Validation of methods for developing typical meteorological years based on future climate models

Seyed Shahabaldin Seyed Salehi[1*]*, Aarne* Männik[2], *Targo* Kalamees[1] and *Martin* Thalfeldt[1]

[1]Department of Civil Engineering and Architecture, Tallinn University of Technology, Tallinn, Estonia
[2]Department of Marine Systems, Tallinn University of Technology, Tallinn, Estonia

Abstract. You Buildings designed for coming decades are more susceptible to global warming, rising average temperatures, and heat waves, which are of a greater risk than in the 1990s. Simulation-based design decisions may be inaccurate if just historical data is used to generate climate files. The building's energy performance, cooling and heating system's design, and room unit sizing require reliable weather data. So far mainly historic meteorological data has been used for simulations. Annual simulations and cooling/heating system design often use a typical meteorological year (TMY) based on historical data using ISO 15927-4 and ISO 15927-2 thorough instructions for TMY for annual simulations and design day data. The above climate models lack the data needed to construct TMYs for building performance simulation (BPS) tools. Climate model forecast accuracy must also be examined from the BPS perspective. This work aims to illustrate and test methods for producing TMYs using three climate models with different resolution and modality in Tartu, Estonia, from 1985 to 2015. The ISO 15927-4 approach selected months from historical meteorological data appropriate for each climate model. Three BPS TMYs were created, and a reference building's energy needs simulated. Long-term historical data was used to simulate the reference building's energy requirement and develop TMYs. The results showed that all the TMYs developed based on NorESM, MPIESM and ECEarth climate models can represent well the energy needs obtained with the TMY based on historical weather data depending on the TMY generation method while MPIESM shows deficiency in cooling season. TMY based on NorESM have the most consistent energy need simulation results while TMY based on ECEarth also performs well both for predicting cooling and heating needs. The future work includes finding other suitable methods for TMY generation that results in even less errors for all models and further validation with other reference buildings for final climate model selection to develop future climate scenarios.

* Corresponding author: shahab.salehi@taltech.ee

1 Introduction

A greater susceptibility of the simulation results of the buildings that will provide service in the following decades have been brought about as a consequence of the impact of global warming, the increase in average temperatures, and the increased occurrence of heat waves. It is worth noting that this vulnerability is noticeably more pronounced than the conditions that were prevalent in the 1990s and early 2000s. (1) It is possible that the simulation-based design decisions made in Estonia could be flawed if only the historical data is taken into consideration for the generation of climate files. This is because of the fact that the climate data for annual energy simulations in Estonia is based on historical data.

In order to facilitate the annual energy simulations of the building, as well as the design of cooling and heating systems, and to determine the appropriate sizing of units for various rooms, it is essential to obtain reliable meteorological data. Over the course of the past nine decades, meteorological data has been collected in a methodical manner for the purpose of conducting simulations. Since historical data may result in inaccurate simulations, a new practice is developed to make use of future climate models in order to avoid conducting simulation-based building design that is inadequate. A few models are already available. (2) The conventional method for designing cooling and heating systems and conducting annual simulations involves making use of a typical meteorological year (TMY) that is derived from historical data. For this reason, the International Organization for Standardization (ISO) standards (3) and (4) offer comprehensive guidance for the collection of weather data and the generation of climate data for annual simulations and design day data. On the other hand, the climate models that were mentioned earlier do not supply sufficient data in a detailed enough manner for the development of TMYs for building performance simulation (BPS) tools. In addition, the accuracy of the predictions made by the climate models needs to be investigated from the point of view of the BPS.

The primary objective of this study is to demonstrate and validate a method for developing TMYs based on three climate models with varying granularity and modality. The location Tartu, Estonia, and the time period between 1985 and 2015 will be used in this study. Selecting months from historical meteorological data that were representative of each climate model data was accomplished through the utilization of the (3) method. Along with the compilation of three TMYs for BPS, a simulation of the energy requirements of a reference building was also performed and the comparative results are presented.

