

Article

# Impact Assessment for Building Energy Models Using Observed vs. Third-Party Weather Data Sets

Eva Lucas Segarra <sup>1,\*</sup>, Germán Ramos Ruiz <sup>1,t</sup>, Vicente Gutiérrez González <sup>1</sup>,  
Antonis Peppas <sup>2</sup>, Carlos Fernández Bandera <sup>1</sup>

<sup>1</sup> School of Architecture, University of Navarra, 31009 Pamplona, Spain; gramrui@unav.es (G.R.R.); vgutierrez@unav.es (V.G.G.); cfbandera@unav.es (C.F.B.)

<sup>2</sup> School of Mining and Metallurgical Engineering, National Technical University of Athens (NTUA), 15780 Athens, Greece; peppas@metal.ntua.gr

\* Correspondence: elucas@unav.es; Tel.: +34-948-425-600 (ext. 800000)

† These authors contributed equally to this work.

Received: 20 July 2020; Accepted: 17 August 2020; Published: 21 August 2020



**Abstract:** The use of building energy models (BEMs) is becoming increasingly widespread for assessing the suitability of energy strategies in building environments. The accuracy of the results depends not only on the fit of the energy model used, but also on the required external files, and the weather file is one of the most important. One of the sources for obtaining meteorological data for a certain period of time is through an on-site weather station; however, this is not always available due to the high costs and maintenance. This paper shows a methodology to analyze the impact on the simulation results when using an on-site weather station and the weather data calculated by a third-party provider with the purpose of studying if the data provided by the third-party can be used instead of the measured weather data. The methodology consists of three comparison analyses: weather data, energy demand, and indoor temperature. It is applied to four actual test sites located in three different locations. The energy study is analyzed at six different temporal resolutions in order to quantify how the variation in the energy demand increases as the time resolution decreases. The results showed differences up to 38% between annual and hourly time resolutions. Thanks to a sensitivity analysis, the influence of each weather parameter on the energy demand is studied, and which sensors are worth installing in an on-site weather station are determined. In these test sites, the wind speed and outdoor temperature were the most influential weather parameters.

**Keywords:** weather file management; weather datasets; weather stations; building energy simulation; sensitivity analysis of weather parameters; thermal zone temperature

## 1. Introduction

The sustainable development goals report of 2019 highlighted the concern of the United Nations toward a more sustainable world where people can live peacefully on a healthy planet. One of the most important areas for the protection of our planet is the actions to mitigate climate change. “If we do not cut record-high greenhouse gas emissions now, global warming is projected to reach 1.5 °C in the coming decades” [1]. This concern was also endorsed by 186 parties in the Paris agreement on climate change in 2015 [2]. One of the strategies for tackling climate change is to reduce energy consumption (by increasing the system efficiency) and increase the use of clean energy so that greenhouse gas emissions are reduced. In this process of decarbonization, the buildings and the construction sector are critical elements, since as the Global Status Report for Buildings and Construction highlighted, they are responsible “for 36% of final energy use and 39% of energy and process-related carbon dioxide (CO<sub>2</sub>)

emissions in 2018, 11% of which resulted from manufacturing building materials and products such as steel, cement and glass.” [3].

For this reason, building energy models (BEMs) play an important role in the understanding of how to reduce the energy consumed by buildings (lighting, equipment, heating, ventilation, and air-conditioning (HVAC) systems, etc.) and how to optimize their use. As Nguyen et al. highlighted, there is a huge variety of building performance simulation tools [4], and EnergyPlus is one of the most used [5]. In any simulation program, to obtain reliable results, the model must not only be adjusted to the behavior of the building, but all the files on which it depends must be reliable and suitable for the use that will be given to the model. Of all these files, the weather file is, perhaps, one of the most important [6].

Bhandari et al. highlighted that there are three kinds of weather data files, *typical*, *actual*, and *future*, which should be selected according to the use of the energy model [6]. The first corresponds to a representation of the weather pattern of a specific place taking into account a set of years (commonly 20–30 years). For each month, the data are selected from the year that was considered most “typical” for that month so that it represents the most moderate weather conditions, excludes weather extremities, and reflects long-term average conditions for a location. In general, they are used to design and study the behavior of the building under standard conditions, to obtain Energy Performance Certificates, to study the feasibility of a building’s refurbishment, etc. The typical files are not recommended in extreme conditions, such as designing HVAC systems for the worst case scenario.

There are two main sources of typical weather files: those developed by the National Renewable Energy Laboratory (NREL), which come from stations in the United States and its territories, where the different versions (typical meteorological year (TMY), TMY2, and TMY3) take into account different numbers of stations, time periods, solar radiation considerations, etc. [7,8]; and those developed by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) as a result of the ASHRAE Research Project 1015 [9,10]: the International Weather for Energy Calculations (IWEC), which takes into account weather stations outside the United States and Canada. The latest version (IWEC2) covers 3012 worldwide locations [11] taking into account a 25 year period (1984–2008) from the Integrated Surface Hourly (ISH) weather database.

The second kind of weather data file, *actual*, corresponds to a specific location and time period. It could be obtained from an on-site weather station or by processing data from several nearby stations. The latter option is commonly used by external data providers. This type of file is usually used to carry out building energy model calibrations, calculate energy bills and utility costs, study the specific behavior of HVAC systems, etc. As can be seen, these files take into account extraordinary situations, such as heat waves, adverse or extraordinary atmospheric phenomena, etc. [12].

Finally, the last kind of weather data file, *future*, is mainly used to simulate how it is possible to adapt the building energy demand to an external energy requirement (demand response [13]) or to obtain better use strategies of HVAC systems in certain situations (model predictive control [14,15]). As Lazos et al. highlighted, there are three common groups of forecasting techniques: statistical, machine learning, and physical and numerical methods [16].

Therefore, each kind of weather data file has a certain purpose; therefore, the reliability of the results will largely depend on the accuracy and suitability of the file selected [17,18].

Many studies highlighted the importance of the weather file when performing building energy analysis, as its accuracy is generally assumed, although it is out of the control of the person in charge of the simulation, its difference being significant. The following are some examples of studies that use *typical* weather files for different applications where the weather data are relevant. It is meaningful in retrofitting scenarios [19], or when quantifying renewable energy as in the cases when sizing photovoltaic solar panels [20], or when used to obtain ground temperatures [21], even when studying climate change; the accuracy of the weather files is also very important, as they are the baseline files in the process of morphing the data [22–24]. Therefore, when using such files, it is important to try to use the most recent [25].

