15**.

The third folder (*03_energy signature*) contains 3 subfolders with the graphs generated after the successful execution of Module 3. The first 2 subfolders store the outputs based on the frequency of data: *01_hourly* and *02_daily*, while the 3rd one (*03_selected*) contains the selected graphs with the most accurate demand curve. Energy signature is an evaluation method in which energy consumption is correlated with climatic variables aimed at representing the actual energy behavior of the building, as well as estimating baseload (not-weather-dependent energy consumption) and base temperature (see more in the subsection 5.4). The effectiveness of the given model is evaluated with the R^2 [67] value for the given model equation. As mentioned above, these outputs can be used to highlight the baseload, as well as the base temperature, the value after which increased energy consumption was recorded. An exemplary energy signature graph is shown in **Figure 16**.

The fourth folder (*04_climate class overview*) contains 2 subfolders with the graphs generated after the successful execution of Module 4. Both folders include the graphical results for the CC assessment of the examined case study. It includes the scatter plots in the *01_overview* folder and psychrometric charts in the *02_psychrometric* folder (see **Figure 17**). The generated plots are named with the examined CC, as well as, for scatter plots, the presented parameter: T or RH. Moreover, in the *01_overview* folder, two subfolders with daily (*MA_daily*) and hourly (*MA_hourly*) CC assessment with moving average[85] are given.

The fifth folder (*05_predictions*) contains up to 6 subfolders with the graphs generated after the successful execution of Module 5. Up to 5 subfolders named with the operation schedules symbol (e.g., *s1*) can be generated, based on the selection in the GUI. In these subfolders, all graphical prediction model outputs (see **Figure 18**) are shown: for each of the examined

[85] see more: https://en.wikipedia.org/wiki/Moving_average

demands, for each prediction variables. Up to 24 graphs can be generated for each operation schedule. The last folder named bests contains only the best performing prediction models for each demand and each operation schedule; e.g., *D1_deltaT_DT_predictions_s1.png* means that for *D1* demand predictions, the DT model with *deltaT* variable is selected for operation schedule *s1*.

The *reports* main folder includes the text-based outputs and one folder named *08_CC_ReEval*. These outputs are generated from each module.

Reports *01_missings_report*, *02_stats_report*, and *03_extremes_report* are generated after the successful computation of Module 0. These 3 reports include the base overview of the provided input data; these data can be matched with the graphical outputs from Module 1.

Both reports *04_correlation_report* and *04_correlation_report_selected* are the outputs from Module 2. The first one includes information on all the examined functions, while the second one gives a summary of the best-performing ones: 5 for each data frequency.

Both reports *05_ES_report* and *05_ES_report_selected* are the outputs from Module 3. The first one includes information on all the examined models, while the second one gives a summary of the best-performing ones: 3 for daily frequencies and the 3 corresponding hourly frequencies.

The *06_CC_report* is the output from Module 4, showing how well each of the CC is fulfilled (see example #9 in the subsection 7.6). This report should be used with the generated graphics showing the CC assessment.

Finally, *07_best_predictions*, *07_prediction_summary*, and *07_predictions_overview* reports are generated after the successful computation of Module 5. The last one includes information on all the examined prediction models (see example #10 in the subsection 7.6), while the first one includes the best-performing model for each demand and variable (see example #11 in the subsection 7.6). The *07_prediction_summary* includes information on only the very best model for each demand, and the recorded savings due to the application of short-term strategies of indoor climate management (see example #12 in the subsection 7.6).

The *08_CC_ReEval* subfolder is generated only if the CC reevaluation was selected in the GUI. The folder includes numerous subfolders presenting the CC assessment for each of the examined models. The contents of each subfolder are the temporary csv file, the CC report file, and two folders for graphical outputs (similar to the one generated in Module 4). These reports can be used to check if the proposed strategy is fulfilling the requirements of the desired CC.

The *others* folder contains all the temporary generated outputs. It includes up to 5 subfolders and 21 temporary csv files. The *old* subfolder stores some of the temporary csv files used in past computing. The *to_use* subfolder contains csv files with the overview of the extreme weeks. The other 3 subfolders are linked with the energy predictions, particularly *predictions_best*, *predictions_outputs*, and *predictions_selected*. The first one contains all types of files (csv, text, and graphs) regarding the very best-performing model for a given demand (one model per demand). The second one contains up to 6 subfolders with the results for a given demand (*D5* is designated as *D5a* and *D5b*, referring to strategies based on offsets for T_i only or simultaneous T_i and RH_i). The third subfolder consists of text reports and csv files for the best-performing models for each demand and the given parameter. Most of the remaining csv files stored in the *others* folders do not require further explanation. The most interesting file is named *prediction_inputs_all.csv,* and it includes all the inputs used during the optimization process performed in Module 5.

The last folder is named *raw_data*, and it stores the original input files, copied before any calculation is performed.

The computing can take even a few hours, depending on the selected scope of assessment, but the single module can be computed as fast as dozens of seconds. User should keep an eye on the computing status presented on the software GUI.

Mistakes can happen, and developer errors are also possible. The *terminal_records.txt* file generated during computation may be helpful in the event of a failure. The file contains printed terminal records of the running Python scripts: sometimes, even without programming knowledge, it might be possible to capture the mistake/error. If users are unable to resolve the issue on their own, they are encouraged to contact the developer via **contact email**: **climate2preserv@gmail.com**. The message should include a brief description of the intended assessment, along with the raw input files and the *terminal_records.txt* file.

7. APPENDICES

In this section, all the appendices of the provided manuals are presented. In subsection 7.1 the quick start user manuals are included; this document is also available separately. Subsections 7.2-7.5 consist of exemplary applications of CEAM software for energy and climate optimization. For all the exemplary applications, the most complex assessment was performed. All the exemplary assessments are performed using the artificially generated input files (i.e., ED and ICD files) by means of *EnergyPlus* software[86]. Keep in mind that the obtained results will slightly vary when re-run. Subsection 7.6 includes all the examples with supplementary explanations mentioned throughout this document.

7.1. QUICK START USER MANUALS

CEAM OVERVIEW

The **C**limate and **E**nergy **A**ssessment for **M**useums (**CEAM**) is an open-source, Python-based decision-support tool designed to help federal institutions and heritage building owners improve energy efficiency while preserving cultural assets. CEAM provides a clear, data-driven overview of energy use and indoor climate conditions, using raw input data to generate useful outputs. Its modular, user-friendly structure enables comprehensive, multicriteria analysis of heritage preservation, indoor climate management, and energy-saving opportunities. The tool employs AI-based optimization methods to evaluate potential energy savings through short-term management strategies, with predicted savings ranging from 10-50% depending on the scenario and selected interventions.

CEAM is meant only as a decision-support tool. The final decision to implement any climate control strategy should be made by the responsible personnel, who must consider the best interests of the housed collection.

This document provides compact information on the quick start for the CEAM software. All the details on required inputs, definition process, capabilities, applications, and limitations of the tool, as well as how to review the received results, can be found in the CEAM manuals available here. The software is available here as a stand-alone Windows app, as well as raw scripts, in the GitHub repository.

CEAM was developed by Marcin Zygmunt as a part of the Climate2Preserv (C2P) project, founded by the Belgian Science Policy Office. The C2P project is the combined effort of project partners: KU Leuven, KIK-IRPA, and ULiège.

Please address all the comments related to the CEAM software via the contact email: **climate2preserv@gmail.com**.

GETTING STARTED WITH CEAM

CEAM software can be downloaded from here as a ready-to-use executable file (no need to install any additional software if the updated version of Windows is used). It is also suggested to run the software via the '*run as administrator*' option. The GitHub repository is given for

---

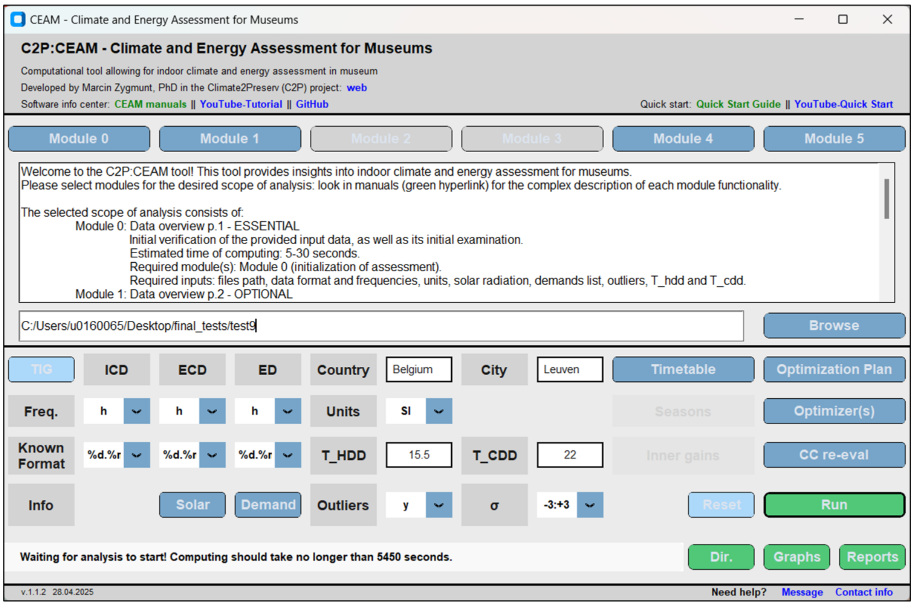
[86] available here: https://energyplus.net/

users who prefer to run the tool from a Python environment (additional libraries and packages are obligatory for a successful operation). The tutorials on how to use the CEAM software are given on the [YouTube channel](#).