2 Methods

To provide an overview of the process that is followed in this study, the steps are briefly listed as follows:
1. Three sets of CMIP6 (Coupled Model Intercomparison Project Phase 6) climate models for the period of 1985-2014 gathered.
2. Historical climate data is pre-processed in the same format as data provided in CMIP6 outputs.
3. Representative months of the historical data Identified by comparing the cumulative distribution of the CMIP6 climate data to cumulative distribution of the historical data. (2 comparison method provided)
4. Representative months gathered and test meteorological years that are representing each CMIP6 models developed. (3 TMY from CMIP6 model and 1 long term historical data)
5. Simulations with long term historical data and the 3 TMYs with four different reference building models with different usage profile and characteristics carried out.

6. Comparison of the annual energy usage of each building type/climate model is conducted to assess the suitability of CMIP6 models.

In the following subsections, a detailed description of how to follow the process is explained.

2.1 CMIP6 global climate Model Input

This study employs three independent global climate models from CMIP6 model pool to generate climate data for the period of interest and extracting a set of input parameters. Historical period runs of global climate models were utilized for the period spanning from 1985 to 2014 (2). The selected models are three global Earth System Models (ESM): ECEarth3 which is a third-generation collaborative effort of several European meteorological services, research institutes, and high-performance computing centres (referred as ECEarth) (5), MPI-ESM1.2 which is the most recent version of the Max Planck Institute ESM (referred as MPIESM) (6) and NorESM2 - the version 2 of the coupled ESM by Norwegian Climate Centre (referred as NorESM) (7). The output of climate models was used without any application of downscaling methods. Model values were extracted for the Tartu location with nearest neighbour approach. The input parameters were average daily temperature and average daily radiation values for the same period from 1985 to 2014. This data was subsequently employed for test meteorological year (TMY) generation as input for simulations, allowing a comparison between the energy simulation results obtained with having these TMYs as input and those derived from actual historical data.

2.2 Data generation and simulation

The research methodology used in this study included the creation of TMY through the utilization of climate model outputs. The TMY datasets were then used as input in simulation models. The simulation procedure was carried out with the IDA ICE 5 software (8). In order to evaluate the performance and accuracy of the future climate models, the annual energy results derived from these simulations were subsequently compared to the outcomes of simulations over the 30-year period, utilizing actual historical climate data from 1985 to 2014. The detailed procedure is presented in the next section.

2.3 Test meteorological year generation method 1

The novel procedure for development of Test meteorological year for the future climate scenarios is an adjusted version of the method derived from a study conducted in Finland (9) which is itself an improved version of the original method in (3). For this procedure, we gathered the hourly climate data (temperature, radiation, humidity, wind speed and direction) of the studied period and daily averages of the climatic parameters of any future climate models. In short, we briefly define the methods as follows (P in the process is temperature, solar radiation and humidity):

1. Calculate the daily means of each climate parameter in both historical data and target climate model, p
2. For each calendar month of historical data ('MM' for all years), the Cumulative Distribution Function (CDF) of daily means over all years is computed, using rank-order statistics.

$$\Phi(p, m, i) = \frac{K(i)}{N+1} \qquad (1)$$

Where K(i) is the rank order of the ith value of the daily means within that calendar month (m) in the whole dataset, and N is the total number of data points for that month in the whole dataset.

3. Data for each calendar month in each year (MM/YYYY) was derived from the targeted climate model results (e.g., NorESM), calculate the F parameter according to formula (2).

$$F(p, y, m, i) = \frac{J(i)}{N+1} \qquad (2)$$

Where J(i) is the rank order of the ith value of the daily means within that month (m) and that year (y)

4. Calculate of the Finkelstein-Schafer (F.S.) values using formula (3) for the climate parameters then incorporating weighting factors ('Wp') based on the significance of climate parameters in the specific region of the study for heating and cooling energy use using formula (4). In this paper, weighting factors for cold climate of Estonia is used.

$$F.S_{(p,y,m)} = F(p, y, m, i) - \Phi(p, m, i) \qquad (3)$$

$$F.S_{(y,m)\ total} = Wp \cdot \sum_{p=1}^{n} F.S_{(p,y,m)} \qquad (4)$$

Where Wp is the weighting factor for the corresponding climate parameter. The weighting factors for cold climate are presented in Table 1.

5. The F.S. values for all parameters are summed (F.S. total) and then sorted. For each calendar month, for the three months with the lowest total ranking, calculate the deviation of the monthly mean wind speed from the corresponding multi-year calendar month mean. The month with the lowest deviation in wind speed is the selected month.