There are other studies that analyze the building energy performance using *actual* weather files, either from commercial vendors [26], or developed using nearby weather stations, not located in the building [6,27–29], or the lesser ones, from weather stations placed in the building or in its surroundings [30]. Finally, regarding the *future* weather files, the uncertainty of the forecast files is closely related to the accuracy of the files on which they are based [31–33]. There are many studies that highlight the importance of the weather files, measuring, for example, their impact on passive buildings [34], on micro-grids [35,36], etc., calculating the loads of district energy systems [37], the electricity consumption with demand response strategies [13], or evaluating the effect on comfort conditions [38]. Some analyzed the effect that certain parameters of the weather file have, emphasizing the temperature as the most sensitive value for load forecasting [39,40].

In terms of temporal resolutions, most research focused on annual results when comparing different sets of weather data. Wang et al. analyzed the uncertainties in annual energy consumption due to weather variations and operation parameters for a reference office prototype, concluding that uncertainties caused by operation parameters were much more significant than weather variations [26]. Crawley et al. analyzed the energy results using measured weather data for 30 years and several weather datasets for a set of five locations in the USA, and the variation in annual energy consumption was on average  $\pm 5\%$  [17]. Seo et al. conducted a similar study, also analyzing the peak electrical demand with similar results: a maximum difference of 5% [41].

In terms of research that focused on monthly criteria, Bhandari et al. found that, when using different weather datasets, the annual energy consumption could vary around  $\pm 7\%$ , but up to  $\pm 40\%$  when monthly analysis was performed [6]. Radhi compared the building's energy performance of using past and recent (up-to-date) weather data with annual and monthly criteria. This showed a difference of 14.5% between the annual electricity consumption simulated with past data and actual consumption, while this difference grew up to 21% for one month when monthly criteria were used [42]. Finally, there were a few studies where the temporal resolution was less than one month. With weekly criteria, Silvero et al. compared five different weather data sources with the observed meteorological data, showing that, for the annual criteria, the results were similar, but for the hottest and coldest weeks of the year, the inaccuracies increased [28].

The aim of this work is to show how to evaluate the impact of using two different *actual* weather datasets on the building energy model simulations, measuring both their effect on energy demand and indoor temperature. The purpose is to analyze if the data provided by the *on-site* weather stations (with a high economic cost and maintenance) could be replaced by the *actual* data provided by a *third-party*. For this study, four test sites based on real buildings were used to compare the existing variations when weather files with data obtained from real stations and external provider were used in the simulations. These test sites are part of an EU funded H2020 research and innovation project SABINA (SmArt BI-directional multi eNergy gAteway) [43], which aims to develop new technology and financial models to connect, control, and actively manage the generation and storage of assets to exploit synergies between electrical flexibility and the thermal inertia of buildings. The energy demand variation analysis was measured by grouping it into different periods (annual, seasonal, monthly, weekly, daily, and hourly) since as explained by ASHRAE [44], "... the aggregated data will have a reduced scatter and associated CV(RMSE), favoring a model with less granular data." The objective of the paper is to highlight these differences in the results when using different granular criteria since, depending on the use of the building energy model, their influence can be significant, for example for calibration purposes, where the monthly or hourly criteria are required. A sensitivity analysis was also performed to evaluate the influence of each weather parameter on the energy demand variation when using the two different *actual* weather datasets.

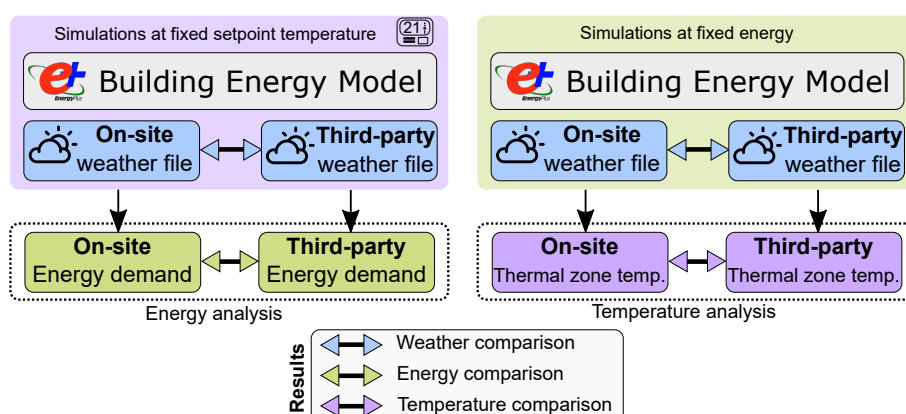
The main contributions of this research are: (1) four real test sites, with different uses and architectural characteristics, located in three different climates were employed in the study; (2) while most of studies that analyze weather data influence in building energy simulation use the typical meteorological year (TMY) [19–24], this study performed a comparison of two *actual* datasets: *on-site*

and *third-party* weather data; (3) when the studies used *actual* weather data, they usually lacked a local weather station due to its expensive installation [6,27–29]; instead, the observed weather data from this study were provided by three weather stations installed on the building roofs or in their surroundings, providing on-site measurements; and (4) the energy results are shown with different temporal resolutions (from annual to hourly) in order to highlight the differences in the variations when the data are accumulated.

The paper is structured as follows. In Section 2, we describe the methodology used to analyze both the differences between the *on-site* and *third-party* weather datasets and the variation produced by these weather files in the results of the simulations in terms of the energy demand and in terms of the indoor temperature. In Section 3, we show the results obtained from the different approaches: the weather datasets comparison (Section 3.1), energy demand (Section 3.2), and indoor temperature (Section 3.3). Finally, in Section 4, we discuss the results obtained in the study, and in Section 5, the conclusions are presented.

## 2. Methodology

Figure 1 shows the diagram of the methodology used in this study and the three analyses that were performed: (1) the weather data comparison between the data provided by the *on-site* weather stations and the *third-party* data; and through energy model simulations using these two weather datasets; (2) the energy demand comparison; and (3) the indoor temperature comparison.



**Figure 1.** Diagram of the methodology used to check the effect of the use of different *actual* weather data in the building energy simulations. Three analyses are performed in the study: weather data comparison (methodology explained on Section 2.1), energy comparison (Section 2.2), and indoor temperature comparison (Section 2.3).

### 2.1. Weather Data Comparison Methodology

Once the weather data from the *on-site* weather station and the *third-party* are gathered, the first analysis is the comparison using a Taylor Diagram [45,46]. This provides a simple way of visually showing how closely a pattern matches an observation, and it is a useful tool to easily compare different parameters at a glance using the same plot. This type of comparison is widely used when weather parameters are analyzed [28,47–51]. Developed by Taylor in 2001, this diagram shows the correspondence between two patterns (in this case, *third-party* weather data as the test field ( $f$ ) and *on-site* weather data as the reference field ( $r$ )) using three statistical metrics: the correlation  $R$ , the centered root-mean-squared difference  $RMS_{diff}$ , and the standard deviation  $\sigma$  of the test and reference field.