The tool has no specific requirements: once downloaded, only the input files are required. Three main input files in the csv format are necessary, particularly the *Exterior Climate Data* (ECD.csv), *Indoor Climate Data* (ICD.csv), and *Energy Data* (ED.csv). A certain input data structure is required: some exemplary input files are available [here](#), with a full introduction in the manuals. Input files with hourly frequencies of records are recommended, yet CEAM also supports daily and monthly granularities. The goal is the use your own data input, and thus a precise definition process is recommended.

RUNNING CEAM

CEAM provides a user-friendly GUI for easy definition and computation. The exemplary filled-in GUI is shown below.



CEAM GUI with exemplary inputs.

The scope of computing is adjustable by the user: the required modules should be selected at the top of the window. When the module is selected, the corresponding description is also shown, with the required inputs to be provided.

The so-called working path should be selected, pointing to the folder containing three input files. Next, all the required information should be provided in the given textboxes, selected from lists, and defined in the additional windows. When all the necessary information is provided, the simulation can be performed by pressing the Run button.

The computing time is dependent on the selected scope of the assessment: a complete simulation can take as long as three hours. The predicted time computing is given before running the simulation, as well as a supporting pop-up window is visualized when calculations are in progress. Users are informed when computing is finished.

All the outputs are stored in the directory path selected before; the original input files are safely copied. The outputs consist of various valuable outputs, including input data overview, the climate and energy consumption evaluations, as well as energy optimization. A proper, predefined file management is provided. Some exemplary results from a successful analysis are available **here**.

A correct understanding of the obtained results is key to a successful assessment of the examined case study. Some active links to the helpful repositories (e.g., manuals or contact info) are given with hyperlinks in the upper and lower parts of the GUI. Users are also welcome to share their feedback.

7.2. EXEMPLARY APPLICATION #1: A COMPLETE INPUT DATA

The exemplary application of the CEAM software is presented using a complete set of input data for a whole calendar year: 8760 hourly records from 1st January to 31st December. It is assumed that the year 2023 is examined. The ICD and ED files are artificially generated, while ECD consists of data from the TMY file for Brussels[87].

For the purpose of a successful analysis, the correct filling out of the software GUI is required. The working directory (with all input files in) was provided and all modules were activated. For all input files hourly frequency was selected, with *%d.%m.%Y* date format. Solar radiation is included in the analysis, as well as *D1*, *D2*, *D4*, and *D5* demands were selected. Belgium (designated as *Be*) was input as a country, and the SI metric system was selected. Heating and cooling degree days temperatures were set as 15.5 °C and 21.5°C respectively. The outliers assessment was used, assuming valid inputs between the standard deviation range of ±3σ. The operation schedule was assumed to mimic the working hours of a typical museum. The optimization schedules *s1* and *s5* were selected, while all optimizers (i.e., DT, NN, LSTM, and GRU) were selected. Finally, CC-reevaluation was selected. All the inputs made are shown in **Figure 19**.

From the input data overview performed in Module 0, we see that the complete set of inputs (8760 records for each input variable) is examined (see the *01_missings_report.txt* file). The outliers assessment affected all the input files, decreasing the number of records down to 8520 records at most (97.26% of the original dataset). The extreme periods were set (see *03_extremes_report.txt*), while the examined data can be characterized with a summary for each parameter (see *02_stats_report.txt*):

- column T_e: max of 31.1°C, min of -5.6°C, and the average of 11.15°C,
- column RH_e: max of 100%, min of 31%, and the average of 78.15%,
- column I_s: max of 824 W/m^2 and the total of 1264.25 kWh/m^2,
- column *HDD*: total of 1974.22°C·days,
- column *CDD*: total of 61.77°C·days,
- column T_i: max of 21.99°C, min of 20.07°C, and the average of 21.03°C,
- column RH_i: max of 61.35%, min of 47.30%, and the average of 50.93%,
- column *D1*: max of 7950.65 Wh, min of 0.0 Wh, average of 2825.94 Wh, and total of 24755.22 kWh,

[87] the used file contains hourly records for a representative year, compiled from measurements taken between 2009 and 2023; the used TMY file is available here: https://climate.onebuilding.org/

- column *D2*: max of 2687.43 Wh, min of 0.0 Wh, average of 679.07 Wh, and total of 5948.63 kWh,
- column *D4*: max of 8798.07 Wh, min of 5.0 Wh, average of 2161.65 Wh, and total of 19080.12 kWh,
- column *D5*: max of 15329.54 Wh, min of 5.0 Wh, average of 5676.25 Wh, and total of 49783.97 kWh.

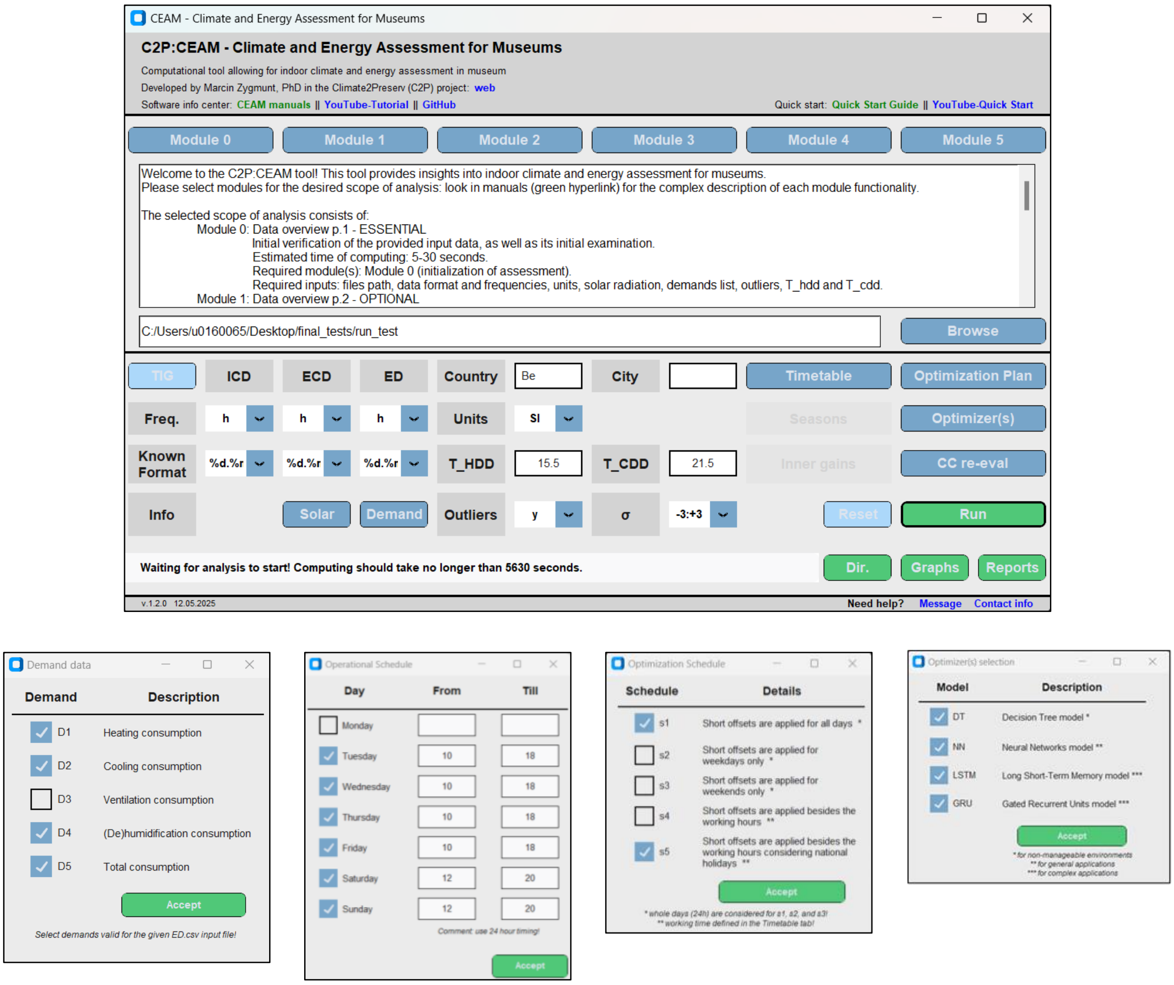


Figure 19. All the input provided into the GUI: main windows (upper part), as well as demand data (far left bottom), operational schedule (middle left bottom), optimization schedule (middle right bottom), and optimizers selection (far right bottom) windows.

Some of the key figures presenting the input data are shown below, examining outputs from Module 1 with hourly frequency. **Figure 20** shows the exterior environment showing a regular evolution of T for localization with 4 seasons, recorded behavior of RH, as well as the presence of Is (denser for the summer period). In **Figure 21** we can see the HDD and CDD throughout the year; it is expected that the examined localization is heating-dominant.

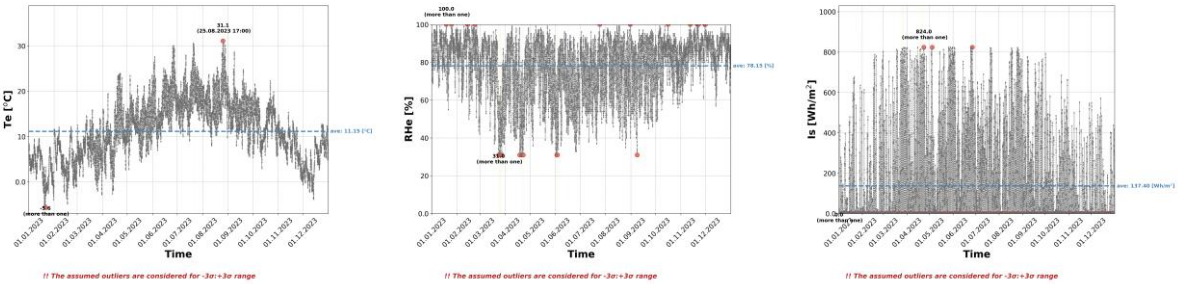


Figure 20. Graphical overview of the exterior climate: the evolution of T_e (on the left), RH_e (in the middle), and I_s (on the right).