Table 1. weighting factors for climate parameters suggested in (9)

Wt, Dry-bulb temperature	Wh, Humidity	Wr, Radiation
1	0	0
1	0	0
0.8	0	0.2
0.8	0	0.2
0.5	0	0.5
0.5	0	0.5
0.5	0	0.5
0.5	0	0.5
0.8	0	0.2
0.8	0	0.2
1	0	0
1	0	0

ensure a smooth transition when selecting the "chosen" months for creating the Test Meteorological Year (TMY), interpolation is used to determine climate parameters for the final eight hours of each month and the initial eight hours of the subsequent month except for radiation, ensuring that the reference year can be repeatedly applied in simulations.

2.4 Test meteorological year generation method 2

The second method has the same principles as the first method, with the difference that instead of ranking climate parameters and calculating Finkenstein-Schafer (F.S) values, the

mean average error between the cumulation distribution curves of the historical data and target climate model was calculated. This means that for this method the whole process shall be followed as previous section, but instead of using formulas (1) to (4), formula (5) and (6) was used to define the comparative parameters.

$$MAE_{(p,y,m)} = \left(\frac{\sum_{i=1}^{n} Ph(i) - Pt(i)}{N * (\max(Ph(i)) - Pt(i))} \right) \qquad (5)$$

$$MAE_{(y,m)\ total} = Wp \cdot \sum_{p=1}^{n} MAE_{(p,y,m)} \qquad (6)$$

Where *Ph(i)* is the daily average values of historical data and *Pt(i)* is the daily average values of the target climate model and *N* is the total number of *Ph(i)-Pt(i)* values.

2.5 Comparative analysis

Two types of evaluation were used and presented in the results section: qualitative and quantitative.

The qualitative analysis was performed to determine the level of similarity or difference between the distribution curves of the energy simulation results. These curves were constructed based on the outcomes obtained from simulations conducted using two types of input files: TMY input files derived from climate models, and 30-year simulations utilizing real climate data.

The quantitative analysis was conducted between the annual energy need outcomes derived from simulations employing typical meteorological year input files and the average annual energy need during a 30-year period of simulations utilizing actual historical climatic data from 1985 to 2014.

2.6 Simulation models and case studies

This study used four building simulation models as reference buildings to include non-residential and residential usage and operating profiles. Two office models have different window sizes to account for solar heat gain differences while the other two models are single-family homes: an older, poorly insulated one and a newly built, well-insulated one. The originate of the model is from another study (10) but the building models were slightly modified from those used in the Estonian cost optimality study.

To reduce solar heat gains on office buildings' thermal balance, smaller windows are half the size of bigger ones. The representative model of the older house has double thermal transmittance for walls and windows.

The interior heat gain profiles for open-plan offices and detached homes from EN 16798-1 (11) were used as usage profiles in this study with detailed information available in the standard. Table 2 and Table 3 include the details of the reference buildings' properties and simulation settings, and Figure 1 shows the four buildings in the simulation software. The solar heat gain coefficient of glazed areas was 0.5 for houses and 0.3 for office buildings. The ventilation fan operation is set as "always on" in house models and "workdays 06-19" for offices.
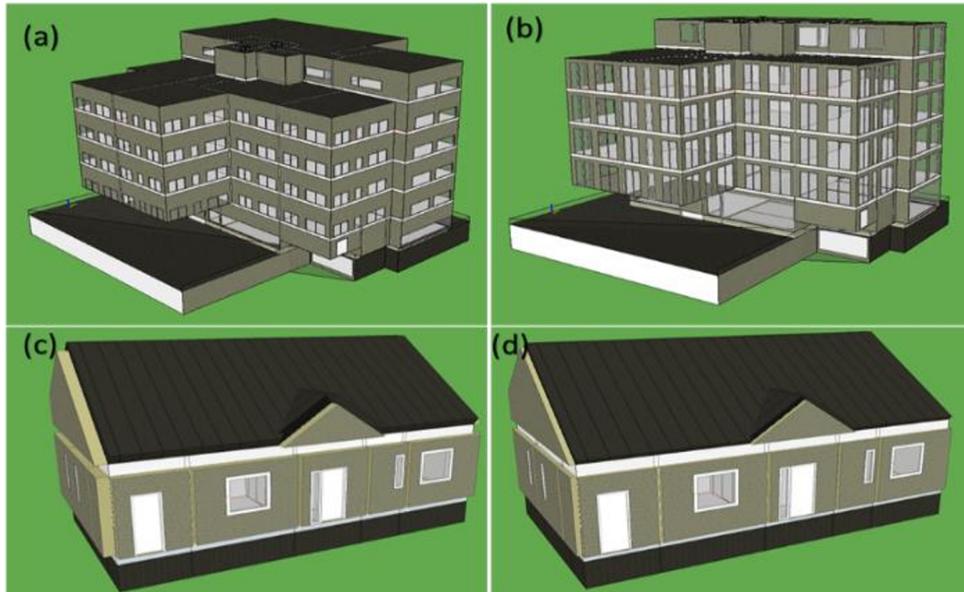
Fig. 1. 3D Model of the reference buildings. (a) office building with small windows, (b) office building with big windows, (c) single-family house with thick walls (lower U value, representing new house), and (d) single-family house with thin walls (representing old house)