The correlation  $R$  (3) is used to show how strongly the two fields are related, and it ranges from  $-1$  to  $1$ . The centered root-mean-squared difference  $RMS_{diff}$  (4) measures the degree of adjustment in amplitude. The closer to  $0$ , the more similar the patterns are. Both indexes provide complementary information quantifying the correspondence between the two fields, but to have a more complete characterization, their variances are also necessary, which are represented by their standard deviations



$\sigma_f$  (1) and  $\sigma_r$  (2) [46]. To allow the comparison between different weather parameters and to show them in the same plot,  $RMS_{diff}$  and  $\sigma_f$  are normalized by dividing both by the standard deviation of the observations ( $\sigma_r$ ). Thus, the normalized reference data have the following values:  $\sigma_r = 1$ ,  $RMS_{diff} = 0$ , and  $R = 1$ .

Figure 2 shows the Taylor diagram baseline plot and how it is constructed. The reference point appears in the x-axis as a grey point. The azimuthal positions show the correlation coefficient  $R$  between the two fields. The standard deviation for the test field  $\sigma_f$  is proportional to the radial distance from the origin, with the solid dashed arc as the reference standard deviation  $\sigma_r$ . Finally, the centered root-mean-squared difference  $RMS_{diff}$  between the test and reference patterns is proportional to the distance to the reference point, and the arcs indicate its value. The diagram allows us to determine the ranking of the *test* fields by comparing the distance to the reference point.

In Figure 2, two test fields are represented as an example: Example 2 has a correlation  $\pm 0.99$ , a  $RMS_{diff} \pm 0.24$ , and a  $\sigma_f \pm 1.15$ . Example 1 has a correlation  $\pm 0.48$ , a  $RMS_{diff} \pm 1.40$ , and a  $\sigma_f \pm 1.55$ . Example 2 performs better than Example 1 since all the metrics are better. The diagram shows this in a visual way as Example 2 is closer to the reference point.

$$\sigma_f^2 = \frac{1}{N} \sum_{n=1}^N (f_n - \bar{f})^2 \quad (1)$$

$$\sigma_r^2 = \frac{1}{N} \sum_{n=1}^N (r_n - \bar{r})^2 \quad (2)$$

$$R = \frac{\frac{1}{N} \sum_{n=1}^N (f_n - \bar{f})(r_n - \bar{r})}{\sigma_f \sigma_r} \quad (3)$$

$$RMS_{diff} = \left[ \frac{1}{N} \sum_{n=1}^N [(f_n - \bar{f}) - (r_n - \bar{r})]^2 \right]^{1/2} \quad (4)$$

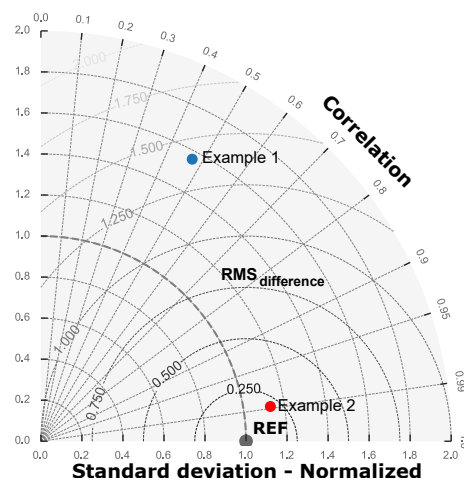


Figure 2. A Taylor diagram example.

## 2.2. Energy Analysis Methodology

The study of the impact on the energy demand due to the use of two different *actual* weather datasets is based on the energy simulation, and therefore, building energy models (BEM) are needed. For this study, detailed BEMs are employed using the EnergyPlus engine [52,53]. BEMs require weather files in the EPW format, which are created using the weather data from both sources (*on-site* and *third-party* weather data) and employing the Weather Converter tool [54] provided as an auxiliary program by EnergyPlus.

As shown in Figure 1, the energy analysis studies the impact on the energy demand when using both *on-site* and *third-party* weather files. BEM provides the energy demand that it requires each weather file to accomplish with the defined requirements (the temperature setpoint of each space). The energy results using the *on-site* weather file are established as the reference as they corresponded with the weather data measured in the building's surroundings. Comparing the variation between the energy demand results using both weather files allows us to analyze the impact of using weather data from *third-party* sources with respect to the reference.

In order to perform a deeper study, a sensitivity analysis is performed to analyze the effects on the energy demand generated by each weather parameter. The methodology consists of replacing variables (one variable at a time) from the *on-site* weather file with data from the *third-party* weather file and generating specific weather files for each parameter. For example, when the dry bulb temperature is analyzed, a weather file is prepared that has the dry bulb temperature data from the *third-party*, but the rest of the weather parameters are maintained the same as in the *on-site* weather file. This way, the impact on energy demand when using only dry bulb temperature data from a *third-party* can be studied. This procedure is done for the weather parameters analyzed in the weather comparison: the dry bulb temperature (*Temp*), relative humidity (*RH*), direct normal irradiation (*DNI*), diffuse horizontal irradiation (*DHI*), wind speed (*WS*), and wind direction (*WD*). These weather parameters are selected for the study as they are all used by EnergyPlus in the simulations, unlike other parameters, such as the global horizontal irradiation [54].

In the case of the other weather parameters provided by the weather stations, such as the atmospheric pressure and precipitation, they are not presented in this study as their impact in the BEMs is low. The process to perform the energy analysis is the same as before: BEM is simulated with the generated weather file with one parameter changed, obtaining an energy demand that is compared to the reference (the energy demand obtained using the *on-site* weather data).

As was explained in the Introduction, most of the studies that analyzed the effect of using different weather datasets in the building energy simulations used only the annual energy results, and only some of them used smaller temporal resolutions (monthly or weekly). This study presents the analysis according to different temporal resolutions and discusses the differences in the results. The time granularity levels proposed are annual, seasonal, monthly, weekly, daily, and hourly. Thus, the uncertainty metrics calculated for the energy results are related to the accumulated energy demand provided by the model in year, season, month, week, day, and hour periods.

For the statistical analysis of the results, three metrics are used in the study: the mean absolute deviation percent (*MADP*) (5), the coefficient of variation of the root-mean-squared error (*CV(RMSE)*) (6), and the coefficient of determination ( $R^2$ ) (7). The equations of these statistical indexes are shown as:

$$MADP = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{\sum_{i=1}^n |y_i|} \quad (5)$$

$$CV(RMSE) = \frac{1}{\bar{y}_i} \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y})^2}{n - p}} \quad (6)$$

$$R^2 = \left( \frac{n \cdot \sum_{i=1}^n y_i \cdot \hat{y} - \sum_{i=1}^n y_i \cdot \sum_{i=1}^n \hat{y}}{\sqrt{(n \cdot \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2) \cdot (n \cdot \sum_{i=1}^n \hat{y}^2 - (\sum_{i=1}^n \hat{y})^2)}} \right)^2 \quad (7)$$

In the equations,  $n$  is the number of observations,  $y_i$  the *on-site* measured data at moment  $i$ , and  $\hat{y}_i$  the *third-party* value at that moment.