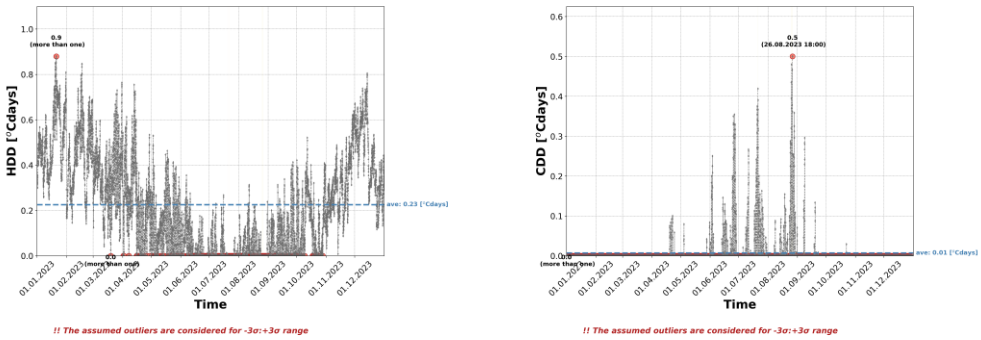


Figure 21. Graphical overview of the exterior climate: the evolution of *HDD* (on the left) and *CDD* (on the right).

The indoor climate can be overviewed with the T_i and RH_i graphs, as shown in **Figure 22**. It can be seen that indoor climate is tightly controlled, especially *T*-wise, as well as keeping the lower bound of RH_i.

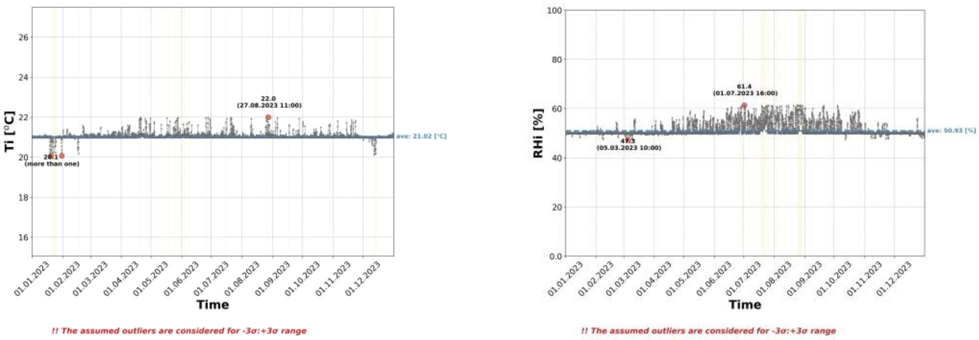


Figure 22. Graphical overview of the interior climate: the evolution of T_i (on the left) and RH_i (on the right).

Finally, the plots of the examined demands are shown in **Figure 23**. As expected, the heating consumption (*D1*) is meaningful, as it occurs throughout the whole year, which is caused due to the necessity of humidification, which uses some part of the heating consumption as well. The peak demand for cooling (*D2*) took place during summer, while some lower consumption is also recorded in other periods, used, e.g., for post-processing of air after adjusting the humidity. The energy consumption used for humidification (*D4*) is mostly linked with the absolute humidity content throughout the year (no significant inner moisture sources for the examined case). Finally, the total consumption (*D5*) represents the sum of all the previously mentioned demands. The northwardly high consumption was recorded during winter, increased one during spring and autumn, while demand similar to the average consumption was recorded during summer.

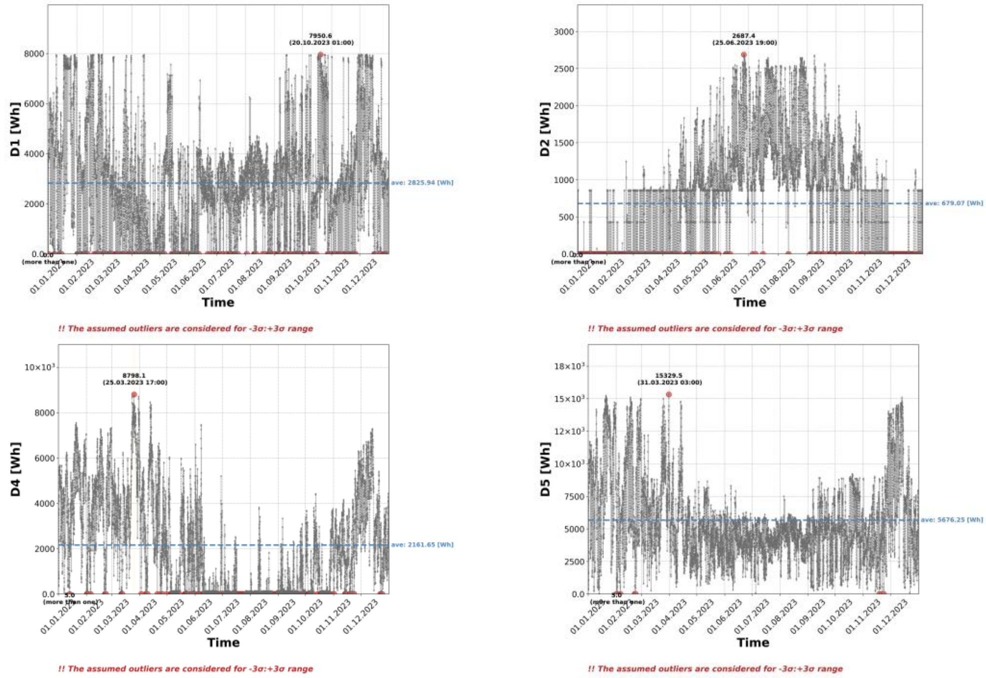


Figure 23. Graphical overview of the examined demands: *D1* (heating, on the upper left) and *D2* (cooling, on the upper right), *D4* (humidification, on the lower left), and *D5* (total, on the lower right).

Outputs from Module 2 provide the correlation assessment of the available variables. According to the *04_correlation_report_selected.txt* file, the strongest correlation was obtained between *deltaTd* and *D4* (see **Figure 24**), for both hourly and daily records (even though monthly correlations are available, they are considered less reliable due to the limited number of data points). This output suggests that the required energy consumption to humidify the incoming air (i.e., the amount of moisture it holds) is highly related. All other correlations can be found in the *04_correlation_report.txt* file. For example, the correlation between T_e and *D1* is, contrary to initial expectations, rather poor (R^2 equals 0.14), but it is much more relevant for T_e and *D2* (R^2 equals 0.72).

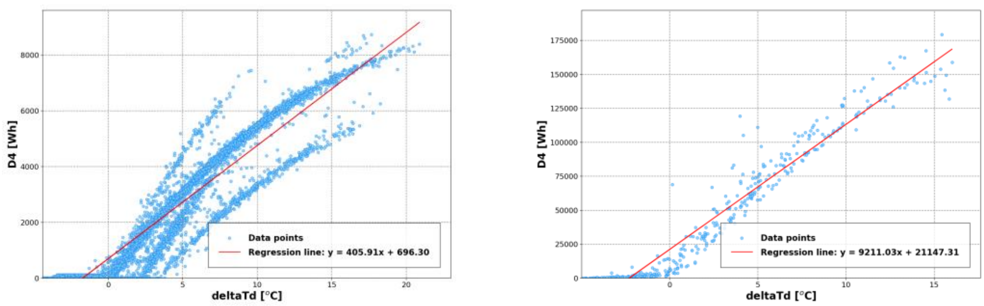


Figure 24. Graphical overview of the best-performing correlation obtained for *deltaTd* and *D2*: for hourly (on the left) and daily (on the right) records.

That type of overview can help with a better understanding of the examined case and the systems used, as well as the zone/building response to various factors (e.g., cooling demand is mostly dictated by the external temperature). Yet, understanding the correlation assessment requires knowledge and some practice. To help with this process, two statistical measures are included for each of the examined correlations: R^2 and r (see explanations in the subsections 5.3 and 7.7).

Outputs from Module 3 provide an overview of energy consumption using the ES method. Based on the information in *05_ES_report_selected.txt* file, the three best-performing models for daily and hourly frequencies are selected based on R^2 values. All the models selected focus on cooling demand (*D2*), aligning with the correlation assessment obtained in Module 2. This confirms that cooling demand is strongly correlated with the examined factors for the studied building. The provided simple methods for analytical evaluation of the data (Modules 2 and 3) offer valuable information on cooling demand management. Correlation assessment highlights a strong dependency between T_e and *D2* dependencies, which is further explored using ES model outputs. As shown in **Figure 25**, the hourly analysis reveals a baseload is 43.86 Wh and a change-point temperature of 7.13°C. The daily plot (recommended for ES assessment) shows a baseload of 2753.03 Wh, and a change-point temperature equals 8.55°C. Using ES outputs building manager can estimate the cooling demand concerning the exterior parameter (e.g., T_e). It is also available for users to define a predicted baseload, as well as to set a more precise overview strategy of the cooling demand with the temperature above the change-point. In *05_ES_report.txt* file, all the examined scenarios for ES assessment are shown and evaluated. As previously discussed for outputs from Module 2, the alignment of heating demand (*D1*) with T_e is not efficient, even for the daily records (R^2 equals 0.621). Analogous to the considerations from outputs in Module 2, the examination of *D1* is less conclusive for the examined case study.

An overview of the results obtained from Modules 2 and 3 can be a helpful approach for a better understanding of the examined case and the systems used. However, interpreting correlation and regression assessments requires some background knowledge and practice.