Table 2. Parameters used in the simulations of the reference buildings.

Parameter	Office building	Single-family house
Gross area, m2	4958	100
Number of stories, -	6	1
Number of zones, -	12	6
Window-to-wall ratio, %	60 & 30	17
Occupants, person/m2	0.0586	0.532
Equipment, W/m2	12	2.4
Lighting, W/m2	12	8
Ventilation airflow rate, L/(s·m2)	2	0.42
Infiltration constant air flow rate, L/(s·m2)	0.056	0.0238
Temperature set points, heating/cooling, °C	21/25	21/25

Table 3. General building model properties

Model	Total thermal transmittance U·A [W/K]	Specific heat loss for heated area H/A, W/(m²·K)
Old house	10.7	1.33
New house	10.7	0.81
Office with big windows	113	0.38
Office with small windows	98	0.32

3 Results and discussion

The climate file generation based on different models are done using 2 methods described in the methods section. In order to validate the simulations input data, the temperature and radiation curves of the generated TMYs are presented in Figure 2, Figure 3 and Figure 4. As it is shown, the cumulation distribution curves of the temperature were qualitatively similar but the radiation curves were not qualitatively close and therefore, the TMY generation

process had limitations. Figure 5 indicates all input data for simulations that are generated based on the climate model files.
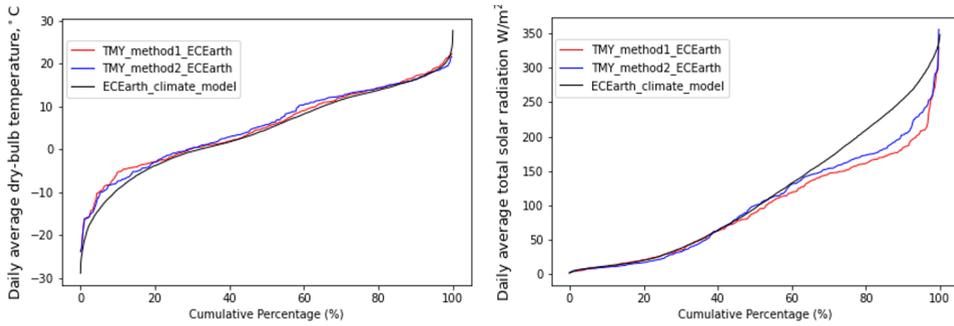

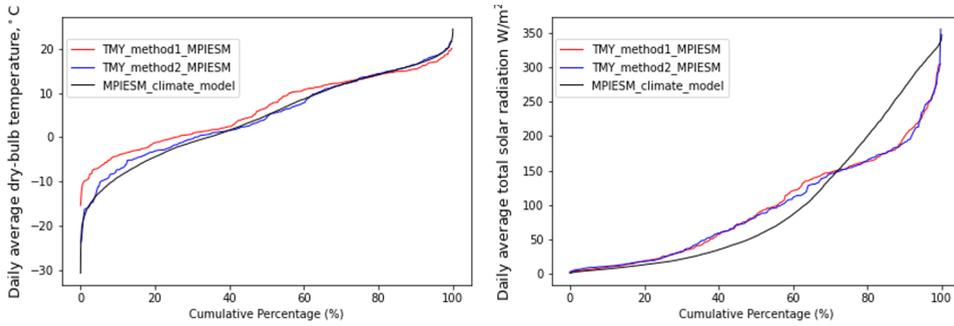**Fig. 2.** Temperature and radiation cumulation distribution curves for ECEarth model