*MADP* and *CV(RMSE)* are both quantitative indexes that show the results in percentage terms. They allow the comparison between different test sites, weather parameters, and time resolutions. *MADP*, which is also called the MAD/mean in some studies [55], has advantages that overcome some shortcomings of other metrics. It is not infinite when the actual values are zero, is very large when actual values are close to zero, and does not take extreme values when managing

low-volume data [55–57].  $CV(RMSE)$ , which gives a relatively high weight to large variations, is the other percentage metric selected for this study because it is a common metric in energy analysis. Indeed, the ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) Guidelines [44], FEMP (Federal Energy Management Program) [58], and IPMVP (International Performance Measurement and Verification Protocol) [59] use it to verify the accuracy of the models.

The coefficient of determination ( $R^2$ ) allows us to measure the linear relationship of the two patterns [60]. It ranges between 0.00 and 1.00, and higher values are better. It should be noted that uncertainty cannot be assessed using only this metric as the linear relationship may be strong, but with a substantial bias.

In the study, the  $MADP$  and  $CV(RMSE)$  metrics are shown for all the temporal resolutions, from annual to *hour*. However,  $R^2$  is only analyzed for the hourly time grain as the study of the linear relationship of larger time grains variations, which has few points, is meaningless.

### 2.3. Indoor Temperature Analysis Methodology

The third comparison analysis studies the impact of using the two weather datasets (*on-site* and *third-party* weather files) for the building's indoor temperature. In order to allow the temperature comparison, the energy used by each model is fixed. In other words, both simulations with *on-site* and *third-party* weather data use the exact same energy; however, due to the differences in the weather parameters, the indoor temperature is different. The methodology consists of performing the first simulation with the *on-site* weather file to obtain the baseline energy demand for each thermal zone of the model. This baseline energy demand is then injected into the model using an EnergyPlus script for an HVAC machine that distributes that energy in each thermal zone. Then, the model is simulated for both the *on-site* and *third-party* weather files. The results of the building temperature—unifying thermal zone temperature, weighing it with respect to its volume—of these two last simulations are compared to analyze the impact on the indoor temperature conditions.

In this case, two quantitative indexes (mean absolute error  $MAE$  (8) and root-mean-squared error  $RMSE$  (9)) and a qualitative index ( $R^2$  (7)) are used to quantify the variation in the shape of the temperature curves.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (8)$$

$$RMSE = \left[ \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right]^{\frac{1}{2}}. \quad (9)$$

The  $MAE$  and  $RMSE$  indexes are used to determine the average variation of the indoor temperature when using the different weather files in the simulations [61]. Both measure the average magnitude of the variation in the units of the variable of interest and are indifferent to the direction of the differences, overcoming cancellation errors. However,  $RMSE$  gives a relatively high weight to large deviations [62–64].  $RMSE$  will always be greater than or equal to  $MAE$  (due to its quadratic nature); thus, the greater the difference between  $MAE$  and  $RMSE$ , the greater the variance between the individual dispersions on the sample. In this case, the three metrics are calculated for the hourly criteria.

## 3. Results

The methodology described in the previous section was applied to four buildings located in three different real test sites for a period of one year (2019). The test sites were an office building at the University of Navarre in Pamplona (Spain), a public school in Gedved (Denmark), and two buildings in the Lavrion Technological and Cultural Park (LTCP) in Lavrion, Greece: H2SusBuild and an administrative building. As shown in Figure 1, three different analyses were performed in the study: weather dataset comparison, energy comparison, and indoor temperature comparison. The following three subsections develop the three analyses.

### 3.1. Weather Data Comparison

For the analysis and comparison of the weather data, the first step was the data gathering from the *on-site* and *third-party* sources for the three locations for the period of study, which is the whole year 2019. *On-site* weather data were provided by weather stations installed in the buildings' surroundings. In Pamplona and Gedved, the weather station was installed on the buildings' roofs. In the case of Lavrion, the weather station was placed in the Technological and Cultural Park where the two test sites were located, near H2SusBuild. Table 1 shows the range, resolution, and accuracy of the sensors that formed part of each weather station. In general, the sensors of Pamplona's weather station had the best accuracy. In the case of Lavrion, the diffuse solar radiation had a manual shadow bar that required readjustment every two days.

The time period of the measured data gathered from the three weather stations is the whole year 2019. Despite the good quality of the measured weather data, the raw data contained small gaps, usually a few hours, so interpolation was performed in order to fill in the missing data. On the other hand, the Weather Converter tool, which is used to generate the weather files, allows undertaking a complementary validation of the data since it produces a warning if data out of the range are used in the weather files' generation process.

The *third-party actual* weather data for the year 2019 and for three locations were provided by meteoblue [65]. They are simulated historic data for a specific place and time calculated with models based on the NMM (nonhydrostatic meso-scale modeling) or NEMS (NOAA Environment Monitoring System) technology, which enables the inclusion of the detailed topography, ground cover, and surface cover. Further information about the computation of the weather data provided by meteoblue is available in [66].

The results of the weather parameter comparison between data from *on-site* weather stations and *third-party* (meteoblue) are shown using the Taylor diagram described in Section 2.1. Figure 3 summarizes all the results: showing the three statistics ( $R$ ,  $RMS_{difference}$ , and  $\sigma$ ) for the whole period of study (2019) calculated on an hourly basis, for the six weather parameters analyzed ( $Temp$ ,  $RH$ ,  $DNI$ ,  $DHI$ ,  $WD$ , and  $WS$ ), and for the three locations (each one represented in a different color).

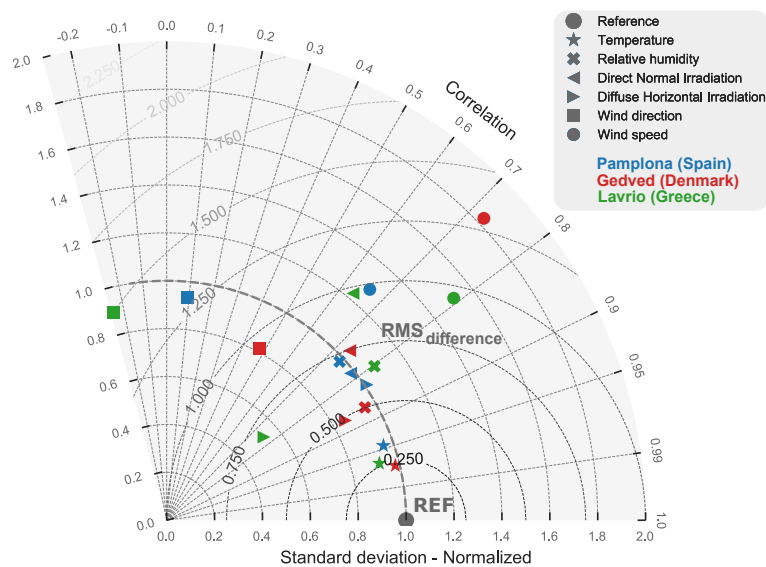
For Pamplona weather (represented in blue), the diagram shows that  $Temp$  provided the weather parameter for this location that better agreed with the *on-site* observations as it had the highest correlation  $R$  of around 0.95, the smallest  $RMS_{diff}$  ( $\pm 0.3$ ), and a very close  $\sigma_f$  (standard deviation) to the reference ( $\pm 0.95$ ).  $RH$ ,  $DNI$ , and  $DHI$  provided similar results with a correlation around 0.7–0.8, an  $RMS_{diff}$  between 0.5–0.6, and a good standard deviation. The parameters that correlated worse with the observed values were the wind parameters, especially  $WD$  ( $R = 0.09$ ).