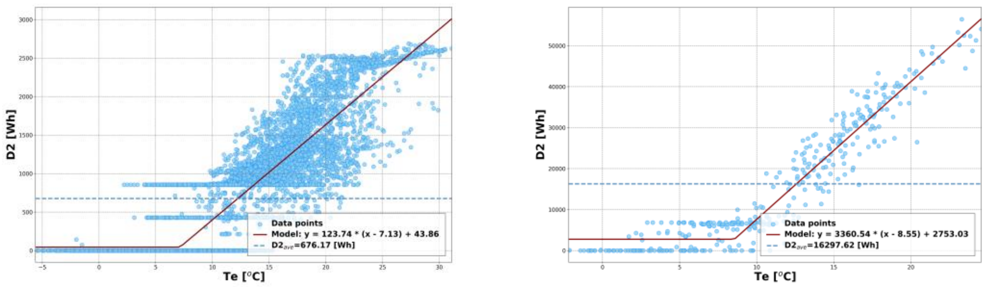


Figure 25. Graphical overview of one of the best-performing ES assessments obtained for *Te* and *D2*: for hourly (on the left) and daily (on the right) records.

Outputs from Module 4 provide an overview of the indoor climate conditions following the requirements for ASHRAE CC. The results are summarized in the *06_CC_report.txt* file, as presented in example #9 in subsection 7.6. The report file includes short information regarding the examined input data, along with the CC assessment for all evaluated classes. The evaluation is given for a common approach, as well as using daily and hourly moving averages[85] (for the short-term criteria).

The high-resolution graphical outputs (see **Figure 26**) are also generated out of Module 4: all of them are stored in the *graphs* main folder, in *04_climate class overview* subfolder. Inside, two additional directories are given, particularly *01_overview* and *02_psychrometric*. In the first one, the overview by means of the traditional scatter plots is given, individually for T and RH for each of the examined CC. In the second folder, CC assessments using psychrometric charts are presented. Psychrometric charts are a great way to evaluate whether the historical records of indoor climate are within the T and RH requirements. The scatter plots give an overview of the time evolution of the indoor climate conditions and thus allow us to examine the selected period. The corresponding CC restrains are shown on each graph, as well as the datapoints that do not fulfill the requirements are highlighted.

Users can compare the numerical results given in the report file and examine them with the graphical outputs. Scatter plots show all the CC restraints (i.e., long-term, seasonal, and short-term), while psychrometric charts highlight long-term and seasonal bounds. In order to know how well requirements for a given CC are fulfilled user should check the first lines in the report file for that class, in particular the result for long-term, seasonal-term, and short-term fulfilments. The final result for a given CC is the lowest of these three values, indicating the share of time the requirements were fulfilled (the time when it was not fulfilled can be easily spotted with scatter plots).

It is strongly recommended for the users to be familiar with ASHRAE Chapter 24[40] to fully understand the results of Module 4. Knowledge on how to read the psychrometric charts is also recommended. CEAM authors are working on the article evaluating the indoor climate assessment based on ASHRAE recommendations[88].

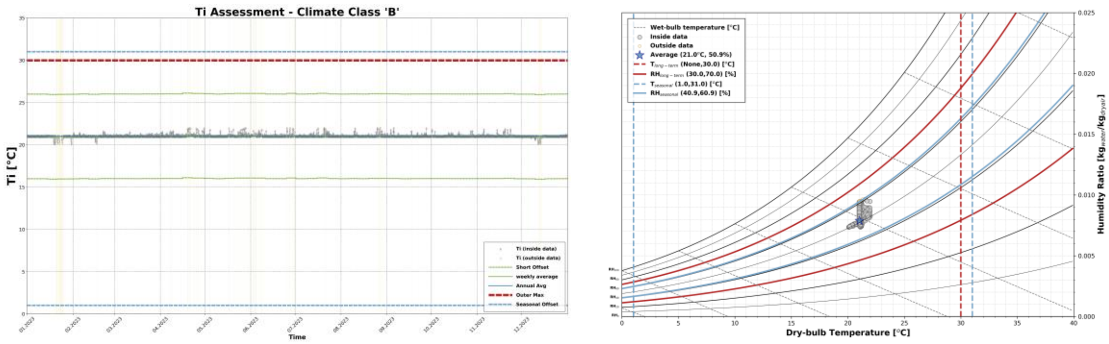


Figure 26. Graphical overview of the CC assessment: exemplary scatter plot (on the left) and the psychrometric chart (on the right).

Outputs from Module 5 are highly dependent on the made definition; in this example, the complete set of results is overviewed. This module is used to evaluate the potential energy savings after applying the acceptable T and RH offsets (see more in subsection 5.6).

Three report files are generated and stored in the *reports* main folder. The *07_predictions_overview.txt* includes the complete overview of all the performed predictions, with various models used. This overview includes the list of assessments of the performed analysis, total results, as well as performance and accuracy assessment. The *07_best_predictions.txt* includes the overview of the best prediction models for each of the examined variables. This report includes only a compact overview of the selected models. The *07_prediction_summary.txt* includes an overview of the expected energy savings when a given

[88] further information regarding this publication will be shared shortly after publishing

offset strategy is applied. The comparison is shown only for the best-performing prediction model (one for each demand). The report includes the base information on the default consumption, as well as the possibility to achieve savings in terms of absolute and relative differences.

Graphical outputs are stored in the *05_predicions* subfolder, in the main *graphs* folder. Inside, additional subfolders are created, allowing for easier orientation within generated outputs. The generated subfolders represent the selected optimization schedules (*s1* and *s5* in this example) as well as the *bests* subfolder. In the *bests* subfolder, graphs for only the best-performing prediction models are shown: for each of the examined demands, all the examined optimization scenarios are distinguished. Each graph consists of three subplots (see **Figure 27**): the upper one showing the full examined period, the distribution for a week with the highest consumption (lower left), as well as the overview of a day with peak demand (lower right).

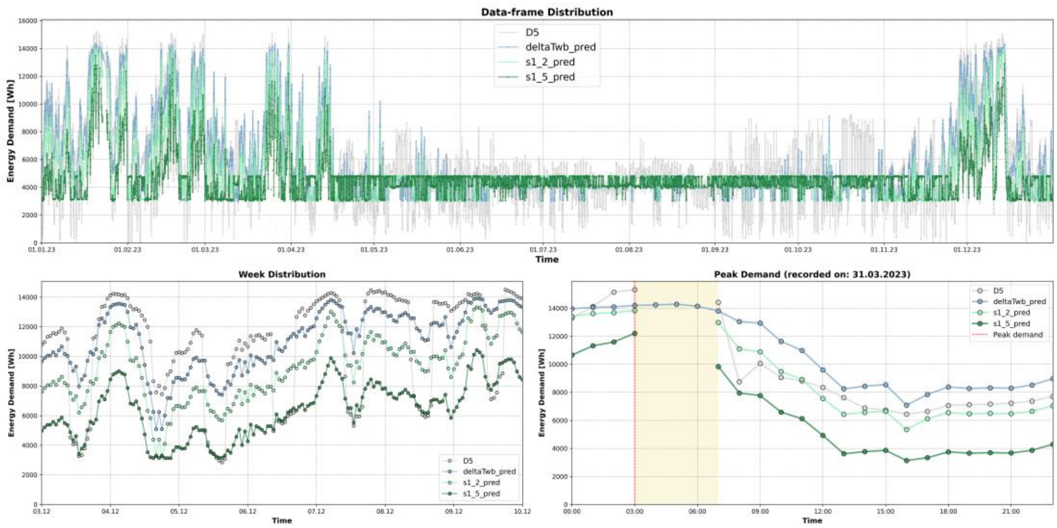


Figure 27. Energy optimization outputs using graphs: energy demand predictions for the examined variant of the total demand.

How to understand the naming of the examined scenarios is explained in subsection 5.6 and in example #12 in subsection 7.6; information provided in **Table 8** is also helpful (see subsection 7.8). The presented graph is named *D5a_deltaTwb_NN_predictions_s1.png*: it examines the total energy consumption (*D5*), considering only T-related strategies (designated with *a*). The selected prediction model is based on the NN application, based on the *deltaTwb* variable. The results assume the *s1* optimization schedule.

The plots consist of real consumption (*D5*, in grey), the base consumption predicted by the trained model (*deltaTwb_pred*, in blue), as well as two energy saving variants, designated as *s1_2_pred* (in light green) and *s1_5_pred* (in dark green) accordingly, allowing for T-based control by ±2°C or ±5°C. The less severe strategy is always marked in lighter green. The information about the used model can be seen in the *07_best_predictions.txt* file. The selected model is characterized by a R^2 value of 0.6365 (indicating decent accuracy) and a CCC of 0.8050 (indicating good agreement). The peak demand is predicted very well, with only a 7.4% difference. The graphs show the corresponding variants summarized in the

07_prediction_summary.txt file. This variant allows to save 10.58% (about 5091 kWh) or 21.00% (about 10105 kWh) for offsets of ±2°C and ±5°C respectively.

In other subfolders, the created graphs show the predictions for all the examined models, for all examined demands (grouped by the optimization scenarios). For a numeric result of the proposed strategy user should use the report, while graphs provide information on the predicted demand behavior.

Please note that the energy optimization module of CEAM is meant only as a decision-support tool. The predicted savings are based exclusively on the input data provided and may be affected by other variables not accounted for in the analysis. However, the energy prediction is performed conservatively, so the expected savings from the proposed short-term offset should be achievable. **The final decision to implement any climate control strategy should be made by the responsible personnel, who must consider the best interests of the housed collection**.

Additionally, the *08_CC_ReEval* subfolder is also generated (if CC re-evaluation is selected) and stored in the main folder. This subfolder includes the CC assessment similar to the one performed in Module 4, yet evaluated based on the modified indoor climate data, following the energy-saving strategies for T and RH offsets. For each of the examined scenarios, a similar CC assessment report is provided, with corresponding scatter plots and psychrometric charts.