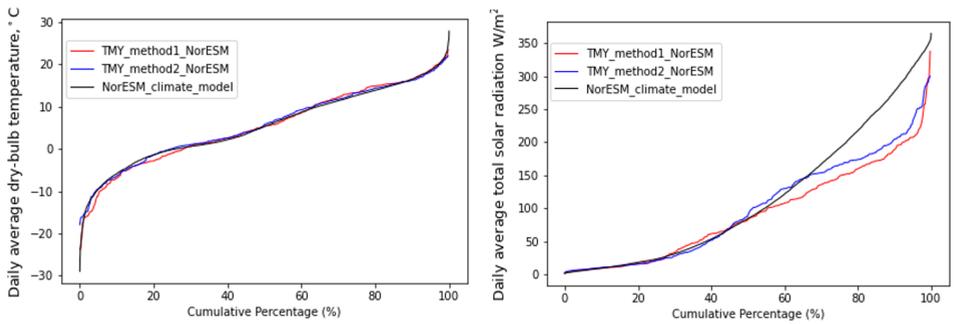**Fig. 3.** Temperature and radiation cumulation distribution curves for MPIESM model

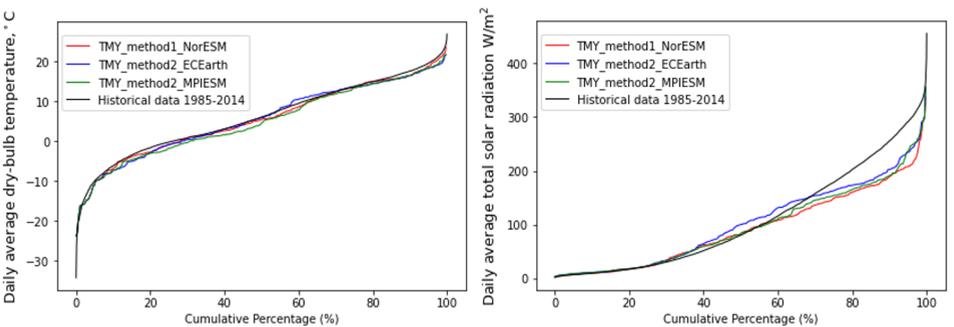**Fig. 4.** Temperature and radiation cumulation distribution curves for NorESM model

Fig. 5. Temperature and radiation duration curves for each input climate file (TMYs) for simulations

Qualitatively, the best result for NorESM occurred when method 1 was used. However, method 2 performed better with ECEarth and MPIESM climate models. The resulted cumulation distribution curves all climate models illustrate the difference of the methods used and helped us choose the method with justifiable results to be used for climate file generation. This process of finding justified model is qualitative based on the charts by focusing on matching minimum and maximum values as well as the general trend of the line to cover both negative and positive differences. After carefully gone through the qualitative assessment, method 1 was chosen for NorESM and method 2 was chosen for MPIESM and ECEarth. However, the simulation results will still have deviations to the real climate model since there were no month/year combination in the historical data to be identical to the three climate models daily averages.

3.1 Simulation results

The distribution curves of the total cooling/heating energy need for each building according to four methods are presented in Figure 6 and Figure 7. Three TMY generated from future climate model (ECEarth, MPIESM, NorESM) and 30-year (1985-2014) period simulation results.

The qualitative assessment of the hourly energy needs curves shows that the curves are all aligned, and the general trend of all models are within an acceptable range. None of the models can be pointed to have obvious deficiency in predicting cooling/heating needs while the choice of TMY generation method is a major factor for them to be aligned with daily averages from climate models or 30-year period simulation results.
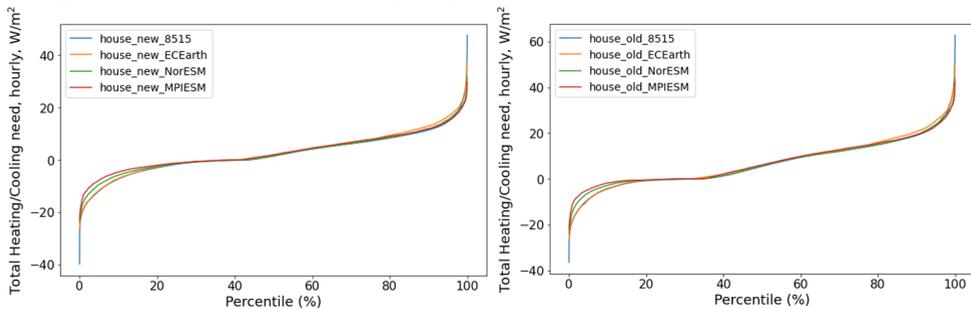