For Gedved weather (red color), the Taylor diagram shows that  $Temp$  was the *third-party* weather parameter that agreed best with the *on-site* observations, with an  $R$  of around 0.97. As in Pamplona, the wind parameters delivered the results furthest from the reference point.  $WS$  had an acceptable correlation, but a very high standard deviation, and  $WD$  performed better for  $\sigma_f$ , but had a low correlation (less than 0.5). For the third location, Lavrion (green color),  $Temp$  also had a good correlation  $R$  (higher than 0.95).  $RH$ ,  $DHI$ , and  $WS$  had a medium  $R$  for the observed data (around 0.8), but they presented differences in the other metrics.  $RH$  had a better standard deviation than the other two, and  $RH$  and  $DHI$  had a lower  $RMS_{diff}$  than  $WS$ . In this location, the parameter that provided the worst results was  $WD$ , which had the worst  $R$  ( $-0.22$ ) and the highest  $RMS_{diff}$  (1.5).

**Table 1.** Technical specifications of the sensors of the weather stations installed in the office building in Pamplona (Spain), in Gedved School (Denmark), and in the Technological and Cultural Park in Lavrion (Greece).

Sensor	Pamplona (Spain)			Gedved (Denmark)			Lavrion (Greece)		
	Range	Resolution	Accuracy	Range	Resolution	Accuracy	Range	Resolution	Accuracy
Temperature (°C)	−50 to +60	0.1	±0.2	−40 to +60	0.1	±0.3	−40 to +65	0.1	±0.5
Relative Humidity (%)	0 to 100	0.1	±2	0 to 100	1	±2.5	0 to 100	1	±3 (0–90%) ±4 (90–100%)
Atmospheric Pressure (mbar)	300 to 1200	0.1	±0.5	150 to 1150	0.1	±1.5	880 to 1080	±0.1	±1
Precipitation (mm)	0.3 to 5.0	0.01	-	-	0.2	±2%	-	0.2	±4%/0.25 (<50 mm/h) ±5%/0.25 (>50 mm/h)
Wind Direction (°)	0 to 359.9	0.1	<3	1 to 360	1	1%	1 to 360	1	±4%
Wind Speed (m/s)	0 to 75	0.1	±3% (0–35) ±5% (>35)	1 to 96	1	0.1 (5–25)	1 to 67	0.44	±1/±5%
Global Solar Radiation (W/m <sup>2</sup> )	0 to ±1300	1	<±10%	0 to ±1300	1	<±10%	0 to 1500	1	<10
Diffuse Solar Radiation (W/m <sup>2</sup> )	0 to ±1300	1	<±10%	0 to ±1300	1	<±10%	0 to 1500	1	<20