The total computing time was about 140 minutes, with 1.62GB of outputs generated. The used input files, as well as all the received outputs, are shared in the *Application_1* subfolder, in the *exemplary application* folder in the GitHub repository[3].

7.3. EXEMPLARY APPLICATION #2: AN INCOMPLETE INPUT DATA

This assessment is performed in the exact same manner as the one presented in the subsection 7.2. The difference is in the input data: all of the input files are incomplete. The incompetence, particularly the empty span or lack of parameters, of the data is random.

The most visible difference is due to the missing values in the input data (see the given *01_missings_report.txt* report): the obtained results vary accordingly, as well as the empty data frames are visible on the generated graphics. Yet, the interpretation of the obtained results is the same as it was explained for the complete input data (subsection 7.2).

The total computing time was about 140 minutes, with 1.56GB of outputs generated. The used input files, as well as all the received outputs, are shared in the *Application_2* subfolder, in the *exemplary application* folder in the GitHub repository[3].

7.4. EXEMPLARY APPLICATION #3: DAILY INPUT DATA

This assessment is performed in the exact same manner as the one presented in the subsection 7.2. The difference is in the input data: the ED file consists of daily records. Thus, the necessary change in definition for ED data is required (see more in the subsection 4.1), particularly the frequency of records (daily) and known format (*%d.%m.%Y* in this example).

The obtained differences are caused due to the process of handling daily ED input: the original data is adjusted to the hourly records based on the HDD/CDD values. As mentioned in the subsection 5.1, the demand adjustment for inputs of lower frequencies than hourly is performed based on the HDD and CDD values. The *D1* records are adjusted according to HDD, while D2 is adjusted according to CDD. Other demands should not be examined (i.e., included in the GUI definition, as shown in subsection **4.1**) if daily or monthly records are provided in the ED file. Keeping the other demands on will unnecessarily increase the computing time, as well as might cause a critical error for the analysis[89]. Due to the performed assumption, the outputted hourly *D1* and *D2* might vary with the original data; for all zeroth HDD/CDD records, no consumption is assumed even if consumption was recorded in the raw file. In this example, the outputted *D1* is 23779.16 kWh and *D2* is 2667.54 kWh, compared with the original sum of 24755.22 kWh and 5948.63 kWh, respectively (with outlier valuation included). The interpretation of the obtained results is the same as it was explained for the complete input data (subsection **7.2**).

The total computing time was about 80 minutes, with 640MB of outputs generated. The used input files, as well as all the received outputs, are shared in the *Application_3* subfolder, in the *exemplary application* folder in the GitHub repository[3].

7.5. Exemplary application #4: monthly input data

This assessment is performed in the exact same manner as the one presented in the subsection **7.2**. The difference is in the input data: the ED file consists of monthly records. Thus, the necessary change in definition for ED data is required (see more in the subsection **4.1**), particularly the frequency of records (monthly) and known format (*%m.%Y* in this example).

The obtained differences are caused due to the process of handling daily ED input: the original data is adjusted to the hourly records following the procedure explained in the subsection **7.4**. It is important to remind that only *D1* and *D2* demands should be selected during the definition process (even if other demands are available in the ED file). In this example, due to the necessary adjustments of the ED file, the outputted *D1* is 24755.22 kWh and *D2* is 5456.32 kWh, compared with the original sum of 24755.22 kWh and 5948.63 kWh, respectively (with outlier valuation included). The interpretation of the obtained results is the same as it was explained for the complete input data (subsection **7.2**).

The total computing time was about 80 minutes, with 640MB of outputs generated. The used input files, as well as all the received outputs, are shared in the *Application_4* subfolder, in the *exemplary application* folder in the GitHub repository[3].

7.6. Examples

In this subsection, all the examples mentioned throughout the document are presented. The examples are included for a better understanding of the developed CEAM software and to simplify the usage process of the tool.

Example #1

The input of:

```
Time,d1,D3,d4,d5,D2
01.01.2024 00:00,10,30,40,50,20
```

[89] it is planned to include some additional user-friendly safety feature into the software to prevent unwanted errors

```
01.01.2024 01:00,100,300,400,500,200
…
31.12.2024 23:00,1000,3000,4000,5000,2000
```

will be restructured as:

```
Time,D1,D2,D3,D4,D5
01.01.2024 00:00,10,20,30,40,50
01.01.2024 01:00,100,200,300,400,500
…
31.12.2024 23:00,1000,2000,3000,4000,5000
```

Thus, the tool will restructure the headers (font and positioning), as well as reorganize the given records for each column to correspond with the default template. The following modifications were made:

- verifying the font of the headers and adjusting it (e.g. from *d1* to *D1*),
- reorganization of the columns of records to follow the default structure.

Example #2

The input of:

```
Time;D1;D2;D3;D4;D5
2024/01/01   00;10;0;;;
2024/01/01   01;100;0;;;
…
2024/12/31   23;1000;0;;;
```

will be restructured as:

```
Time,D1,D2,D3,D4,D5
01.01.2024 00:00,10,0,,,
01.01.2024 01:00,100,0,,,
…
31.12.2024 23:00,1000,0,,,
```

Thus, the tool will change the used separator (from semicolon to comma) and adjust the *Time* column to correspond with the default template. The following modifications were made:

- swapping components separator from a forward slash (/) to a dot (.),
- changing the date format (from *yyyy.mm.dd* to *dd.mm.yyyy*),
- changing the time format (from *HH* to *HH:MM*),
- adjust the gap between date and time (from three to one white space).

Example #3

The input for the ED file consists of *D1*, *D2*, and *D4* consumptions. Users can define the *ED.csv* file as:

```
Time,D1,D2,D3,D4,D5
01.01.2024 00:00,1000,0,,200,
01.01.2024 01:00,1000,10,,200,
…
31.12.2024 23:00,1000,0,,200,
```

or with an additional attachment of *D5*, as a sum of the preceding demands, as follows:

Time,D1,D2,D3,D4,D5
01.01.2024 00:00,1000,0,,200,1200
01.01.2024 01:00,1000,10,,200,1210
…
31.12.2024 23:00,1000,0,,200,1200

Note! The energy consumption should be from one fuel (e.g., electricity); otherwise, combining demands is not possible.

Example #4

The user has records of heating (*D1*) and the total (*D5*) consumptions, yet there is no concrete knowledge of allocates from total demands. Users can use both of the records, yet the assessment based on *D5* might be inaccurate (no information on constituents).

Time,D1,D2,D3,D4,D5
01.01.2024 00:00,1000,,,,1200
01.01.2024 01:00,1000,,,,1210
…
31.12.2024 23:00,1000,,,,1200

In that case, the recommended approach is to examine the heating consumption (*D1*) only, with potential *D5* assessment in the future (when more comprehensive knowledge on the energy consumption is gained):

Time,D1,D2,D3,D4,D5
01.01.2024 00:00,1000,0,,,
01.01.2024 01:00,1000,,,,
…
31.12.2024 23:00,1000,0,,,

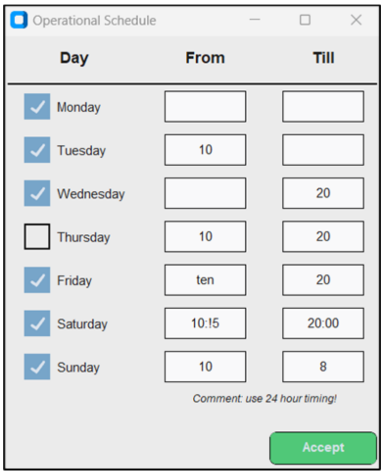
Example #5

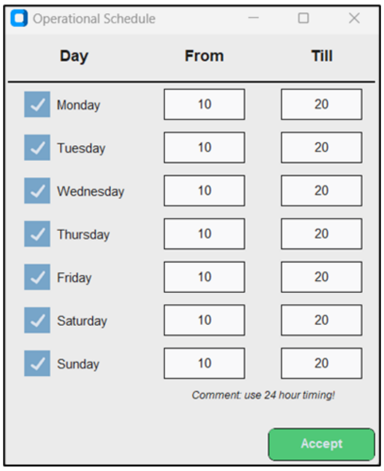
This example shows the allowed variability in the *Country* input: the correct spelling is required, following the available list in **Table 5**. If the provided schema is not followed, Belgium (BE) is assumed as the default.

input: Poland output: Poland
input: poland output: Poland
input: pOlAND output: Poland
input: Polaand output: Belgium (**spelling mistake**)
input: PL output: PL
input: pl output: PL
input: POL output: BE (**incorrect abbreviation**)

Example #6

This example shows how to correctly fulfill the Operational Schedule window in CEAM software. The exemplary input is shown below (on the left) with the corresponding correct version (on the right).





The explanation of the performed assessments by the CEAM software is as follows:

- Monday was selected, yet no working hours were provided; this day would not be included in the examined timetable,
- Tuesday was selected, yet no closing hour was provided; this day would not be included in the examined timetable,
- Wednesday was selected, yet no opening hour was provided; this day would not be included in the examined timetable,
- Thursday was not selected despite the input working hours; this day would not be included in the examined timetable,
- Friday was selected and the working hours were provided, but the opening hour is spelled (not as a number); this day would not be included in the examined timetable,
- Saturday was selected and the working hours were provided, but the time is defined with *HH:MM* format; this day would not be included in the examined timetable,
- Sunday was selected and the working hours were provided, but the time is input with the AM/PM format; this day would not be included in the examined timetable.

Example #7

This example shows the output of the TIG sub-software following the inputs shown in **Figure 8** on the right. The outputted data frame starts on January 1st 2025, at midnight, and ends on April 30th 2025, at 23:59. The selected frequency in minutes is 5; thus, the very last record in the generated file is 30.04.2025 23:55. Part of the generated TIG file is shown below.