Fig. 6. The distribution curve of the Total Heating (positive) and Cooling (negative) needs in the house models for every climate data input.
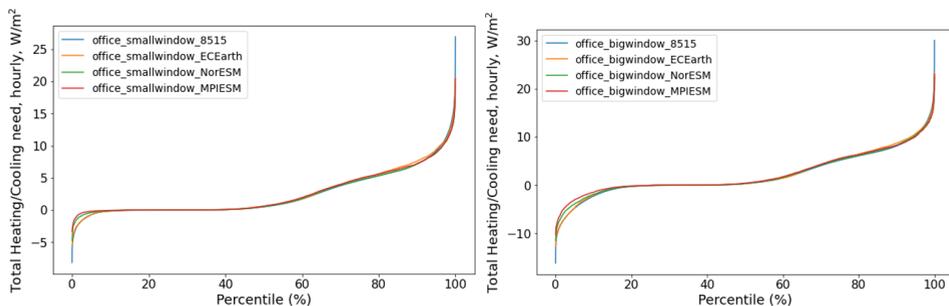

Fig. 7. The distribution curve of the Total Heating (positive) and Cooling (negative) needs in the office models for every climate data input.

Starting with the office featuring large windows, the 30-year average energy need for heating ranges from 31.4 to 34.7 kW/m²/y across different climate models and long term simulation, while cooling need varied from 7.9 to 10.8 kW/m²/y. Notably, the MPIESM

model consistently exhibited lower cooling need, suggesting potential challenges in accurately predicting cooling needs or that the methods did not provide the perfect month/year that is fit to the model. This issue was already discussed in the first parts of the results section. The long-term heating need (32.1 kW/m²/y) is slightly higher than MPIESM result (31.4 kW/m²/y) and lower than the other two climate models while the cooling need (10.6 kW/m²/y) is higher than ECEarth and MPIESM, but slightly lower than NorESM result (10.8 kW/m²/y).

On the contrary, the office model with small windows demonstrated heating and cooling need values ranging from 32.3 to 36.3 kW/m²/y and 4.0 to 6.2 kW/m²/y, respectively. Again, the MPIESM model stands out with the lowest cooling need, prompting consideration for additional research to validate its future use for BPS. Moreover, both heating and cooling needs of the long-term simulation is higher than all climate models while NorESM performs better for cooling need and MPIESM has the closest value to the long-term results.

Focusing on the newly built house, heating and cooling energy need spans from 42.7 to 46.1 kW/m²/y and 6.9 to 7.5 kW/m²/y, respectively. Minor variations across different climate models suggest that the design attributes of the newly built house could have potentially adapted well to diverse climate scenarios.
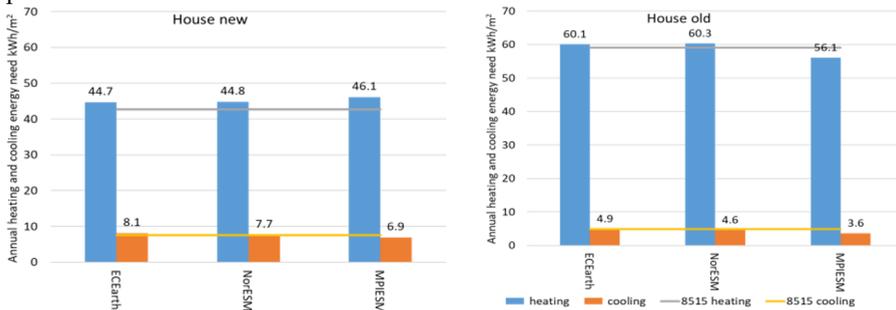