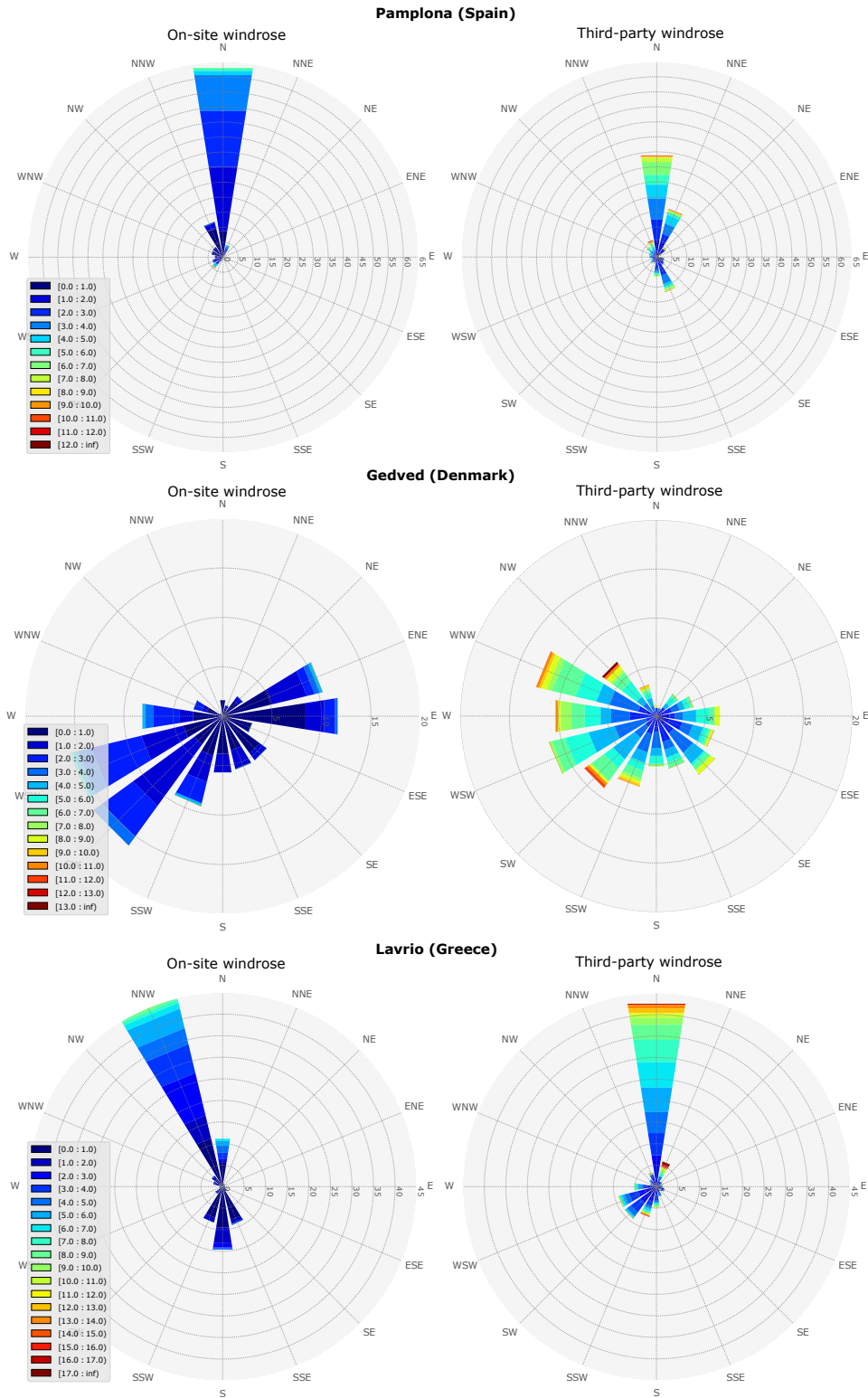


**Figure 3.** The normalized Taylor diagram for the three weather locations (Pamplona (Spain), Gedved (Denmark), and Lavrio (Greece)) compared on an hourly basis for *on-site* and *third-party* weather data for the year 2019. This shows the dry bulb temperature, relative humidity, direct normal irradiation, diffuse horizontal irradiation, wind speed, and wind direction.

Comparing the statistical results for the three locations, the performance of data provided by the *third-party* varied for each location. Gedved provided the best results for four of the six weather parameters (*Temp*, *RH*, *DHI*, and *WD*). In the three cases, *Temp* was the parameter that best matched the reference ( $R$  around 0.95,  $RMS_{diff}$  lower than 0.4, and  $\sigma_f$  near one), and *WD* was the worst (correlations lower than 0.5 and  $RMS_{diff}$  higher than 0.9). *WS* also provided poor correspondence with the observations, especially for the Gedved location. The rest of the parameters were scattered in the medium part of the diagram.

In Appendix A, a deeper analysis is shown where the statistical indexes for the monthly and seasonal data are represented in order to analyze their homogeneity. Figures A1–A3 show that the *Temp*, *RH*, *DNI*, *DHI*, and *WD* parameters for the three weathers were quite homogeneous, with the seasonal and monthly indexes quite concentrated, providing similar  $R$ ,  $RMS_{diff}$ , and  $\sigma_f$ . There are some exceptions, such as *DNI* for November in Pamplona and January in Gedved, which agreed worse with the observations than for the rest of the months. On the other hand, *WS* had more heterogeneous monthly and seasonal results since more scattered points were seen in the diagrams. In general, the winter and autumn months correlated the worst with the observed data.

Since the wind parameters produced higher variations when comparing *on-site* and *third-party* weather datasets, a deeper comparison analysis was performed using wind rose diagrams (see Figure 4). This diagram shows the distribution of the wind speed and wind direction. The rays point to the direction the wind is coming from, and their length indicates the frequency in percentage. The color depends on the wind speed, growing from blue to red colors. Pamplona's wind rose shows that *WS* from the *third-party* data was much higher than observed, and although the prevailing direction was north in both cases, there were important differences in the frequency percentages. For Gedved, the *third-party* data provided much higher wind speed (yellow to red colors in the wind rose) than the observed data (blue colors) and a different prevailing wind direction. Finally, the wind roses for Lavrion show differences in the prevailing wind direction and very different wind speeds, being higher in the *third-party* wind rose.



**Figure 4.** Wind rose comparison between the weather station and *third-party* data of the Pamplona (Spain), Gedved (Denmark), and Lavrio (Greece) locations.

### 3.2. Energy Analysis Results

After analyzing the variations in the different weather parameters between the *on-site* and *third-party* weather data for the three weathers, the following analysis consisted of the study of the impact produced by these variations in the test sites' building energy demands using detailed BEMs. Figure 5 shows

the four test sites analyzed in this study showing a real image from the building and an image of the performed model in EnergyPlus (in colors for the different thermal zones of each building).



**Figure 5.** The test sites analyzed. From left to right: office building (University of Navarre, Spain); Gedved school (Denmark); H2SusBuild and administration building in Lavrio (Greece). On top, the real buildings and, on the bottom, the building energy models (SketchUp thermal zone representation, OpenStudio plugin [67]).

The first test site was the office building attached to the Architecture School at the University of Navarre in Pamplona (Spain). It hosts administration uses and classrooms for the postgraduate students. This building is a 755 square meter single-story building built in 1974. It has a concrete structure; the outdoor walls are built of red brick fabric ( $U$  value =  $0.3 \text{ W/m}^2\text{K}$ ); the flat roof has the insulation above the deck ( $U$  value =  $0.2 \text{ W/m}^2\text{K}$ ); and aluminum window frames were installed in situ with an air chamber. The Gedved public school (Denmark) consists of six buildings and was built in 1979 and then renewed in 2007. The library of one of the school buildings was selected as the test site. It is a one-story building with a total surface area of  $1138 \text{ m}^2$ , with a big central space—the library—and nine classrooms and serving spaces around it. The building walls consist of two brick layers with 150 mm insulation in between ( $U$  value =  $0.27 \text{ W/m}^2\text{K}$ ). The windows are two-layer double-glazed windows with cold frames. The ceiling is insulated with 200–250 mm mineral wool for the sloping and flat ceiling, respectively ( $U$  value =  $0.07 \text{ W/m}^2\text{K}$  and  $0.16 \text{ W/m}^2\text{K}$ , respectively), and the floor is made of concrete and contains 150 mm insulation under it ( $U$  value =  $0.21 \text{ W/m}^2\text{K}$ ).

In Lavrion (Greece), two buildings from the Lavrion Technological and Cultural Park (LTCP) were used as test sites: H2SusBuild and the administration building. H2SusBuild has a ground floor and an attic floor with a total surface area of approximately  $505 \text{ m}^2$ . The ground floor hosts a small kitchen, toilets, the control room, and the main area. The attic also hosts two offices and a meeting room. Its envelope consists of a concrete structure with double concrete block walls and single-glazed windows with aluminum frames. It also has external masonry consisting of double brick walls with 10 cm expanded polystyrene (EPS) insulation ( $U$  value =  $0.25 \text{ W/m}^2\text{K}$ ). The roof consists of metallic panels with a 2.5 cm polyurethane insulation layer in the middle ( $U$  value =  $0.75 \text{ W/m}^2\text{K}$ ). The administration building hosts the LTCP managing authority and administrative services. It is a two-story renovated neoclassic building with a surface area of about  $644 \text{ m}^2$ . The building envelope is made of stone approximately 70 cm thick ( $U$  value =  $1.85 \text{ W/m}^2\text{K}$ ) with wooden-framed double-glazed large windows. The roof consists of a wooden frame with gutter tiles placed on top ( $U$  value =  $0.49 \text{ W/m}^2\text{K}$ ).