Time
01.01.2025 00:00
01.01.2025 00:05
01.01.2025 00:10
…
30.04.2025 23:50
30.04.2025 23:55

Example #8

The user has records of heating (*D1*) consumptions in 15-minute frequencies, yet there are some missing, as shown below.

Time,D1,D2,D3,D4,D5
01.01.2024 00:00,1000,,,,
01.01.2024 00:15,500,,,,

```
01.01.2024 00:30,500,,,,
01.01.2024 00:45,1000,,,,
01.01.2024 01:00,600,,,,
01.01.2024 01:15,800,,,,
01.01.2024 01:45,700,,,,
01.01.2024 04:00,1000,,,,
...
31.12.2024 23:00,2000,,,,
31.12.2024 23:45,2000,,,,
```

CEAM will automatically adjust the provided input by adding the missing records (rows) and doing the hourly averaging (if any of the records for the given is provided; empty rows are not considered). The adjusted file is shown below.

```
Time,D1,D2,D3,D4,D5
01.01.2024 00:00,750,,,,
01.01.2024 01:00,700,,,,
01.01.2024 02:00,,,,,
01.01.2024 03:00,,,,,
01.01.2024 04:00,1000,,,,
...
31.12.2024 23:00,2000,,,,
```

Example #9

Part of the outputted report for the CC assessment (particularly class A2) is shown below.

```
Climate Class 'A2':
 Long-term fulfillment: 80.93% of the time
 Seasonal-term fulfillment: 0.00% of the time
 Short-term fulfillment: 87.77% of the time
 Short-term fulfillment for daily moving average: 86.96% of the time
 Short-term fulfillment for hourly moving average: 86.05% of the time

 Details:
  Ti records within outer min and max range: 7405 (91.21%)
  RHi records within outer min and max range: 7227 (88.94%)
  Ti and RHi records (simultaneously) within outer min and max range: 6571 (80.93%)

  Ti records within annual average ± seasonal offset: 8081 (99.53%)
  RHi records within annual average ± seasonal offset: 0 (0.00%)
  Ti and RHi records (simultaneously) within annual average ± seasonal offset: 0 (0.00%)

  Ti records within short-term offset: 7933 (97.71%)
  RHi records within short-term offset: 7281 (89.60%)
  Ti and RHi records (simultaneously) within short-term offset: 7126 (87.77%)

  Ti_MA_d records within short-term offset: 7903 (97.34%)
  RHi_MA_d records within short-term offset: 7249 (89.21%)
  Ti_MA_d and RHi_MA_d records (simultaneously) within short-term offset: 7060 (86.96%)

  Ti_MA_h records within short-term offset: 7860 (96.81%)
  RHi_MA_h records within short-term offset: 7213 (88.76%)
  Ti_MA_h and RHi_MA_h records (simultaneously) within short-term offset: 6986 (86.05%)
```

This report examines precisely each of the CC proposed in the ASHRAE document, highlighting all time-dependent offsets. Moreover, it shows the requirement fulfillment for T and RH separately, as well as simultaneously (i.e., environment assessment). Additionally, the short-term assessment is provided via traditional assessment, as well as via a moving average[85] (MA) evaluation for daily and hourly calculation steps.

The report is divided into two parts: the summary (under the *Climate Class* '*A2*' header) and detailed information in the following section. The user can read the fulfillment of the CC requirements in % of time.

Example #10

Part of the outputted report for the prediction models evaluation (GRU model for *D1* analysis with *deltaT* as prediction parameter) is shown below.

```
==============================================
=== Energy savings assessment for D1 consumption ===
==============================================

Results for file: D1_deltaT_GRU_predictions.csv
        Examined consumption: D1
        Examined variable: deltaT
        Used prediction model: GRU
        Total Base Consumption: 22496487.58
        Total Predicted Consumption: 23023227.49
        Difference: -526739.91 (-2.34%)
        Model performance assessment:
                R²: 0.6563
                RMSE: 1140.3878
                MAE: 790.9108
                CCC: 0.8010
        Peak Consumption: 8819.70 (recorded at 24.01.2023 23:00)
        Predicted Peak: 7227.00
        Peak Diff: 1592.70 (18.06%)
```

The report shows how accurate the given model is in terms of energy prediction, as well as mimicking the energy profile of the examined case.

Example #11

Part of the outputted report for the selected best models for predictions.

```
=======================================================
=== Selection of the best prediction models for D1 assessment ===
=======================================================

Best model for deltaT variable: GRU
        File: D1_deltaT_GRU_predictions.csv
        R²: 0.6563
        CCC: 0.8010
        Peak diff: 1592.70 (18.06%)
------------------------------------------------------------------
Best model for Ti variable: GRU
        File: D1_Ti_GRU_predictions.csv
        R²: 0.5281
        CCC: 0.6936
        Peak diff: 3751.81 (42.54%)
------------------------------------------------------------------
```

The report shows that GRU algorithm was selected for both models: with *deltaT* and T_i prediction variables; the corresponding statistic factors are shown.

Example #12

The exemplary summary report of the performed energy optimization is shown below.

```
Processing: D1_deltaT_GRU_predictions.csv
D1: 22496487.58
deltaT_pred: 23624823.05
Difference: -1128335.47 (-4.78%)

deltaT_Heating_neg_T-D1_2_s1_pred: 20902351.63 | Savings: 2722471.41 (11.52%)
deltaT_Heating_neg_T-D1_5_s1_pred: 19285324.35 | Savings: 4339498.70 (18.37%)
deltaT_Heating_neg_T-D1_2_s5_pred: 21348095.86 | Savings: 2276727.19 (9.64%)
deltaT_Heating_neg_T-D1_5_s5_pred: 20161734.22 | Savings: 3463088.83 (14.66%)

Processing: D2_deltaTwb_NN_predictions.csv
D2: 13133928.90
deltaTwb_pred: 11958362.20
Difference: 1175566.70 (9.83%)

deltaTwb_Cooling_pos_Twb-D2_2_s1_pred: 11821082.72 | Savings: 137279.49 (1.15%)
deltaTwb_Cooling_pos_Twb-D2_5_s1_pred: 11534482.79 | Savings: 423879.42 (3.54%)
deltaTwb_Cooling_pos_Twb-D2_2_s5_pred: 11845992.74 | Savings: 112369.46 (0.94%)
deltaTwb_Cooling_pos_Twb-D2_5_s5_pred: 11609898.86 | Savings: 348463.35 (2.91%)
```

The report shows the best performing model for *D1* and *D2* assessment, in particular GRU algorithm based on *deltaT* variable, and NN algorithm based on *deltaT*$_{wb}$ variable. In the heading, a real total consumption and the obtained prediction is shown, as well as the corresponding differences. In the second part, the potential saving after applying the given offset strategy is shown, highlighting the savings.

Example: *deltaT_Heating_neg_T-D1_2_s1_pred* presents the heating demand for the *deltaT* model, for ±2^{o}C offset, for s1 operation strategy. The *neg_T-D1* means that *D1* demand was examined, and the offset was applied, potentially lowering the T_i (e.g., from 21.2^{o}C down to 19.2^{o}C). For cooling assessment, it shows *pos_Twb-D2,* which indicates that *D2* demand was examined, and the offset was applied, potentially increasing the T_i (e.g., from 21.2^{o}C down to 23.2^{o}C), and the following T_{wb} was calculated.

Other possibilities: *deltaAbsH_Humidify_mix_AbsH-D4_5_s1_pred* presents the optimization based on *D4* demand, using the *deltaAbsH* as a variable. The offset of ±5% was applied for the s1 operation scenario, allowing for mix offsets, i.e., both positive and negative (for records above the average RH_i the positive was used, and the negative one for records below the average RH_i). The *TISTd_Total_complex_Td_T2RH5_s1_pred* prediction presents the optimization on *D5* demand (indicated with *Total*), using the *TISTd* as a variable. The complex offset strategy was applied for the s1 operation scenario, assuming simultaneous application of ±2^{o}C and ±5% to T_i and RH_i (as indicated with *T2RH5*). The offset adjustment for RH_i is performed in the same way as for *D4* assessment, while T_i offset is based on HDD or CDD presence: negative for heating and positive for cooling seasons.

7.7. EQUATIONS

A full collation of the used equations in the CEAM software is presented below.

$$\sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{n}} \qquad \textbf{eq.1}$$

where: x_i is the individual value, n is the total number of records, μ is the mean value, and σ is the standard deviation.

$$p_{sv} = A \cdot e^{\left(\frac{B \cdot T}{C+T}\right)} \qquad \textbf{eq.2}$$

where: p_{vs} is saturation vapor pressure [hPa], T is temperature [°C], A is a constant equals to 6.112 [hPa], B is empirical constant for water vapor equals 17.67, and C is temperature constant equals 243.15 [°C].

$$p_v = \frac{RH}{100} \cdot p_{sv} \qquad \textbf{eq.3}$$

where: p_v is vapor pressure [hPa], p_{vs} is saturation vapor pressure [hPa], and RH is relative humidity [%].

$$HDD = \sum max\left(0, \frac{T_{HDD} - T_{e,i}}{k}\right) \qquad \textbf{eq.4}$$

where: HDD is heating degree days [oC·days], T_{HDD} is heating base temperature [°C], $T_{e,i}$ is exterior temperature [°C], i is time (hour) index, and k is the number of records per day [-].

$$CDD = \sum max\left(0, \frac{T_{e,i} - T_{CDD}}{k}\right) \qquad \textbf{eq.5}$$

where: CDD is cooling degree days [oC·days], T_{CDD} is cooling base temperature [°C], $T_{e,i}$ is exterior temperature [°C], i is time (hour) index, and k is the number of records per day [-].

$$T_d = \frac{1}{\left(\frac{1}{T_0} - \frac{ln(E/E_0)}{L_{RV}}\right)}$$

$$E = E_s \cdot \frac{RH}{100} \qquad \textbf{eq.6}$$

$$E_s = E_0 \cdot e^{\left(L_{RV} \cdot \left(\frac{1}{T_0} - \frac{1}{T}\right)\right)}$$

where: T_d is dew point temperature [K], T_0 is reference temperature equal to 273.15 [K], T is air temperature [K], E is actual vapor pressure [Pa], E_s is saturation vapor pressure [Pa], E_0 is reference vapor pressure equal to 611 [Pa], RH is relative humidity [%], and L_{RV} is a ratio of latent heat of vaporization to the specific gas for water vapor equal to 5423 [K].