Fig. 8. The annual heating and cooling needs per m2 area in the house for every climate model.
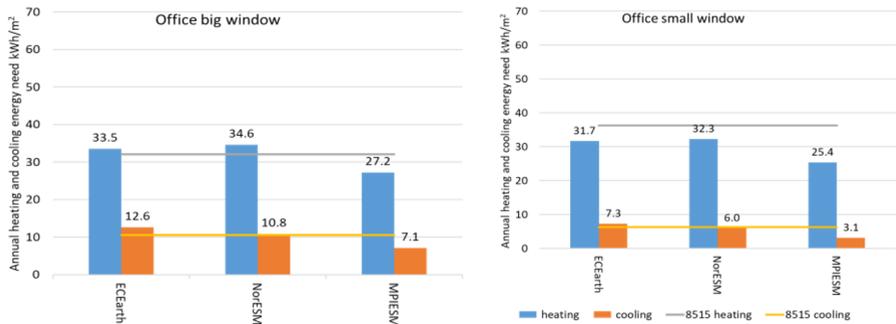

Fig. 9. The annual heating and cooling needs per m2 area in the office for every climate model.

In contrast, the old house model exhibited higher heating and lower cooling needs, with values ranging from 56.1 to 60.3 kW/m²/y for heating and 3.6 to 4.9 kW/m²/y for cooling. Once again, the MPIESM model displayed the least cooling need, emphasizing its role in predicting cooling needs for buildings with significant thermal gains. Figure 8 and Figure 9 indicate all the annual heating and cooling needs per m2 area in the home and office models for every climate model. The absolute differences between the annual heating and cooling needs based on climate models and long term simulations are gathered in Table 4. As it is shown, MPIESM is more likely to underestimate energy needs compared to other two models and the deviation values for NorESM and ECEearth are similar, while NorESM has low and consistent deviation values for cooling needs.

Table 4. Absolute differences between the annual heating and cooling needs per m2 from simulation using TMY based on different future climate models and long term simulation results. Red boxes means the value was lower than the long term results and green boxes are for higher values compared to long term results.

Models		ECEarth	NorESM	MPIESM
Office Big windows	heating	2.6	2.5	-0.7
	cooling	-1.2	0.2	-2.7
Office Small windows	heating	-3.9	-4.0	-2.0
	cooling	-1.7	-0.2	-2.2
House Newly built	heating	2.0	2.1	3.4
	cooling	0.6	0.2	-0.6
House Old	heating	1.0	1.2	-3.0
	cooling	0.0	-0.3	-1.3

A scatterplot of all results is provided in Figure 10 to summarize the heating and cooling needs and demonstrate the methods performances in both qualitative and quantitative matter combined. The data markers that are above the long-term values line mean that the cooling/heating need is higher than the long-term simulation results. Among different building models and climate models, the markers are spread both above and below the line but the majority of the markers from NorESM results are closer to the long-term line in both heating and cooling need scatterplot compared to other methods. In the heating chart, most of the ECEarth markers are under NorESM markers, while MPIESM markers are not that far from the line in heating. Regarding the cooling scatterplot, NorESM results are very close to the line which means that the TMY based on NorESM is consistent among different building type and usage profile following by ECEarth TMY which also have strong performance except for offices with big windows.
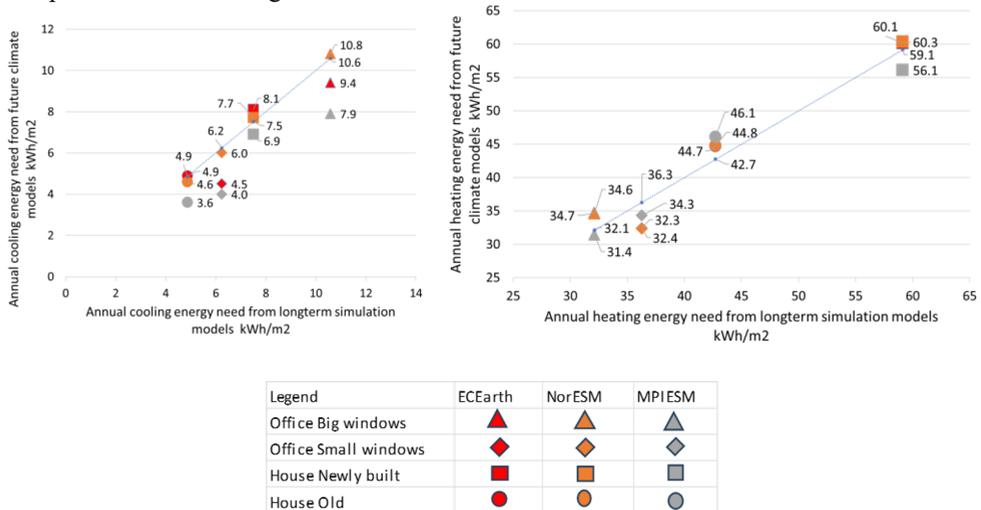