For each building, an individual pattern of use corresponding to the actual use of the building was implemented in the simulation model. Each building had its own calendar of use, occupancy, and internal loads of electric equipment and lighting. Regarding the HVAC systems, setpoints and usage hours were defined for each. The office building in Pamplona and H2SusBuild and administration building in Lavrion implemented both heating and cooling systems in the models, and the Gedved school had only a heating system. Table 2 shows the input data of the four models.

**Table 2.** Input data of the four models.

	Office Building Pamplona (Spain)	Public School Gedved (Denmark)	H2SusBuild Lavrion (Greece)	Administration Building Lavrion (Greece)
Lighting (W/m <sup>2</sup> )	10	10	8.5	8.5
Equipment (W/m <sup>2</sup> )	8	15	8	8
Occupation schedule	Wd 9–21 h/Sat 9–14 h	Wd 8–16 h/Sat 8–13 h	Wd 9–20 h/Sat 9–14 h	Wd 9–20 h/Sat 9–14 h
Heating setpoint (°C)	20	Day 21/Night 15.6	21	21
Cooling setpoint (°C)	26	No cooling	24	24

Wd: weekdays. Sat: Saturdays.

The results of the statistical analysis for the energy study are presented in Table 3. The table is divided into four sub-tables, one for each test site. They show the three uncertainty metrics calculated for the energy demand obtained from simulations using the *on-site* and *third-party* weather files (*TPW*), with *on-site* as the reference. The first column of each test site's table, designated as *TPW*, shows the difference in percentage (*MADP*) of the energy demand when the third-party weather file is used in the simulation with respect to the reference simulated with *on-site* weather data. With the inputs shown in Table 2, the models provide the following annual energy demand: 21.9 kWh/m<sup>2</sup> for the office building, 91.5 kWh/m<sup>2</sup> for the Gedved school, 142.2 kWh/m<sup>2</sup> for H2SusBuild, and 91.6 kWh/m<sup>2</sup> for the administration building.

The table allows performing two different analyses depending on how it is read. The vertical interpretation of the table shows the percentage results as a function of the time resolutions (from annual to hourly criteria) employed for the analysis. On the other hand, horizontally, the variations in the energy demand for the sensitivity analysis changing only one weather parameter at a time are presented (*DHI*, *DNI*, *RH*, *Temp*, *WD*, and *WS*).

**Table 3.** Uncertainty metrics (*MADP*, *CV(RMSE)*, and *R*<sup>2</sup>) were used in the energy analysis for the four test sites. *TPW*: the results using the weather file with all the parameters from the *third-party* weather data source. The rest changed only one parameter at a time: *DHI* (diffuse horizontal irradiation), *DNI* (direct normal irradiation), *RH* (relative humidity), *Temp* (temperature), *WD* (wind direction), and *WS* (wind speed).

		Office Building, Pamplona (Spain)							School, Gedved (Denmark)						
Statistic Index	Time	TPW	DHI	DNI	RH	Temp	WD	WS	TPW	DHI	DNI	RH	Temp	WD	WS
<i>MADP</i> (%)	Year	1.42	0.79	4.33	0.50	4.07	0.03	7.95	43.82	0.82	6.24	1.16	0.68	0.01	46.91
	Season	9.61	3.93	9.01	3.17	4.07	0.03	11.74	43.82	0.82	6.24	1.16	1.71	0.01	46.91
	Month	18.14	4.65	10.29	2.97	12.07	0.03	13.60	43.82	0.82	6.24	1.16	3.59	0.01	46.91
	Week	17.88	5.24	11.17	3.40	14.39	0.04	13.99	43.82	0.90	6.24	1.19	5.16	0.02	46.91
	Day	26.03	6.25	11.80	4.02	22.53	0.05	14.08	44.41	0.98	6.25	1.37	8.78	0.02	46.91
	Hour	29.96	6.64	12.27	4.71	25.58	0.08	14.31	45.17	1.03	6.27	1.50	10.10	0.02	46.91
<i>CV(RMSE)</i> (%)	Year	1.42	0.79	4.33	0.50	4.07	0.03	7.95	43.82	0.82	6.24	1.16	0.68	0.01	46.91
	Season	10.78	4.72	10.11	4.00	4.54	0.03	15.48	58.42	0.95	7.61	1.49	2.25	0.02	60.25
	Month	26.12	5.55	13.15	4.02	14.61	0.04	17.96	52.44	0.99	7.05	1.41	4.39	0.02	54.62
	Week	26.72	7.12	15.04	5.17	19.68	0.06	20.01	54.15	1.19	7.10	1.45	7.55	0.02	55.29
	Day	42.90	9.29	18.18	7.60	34.96	0.08	22.07	60.39	1.44	7.74	1.86	13.11	0.03	59.33
	Hour	61.42	13.14	22.37	12.56	50.71	0.20	30.55	91.79	2.75	14.04	3.54	24.99	0.05	88.81
<i>R</i> <sup>2</sup> (%)	Hour	90.86	99.48	98.64	99.53	92.54	100.00	98.43	95.18	99.99	99.82	99.98	98.55	100.00	96.84
		H2SusBuild, Lavrion (Greece)							Administration Building, Lavrion (Greece)						
		TPW	DHI	DNI	RH	Temp	WD	WS	TPW	DHI	DNI	RH	Temp	WD	WS
<i>MADP</i> (%)	Year	32.86	5.95	5.22	1.97	5.21	1.05	39.40	1.29	11.40	7.27	3.45	9.62	1.93	12.19
	Season	44.58	8.14	9.49	3.21	7.25	2.16	39.40	27.75	13.83	13.40	4.20	9.62	2.73	21.63
	Month	45.83	10.32	11.20	3.95	11.82	2.94	41.66	38.70	15.56	15.80	4.41	14.73	2.99	29.97
	Week	47.80	10.50	11.15	4.05	13.02	2.93	42.14	38.70	15.69	15.80	4.52	14.78	2.99	29.97
	Day	49.46	10.75	11.40	4.46	17.17	2.94	42.49	38.91	15.72	15.80	4.91	16.60	3.00	29.97
	Hour	51.58	10.93	11.85	5.12	19.97	2.95	43.95	39.45	15.79	15.83	5.49	17.88	3.01	30.22
<i>CV(RMSE)</i> (%)	Year	32.86	5.95	5.22	1.97	5.21	1.05	39.40	1.29	11.40	7.27	3.45	9.62	1.93	12.19
	Season	61.65	9.69	10.71	3.91	12.40	2.58	63.61	36.21	15.57	14.26	4.78	15.61	2.93	34.39
	Month	65.24	13.41	13.37	5.04	15.66	3.44	65.89	43.53	20.89	18.63	6.20	20.23	3.83	38.06
	Week	72.40	14.11	13.97	5.46	19.34	3.52	73.29	46.18	21.35	19.09	6.62	21.59	3.85	40.61
	Day	82.65	14.77	14.74	6.37	24.90	3.68	81.45	49.81	21.93	19.63	7.76	24.17	3.97	43.01
	Hour	90.37	18.35	17.92	8.83	34.68	4.85	85.67	90.81	37.52	33.69	15.03	47.91	6.67	72.69
<i>R</i> <sup>2</sup> (%)	Hour	85.58	98.41	98.43	99.61	93.76	99.88	92.43	81.85	97.30	97.80	99.51	94.91	99.92	91.18