Note: CEAM is converting temperature values to [°C] (after computing).

$$\begin{aligned} T_{wb} = T \cdot \tan^{-1}\left(0.151977 \cdot \sqrt{RH + 8.313659}\right) + \tan^{-1}(T + RH) \\ - \tan^{-1}(RH - 1.676331) + 0.00391838 \cdot RH^{1.5} \\ \cdot \tan^{-1}(0.023101 \cdot RH) - 4.686035 \end{aligned} \qquad \textbf{eq.7}$$

where: T_{wb} is wet-bulb temperature [°C], T is air temperature [°C], and RH is relative humidity [%].

$$\Delta T = T_i - T_e \qquad \textbf{eq.8}$$

where: ΔT is temperature difference [°C], T_i is indoor air temperature [°C], and T_e is outdoor air temperature [°C].

$$\Delta T_d = T_{d,i} - T_{d,e} \qquad \textbf{eq.9}$$

where: ΔT_d is dew point temperature difference [°C], $T_{d,i}$ is indoor dew point temperature [°C], and $T_{d,e}$ is outdoor dew point temperature [°C].

$$\Delta T_{wb} = T_{wb,i} - T_{wb,e} \qquad \textbf{eq.10}$$

where: ΔT_{wb} is wet-bulb temperature difference [°C], $T_{wb,i}$ is indoor wet-bulb temperature [°C], and $T_{wb,e}$ is outdoor wet-bulb temperature [°C].

$$\Delta RH = RH_i - RH_e \qquad \textbf{eq.11}$$

where: ΔRH is relative humidity difference [%], RH_i is indoor relative humidity [%], and RH_e is outdoor relative humidity [%].

$$TIS = \Delta T \cdot \left(1 - \frac{I_{s,i}}{I_{s,max}}\right) \quad if \quad T_e < T_{HDD}$$
$$TIS = \Delta T \quad if \quad T_{HDD} < T_e < T_{CDD} \qquad \textbf{eq.12}$$
$$TIS = \Delta T \cdot \left(1 + \frac{I_{s,i}}{I_{s,max}}\right) \quad if \quad T_e > T_{CDD}$$

where: *TIS* is an optimization factor including the simultaneous impact of air temperature and solar radiation [°C], ΔT is temperature difference [°C], $I_{s,i}$ is solar radiation [W/m^2], $I_{s,max}$ is the maximal recorded solar radiation [W/m^2], T_e is outdoor air temperature [°C], T_{HDD} is heating base temperature [°C], T_{CDD} is cooling base temperature [°C], and *i* is time (hour) index.

$$AbsH = \frac{P_a}{R_w \cdot T}$$
$$P_a = P_s \cdot \frac{RH}{100}$$
$$P_s = P_c \cdot e^{\left[\frac{T_c}{T} \cdot \left(a_1\tau + a_2\tau^{1.5} + a_3\tau^3 + a_4\tau^{3.5} + a_5\tau^4 + a_6\tau^{7.5}\right)\right]} \qquad \textbf{eq.13}$$
$$\tau = 1 - \frac{T}{T_c}$$

where: *AbsH* is absolute humidity [kg/m^3], P_a is actual vapor pressure [Pa], P_s is saturation vapor pressure [Pa], P_c is critical vapor pressure equal to 22.064 [MPa], R_w is specific gas constant for water vapor equal to 461.5 [J/kgK], *T* is air temperature [°C], *RH* is relative humidity [%], T_c is a critical temperature for water equal to 647.096 [K], τ is a dimensionless temperature difference parameter used in the *Wagner & Pruss* equation [-], and a_i are empirical constants [-].

Note: CEAM is converting absolute humidity values to [g/m^3] (after computing), as well as temperature converted to [K] (for computing).

$$D = N \cdot \left(T_{cp,1} - T_e\right)^+ + P \cdot \left(T_e - T_{cp,2}\right)^+ + C \qquad \textbf{eq.14}$$

where: *D* is energy demand [kWh], *N* and *P* are negative and positive slopes [kWh/°C], T_e is outdoor air temperature [°C], $T_{cp,1}$ and $T_{cp,2}$ are change-point temperatures [°C], and *C* is a baseload [kWh].

Note: the $^+$ superscript indicates only positive values allowed within the given brackets.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad \textbf{eq.15}$$

where: R^2 is coefficient of determination [-], *yi* is observed value, $\hat{y_i}$ is predicted value, $\bar{y}$ is mean of the observed values, and *n* is numbers of observations.

Note: the R^2=1 means perfect prediction, R^2=0 means no predictive power, and negative values indicates that using mean value is more accurate than predictions.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad \textbf{eq.16}$$

where: *RMSE* is root mean square error [-], y_i is observed value, $\hat{y_i}$ is predicted value, and *n* is numbers of observations.

Note: the lower the value the better the accuracy. Use both *RSME* and *MAE* to understand error characteristics:

- *RSME* >> *MAE*: large outliers are present,
- *RSME* ≈ *MAE*: errors are relatively uniform.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad \textbf{eq.17}$$

where: *MAE* is mean absolute error [-], y_i is observed value, $\hat{y_i}$ is predicted value, and *n* is numbers of observations.

Note: the lower the value the better the accuracy. Use both *RSME* and *MAE* to understand error characteristics:

- *RSME* >> *MAE*: large outliers are present,
- *RSME* ≈ *MAE*: errors are relatively uniform.

$$CCC = \frac{2s_{y\hat{y}}}{s_y^2 + s_{\hat{y}}^2 + (\bar{y} - \hat{y})^2} \qquad \textbf{eq.18}$$

where: *CCC* is concordance correlation coefficient [-], $\bar{y}$ is mean observed values, $\hat{y}$ is mean predicted values, s_y^2 is variance of observed values, $s_{\hat{y}}^2$ is variance of predicted values, and $s_{y\hat{y}}$ is covariance between observed and predicted values.

7.8. ADDITIONAL INFORMATION

All sorts of additional information (e.g., lists or specs) mentioned throughout the document are listed in this subsection.

The exemplary structure of the var.txt file is as shown below. It contains information provided by the user in the CEAM GUI, in particular: path to the working directory, module selection (scope of analysis), records frequencies and time format for input files, information on solar radiation, and the examined demands, country and city info, units selection, heating and cooling degree days temperatures, outliers assessment with normal distribution range, working hours overview, as well as a selection of the optimization schemas.

```
Working directory: C:/Users/u0160065/Desktop/CEAM_example

Module 0: on
Module 1: on
Module 2: on
Module 3: on
Module 4: on
Module 5: on

ICD_freq: h
ECD_freq: h
ED_freq: h

ICD_format: %d.%m.%Y
ECD_format: %d.%m.%Y
ED_format: %d.%m.%Y

Solar radiation: n

Country: Belgium
City: Not provided

Units: SI

T_HDD: 16.0
T_CDD: 21.0

Outlier: y
Normal distribution: -3:+3

Operation day: Tuesday, from: 10, till: 18
```

Operation day: Wednesday, from: 10, till: 18
Operation day: Saturday, from: 10, till: 20

Optimization Schedule: s1, s4, s5

Examined demands: D1

Table 5. Country headers with the corresponding abbreviations used by the CEAM software.

Albania: AL, Algeria: DZ, Argentina: AR, Armenia: AM, Australia: AU, Austria: AT, Azerbaijan: AZ Bangladesh: BD, Belarus: BY, Belgium: BE, Bolivia: BO, Bosnia and Herzegovina: BA, Brazil: BR, Bulgaria: BG Canada: CA, Chile: CL, China: CN, Colombia: CO, Costa Rica: CR, Croatia: HR, Czech Republic: CZ, Czechia: CZ Denmark: DK, Dominican Republic: DO Ecuador: EC, Egypt: EG, Estonia: EE Finland: FI, France: FR Georgia: GE, Germany: DE, Greece: GR	Honduras: HN, Hong Kong: HK, Hungary: HU Iceland: IS, India: IN, Indonesia: ID, Ireland: IE, Israel: IL, Italy: IT Jamaica: JM, Japan: JP Kazakhstan: KZ, Kenya: KE, Kyrgyzstan: KG Latvia: LV, Liechtenstein: LI, Lithuania: LT, Luxembourg: LU Malaysia: MY, Malta: MT, Mexico: MX, Moldova: MD, Mongolia: MN, Morocco: MA Netherlands: NL, New Zealand: NZ, Nicaragua: NI, Nigeria: NG, North Macedonia: MK, Norway: NO	Pakistan: PK, Panama: PA, Paraguay: PY, Peru: PE, Philippines: PH, Poland: PL, Portugal: PT Romania: RO, Russia: RU San Marino: SM, Saudi Arabia: SA, Serbia: RS, Singapore: SG, Slovakia: SK, Slovenia: SI, South Africa: ZA, South Korea: KR, Spain: ES, Sri Lanka: LK, Sweden: SE, Switzerland: CH Taiwan: TW, Thailand: TH, Tunisia: TN, Turkey: TR Ukraine: UA, United Arab Emirates: AE, United Kingdom: GB, United States: US, Uruguay: UY, Uzbekistan: UZ Venezuela: VE, Vietnam: VN Zambia: ZM, Zimbabwe: ZW

Table 6. List of parameter combinations for a correlation assessment by CEAM.