Fig. 9. Scatterplot of all results to summarize all results and reflect the possible climate model or method deficiencies. Numbers refer to climate models and the line refers to the long-term values.

3.2 Discussion

The approach of this study was novel and has not been tested before in the literature. The major obstacle in this study was to be able to gather hourly data that are suitable for simulations since the future climate models included only average daily or average 3-hour values. While they cannot be used themselves as the simulation input, finding a representative month from the historical data to match to the specifications of the future climate models can be controversial. All climate models are still relevant while MPIESM needs to be re-evaluated and possibly phased out for cooling seasons. The results from NorESM and ECEarth are close to each other, which may lead to the conclusion that they are better models for BPS studies in the future.

4 Conclusion

This study addresses the critical challenge of developing test meteorological year based on future climate models. The vulnerability of simulations relying solely on finding month/year combination from historical data that fits with the climate models. The TMY based on MPIESM model consistently exhibits lower cooling need, suggesting its potential deficiency in predicting cooling needs for buildings exposed to the future climate. TMY based on NorESM have the most consistent energy need simulation results while TMY based on ECEarth also performs well both for predicting cooling and heating needs. Since the TMYs are not completely aligned with the climate models (Specifically the radiation duration curves that have direct impact in cooling needs), the deficiencies can reduce and the performance of the TMYs develops if other methods are suggested that can produce TMYs with closer parameter duration curves to the climate models.

The study highlights the necessity of adopting future climate models in simulation-based design decisions to ensure resilience in the face of evolving climate conditions. Further research is recommended to propose new methods or refine the proposed methods as well as to evaluate the long-term suitability of certain climate models, particularly MPIESM, for accurate predictions in cooling scenarios.

5 Acknowledgement

References

1. International Energy Agency. Global Energy & CO2 Status Report. IEA Publications, France, 2019.
2. Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development. 2016 May 26;9(5):1937–58.
3. EN ISO 15927-4. Hygrothermal Performance of Buildings. Calculation and Presentation of Climatic Data. Hourly Data for Assessing the Annual Energy Use for Heating and Cooling. 2005.

4. EN ISO 15927-2. Hygrothermal performance of buildings — Calculation and presentation of climatic data — Part 2: Hourly data for design cooling load. 2009.

5. EC-Earth Consortium (EC-Earth). EC-Earth-Consortium EC-Earth3 model output prepared for CMIP6 CMIP historical. Version YYYYMMDD. Earth System Grid Federation. 2019.

6. Jungclaus J, Bittner M, Wieners KH, Wachsmann F, Schupfner M, Legutke S, et al. CMIP6.CMIP.MPI-M.MPI-ESM1-2-HR.historical. Earth System Grid Federation. 2019 Jan 1

7. Bentsen M, Jan D, Seland Ø, Toniazzo T, Gjermundsen A, Graff LS, et al. NCC NorESM2-MM model output prepared for CMIP6 CMIP historical. Version YYYYMMDD. Earth System Grid Federation. 2019

8. EQUA. IDA ICE - Indoor Climate and Energy. Stockholm. 2013.

9. Targo Kalamees, Kirsti Jylhä, Tietäväinen H, Juha Jokisalo, Simo Ilomets, Reijo Hyvönen, et al. Development of weighting factors for climate variables for selecting the energy reference year according to the EN ISO 15927-4 standard. Energy and buildings [Internet]. 2012 Apr 1 ;47:53–60

10. Arumägi E, Simson R, Kuusk K, Kalamees T, Kurnitski J. Analysis of cost-optimal minimum energy efficiency requirements for buildings. Tallinn University of Technology: Tallinn, Estonia. 2017.

11. ISO EN 16798-1. Energy performance of buildings - Part 1: Indoor environmental input parameters for design and assessment of energy performance of buildings. Brussels. 2019.