(Te ; RHe) (Te ; Is) (Ti ; RHi) (Ti ; Is) (Ti ; D1) (Ti ; D2) (Ti ; D5) (Te ; D1) (Te ; D2)	(HDD ; D1) (CDD ; D2) (HDD ; D5) (CDD ; D5) (Tdi ; D1) (Tdi ; D2) (Tdi ; D4) (Tdi ; D5) (Tde ; D1)	(deltaT ; D1) (deltaT ; D2) (deltaT ; D5) (TIST ; D1) (TIST ; D2) (TIST ; D5) (deltaRH ; D4) (deltaRH ; D5) (deltaAbsH ; D4)

(Te ; D5) (RHi ; D4) (RHi ; D5) (AbsHi ; D4) (AbsHi ; D5) (pvi ; D4) (pvi ; D5) (RHe ; D4) (RHe ; D5) (AbsHe ; D4) (AbsHe ; D5) (pve ; D4) (pve ; D5)	(Tde ; D2) (Tde ; D4) (Tde ; D5) (Twbi ; D1) (Twbi ; D2) (Twbi ; D4) (Twbi ; D5) (Twbe ; D1) (Twbe ; D2) (Twbe ; D4) (Twbe ; D5)	(deltaAbsH ; D5) (deltaTd ; D1) (deltaTd ; D2) (deltaTd ; D4) (deltaTd ; D5) (TISTd ; D1) (TISTd ; D2) (TISTd ; D4) (TISTd ; D5) (deltaTwb ; D1) (deltaTwb ; D2) (deltaTwb ; D4) (deltaTwb ; D5) (TISTwb ; D1) (TISTwb ; D2) (TISTwb ; D4) (TISTwb ; D5)

Table 7. A short summary of ASHRAE's CC specifications.

CC	SHORT-TERM FLUCTUATIONS	SEASONAL ADJUSTMENTS	COLLECTION RISK
AA	RH ± 5% T ± 2°C	RH const. T ± 5°C	no risk at all
A1	RH ± 5% T ± 2°C	RH ± 10% T + 5°C T – 10°C	no risk to most
A2	RH ± 10% T ± 2°C	RH const. T + 5°C T – 10°C	
B	RH ± 10% T ± 5°C	RH ± 10% T + 10°C T – 20°C	small risk to most
C	none	none	prevent high risk
D	none	none	prevent dampness

Table 8. Symbols and abbreviations used in the naming of the generated graphical outputs.

d – daily frequency *h* – hourly frequency *m* – monthly frequency *EDW* – Extreme Dry Week *EHW* – Extreme Humid Week *ESW* – Extreme Summer Week *EWW* – Extreme Winter Week *ESunW_n* – Extreme Sunny Week with the lowest solar radiation *ESunW_p* – Extreme Sunny Week with the highest solar radiation	*D1* – heating demand [Wh] *D2* – cooling demand [Wh] *D4* – humidification demand [Wh] *D5* – total demand [Wh] *D5a* – total demand when only *T*-related strategies are considered [Wh] *D5b* – total demand when simultaneous *T*- and *RH*-related strategies are considered [Wh] *HDD* – heating degree days [°C·days] *CDD* – cooling degree days [°C·days]

Te – exterior temperature [°C]
Ti – indoor temperature [°C]
Tde – exterior dry-bulb temperature [°C]
Tdi – interior dry-buld temperature [°C]
Twbe – exterior wet-bulb temperature [°C]
Twbi – interior wet-bulb temperature [°C]
deltaT – difference between *Ti* and *Te* [°C]
deltaTd – difference between *Tdi* and *Tde* [°C]
deltaTwb – difference between *Twbi* and *Twbe* [°C]

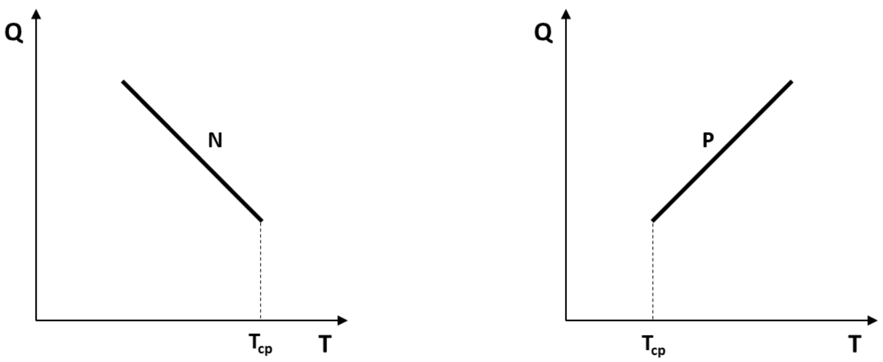
RHe – exterior relative humidity [%]
RHi – interior relative humidity [%]
AbsHe – exterior absolute humidity [g/m³]
AbsHi – interior absolute humidity [g/m³]
pve – exterior vapor pressure [Pa]
pvi – interior vapor pressure [Pa]
deltaRH – difference between *RHi* and *RHe* [%]
deltaAbsH – difference between *AbsHi* and *AbsHe* [g/m³]

Is – total solar radiation [Wh/m²]
TIST – the index used for energy consumption assessment considering the mutual impact of *T* and *Is* [°C]
TISTd – the index used for energy consumption assessment considering the mutual impact of *Td* and *Is* [°C]
TISTwb – the index used for energy consumption assessment considering the mutual impact of *Twb* and *Is* [°C]

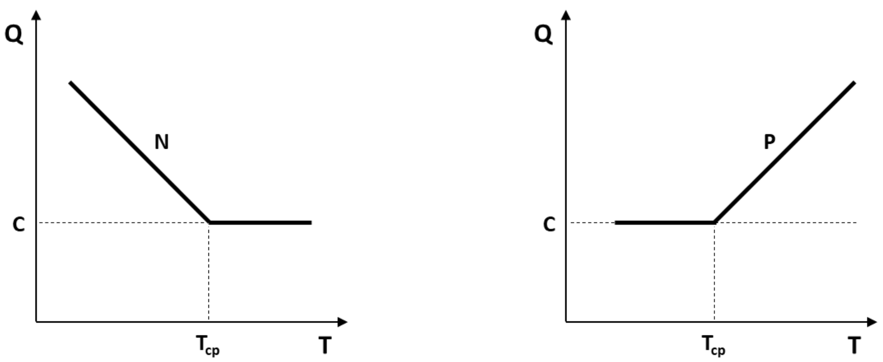
Cor – correlation

2pC – prefix used to describe the used model of Energy Signature: 2 parameters for cooling
2pH – prefix used to describe the used model of Energy Signature: 2 parameters for heating
3pC – prefix used to describe the used model of Energy Signature: 3 parameters for cooling
3pH – prefix used to describe the used model of Energy Signature: 3 parameters for heating
4pC – prefix used to describe the used model of Energy Signature: 4 parameters for cooling
4pH – prefix used to describe the used model of Energy Signature: 4 parameters for heating
5pTotal – prefix used to describe the used model of Energy Signature: 5 parameters for total consumption

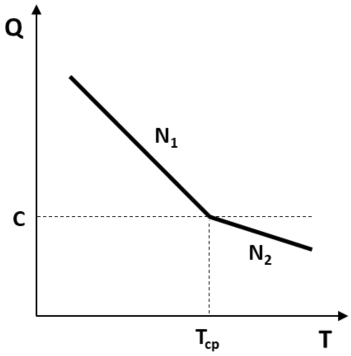
DT – Decision Tree model
NN – Neural Network model
LSTM – Long Short-Term Memory model of Recurrent Neural Network
GRU – Gated Recurrent Unit model of Recurrent Neural Network



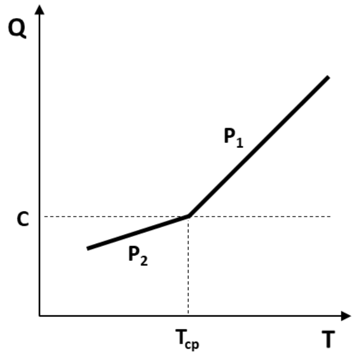
$D = N \cdot (T_{cp} - T_e)^+$ $\quad$ $D = P \cdot (T_e - T_{cp})^+$



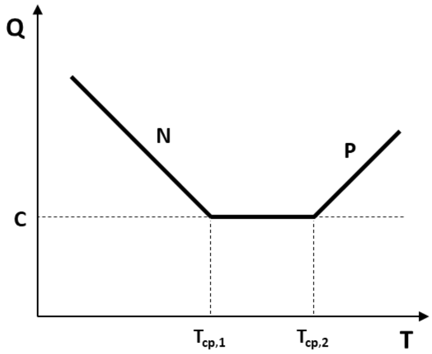
$D = N \cdot (T_{cp} - T_e)^+ + C$ $\quad$ $D = P \cdot (T_e - T_{cp})^+ + C$



$$D = N_1 \cdot (T_{cp} - T_e)^+ - N_2 \cdot (T_e - T_{cp})^+ + C$$



$$D = P_1 \cdot (T_e - T_{cp})^+ - P_2 \cdot (T_{cp} - T_e)^+ + C$$



$$D = N \cdot (T_{cp,1} - T_e)^+ + P \cdot (T_e - T_{cp,2})^+ + C$$

Figure 28. Overview of the available models for ES assessment using CEAM software: 2-parameter models (1st row), 3-parameter models (2nd row), 4-parameter models (3rd row), and the universal 5-parameter model (4th row).

7.9. LIST OF FIGURES

7.10. LIST OF TABLES