The first analysis obtained from Table 3 was the influence of the time resolution used in the study of the energy demand variation. In this case, the analysis was done in the vertical from the annual to hourly criteria. The percentage metrics  $MADP$  and  $CV(RMSE)$  allowed us to compare the results for each time grain and study its influence in the results. Both indexes were closely related; however,  $CV(RMSE)$  gave a relatively high weight to large variations. It is remarkable that the differences between  $CV(RMSE)$  and  $MADP$  decreased as the time grain increased (from hourly to annual criteria) as, when the energy demand was accumulated, the outliers were minimized. The results for the hourly basis showed that the  $CV(RMSE)$  values were around twice the  $MADP$  values for the four sites and all the weather parameters. This indicates that significant outliers were present in the energy demand results when both simulations based on the *on-site* and *third-party* weather datasets were compared.

On the other hand, both the  $MADP$  and  $CV(RMSE)$  metrics showed how, in the four test sites, the variations in the energy demand grew as the time grain decreased, which matches with ASHRAE's statement about the energy data granularity [44]. If the results were analyzed with the accumulated energy demand for a period of time (i.e., monthly, annual, etc.), the energy variation was minimized with respect to the hourly analysis. For example, differences of  $MADP$  up to  $\pm 38\%$  between the annual and hourly criteria are seen in the results for the administration building. In this case, the  $MADP$  for the accumulated data for the year was only 1.29%; thus, the annual building energy demand simulated for the *third-party* weather file was very similar to the reference, simulated with the *on-site* weather file.

However, for the hourly basis, this variation grew up to 39.45%, which is a significant deviation. The reason is because, alternately, in some cases, the model simulation overestimated the energy demand needed by the building (the model demanded more energy than the reference), and in other cases, the model underestimated it. When the data were accumulated from the hourly basis to longer periods of time (daily, weekly, monthly, seasonal, and annual), a compensation effect occurred by canceling each other out, which resulted in the minimization of the energy demand variation. As the length of the periods increased, so did the compensation effect and, therefore, also the minimization of the variations.

It is also remarkable that for all the test sites, the  $CV(RMSE)$  results showed high values for the monthly and hourly resolutions, which are the time criteria commonly used by the energy analysis standards.

The second analysis was the study of the influence of each weather parameter in the energy demand variation. In this case, the interpretation of the tables from 3 was done horizontally: the first column presents the results for the simulation with the *third-party* weather file (*TPW*), which had all the weather parameters changed, and the following columns show the results for the different weather parameters.

Comparing the results of the four test sites using the  $MADP$  and  $CV(RMSE)$  indexes, some common observations can be made. In all of them, the weather parameter that generated less impact in the simulated energy demand was *WD*, even though it was the weather parameter that worst fit the *on-site* weather data, as was shown in the Taylor diagram (Figure 3). The reason is because the mechanical ventilation and infiltration EnergyPlus objects used in these models did not account for *WD* in the simulations [68].

On the other hand, in the four test sites, when *WS* was analyzed, it showed an important impact in the energy demand simulations' outputs. This was mainly due to two causes: The first was the use of dynamic infiltrations introduced by using the EnergyPlus object *ZoneInfiltration:EffectiveLeakageArea*, which took into account the *WS* parameter in the calculations [68]. The leakage area in  $\text{cm}^2$  was calculated by the calibration process previously developed by the authors [69–71]. The second was because the differences between the *third-party* *WS* data and *on-site* data (see the Taylor diagram in Figure 3 and the wind roses from Figure 4) were significant.



In both the Gedved school and H2SusBuild, the *third-party* wind speed provided faster values, which generated a significant increase in the energy demand during almost all the year, but there were a few moments with a decrease in the energy demand. Therefore, the compensation effect between the overestimated and underestimated energy demand was reduced. This explains why the variation due to *WS* was high for all the time grains for these test sites. This effect was especially clear in the Gedved school, which did not have a cooling system. In this case, all the time grains for *WS* provided the same *MADP* because the higher *WS* of the *third-party* data always meant a higher heating energy demand and no energy demand compensation existed.

Regarding the *Temp* parameter, in the weather data analysis (Section 3.1), based on an hourly time grain, it was the parameter with less variation between *on-site* and *third-party* data for the three sites. However, the *MADP* and *CV(RMSE)* results, especially for the hourly criteria, showed that it had a significant impact on the energy demand in the four test sites. It was the second parameter of influence for the Gedved school, H2SusBuild, and administration building after wind speed and the first one in the office building with an *MADP* of  $\pm 26\%$ . In relation to the solar irradiation, the Taylor diagram (Figure 3) showed that *DHI* from the *third-party* weather data provided a better correlation than *DNI* for the three locations, and this was reflected in the sensitivity energy analysis. For the Gedved school, these two parameters had less impact on the energy demand than for the other three models. The reason is because the school lacked a cooling system; therefore, in summer, when more solar access was available, no energy demand was taken into account.

To conclude the explanation of Table 3, factor  $R^2$  was analyzed. It compared the shape of the energy demand curves from the different simulations and showed that the energy demand simulated with the *third-party* weather data fit quite well with the energy demand simulated from the *on-site* data for the four test sites (with  $R^2$  between  $\pm 82\%$  and  $\pm 95\%$ ). Regarding the different weather parameters, the results for each parameter matched with the analysis of the hourly percentage indexes. The parameters with lower hourly *MADP* and *CV(RMSE)* values had higher  $R^2$  values.

Finally, to show in a visual way the previous analysis of the influence of the time resolution employed in the study and the sensitivity of each weather parameter, the *MADP* index results are plotted in Figure 6. Each graph presents the *MADP* result for each test site. In the graphs, the six time resolutions are shown on the *x* axis, and the dashed line presents the results for the simulation with all the *third-party* weather parameters (*TPW*). Each color represents the results for each weather parameter of the sensitivity analysis. The graphs show how the variations in the energy demand grew as the time grain decreased, especially *Temp*. They also show that *WS* was the most sensitive weather parameter for the Gedved school, H2SusBuild, and administration building. Only in the case of the office building in Pamplona was *WS* the most sensitive weather parameter taking into account an annual criterion; however, per hour, it changed to *Temp*.

Previous analyses showed the significant variations in the energy demand when using different *actual* weather datasets. In order to study if these differences in the energy demand were mostly due to the building architectures or to the weather, a complementary theoretical analysis was performed, and this is presented in the following section.

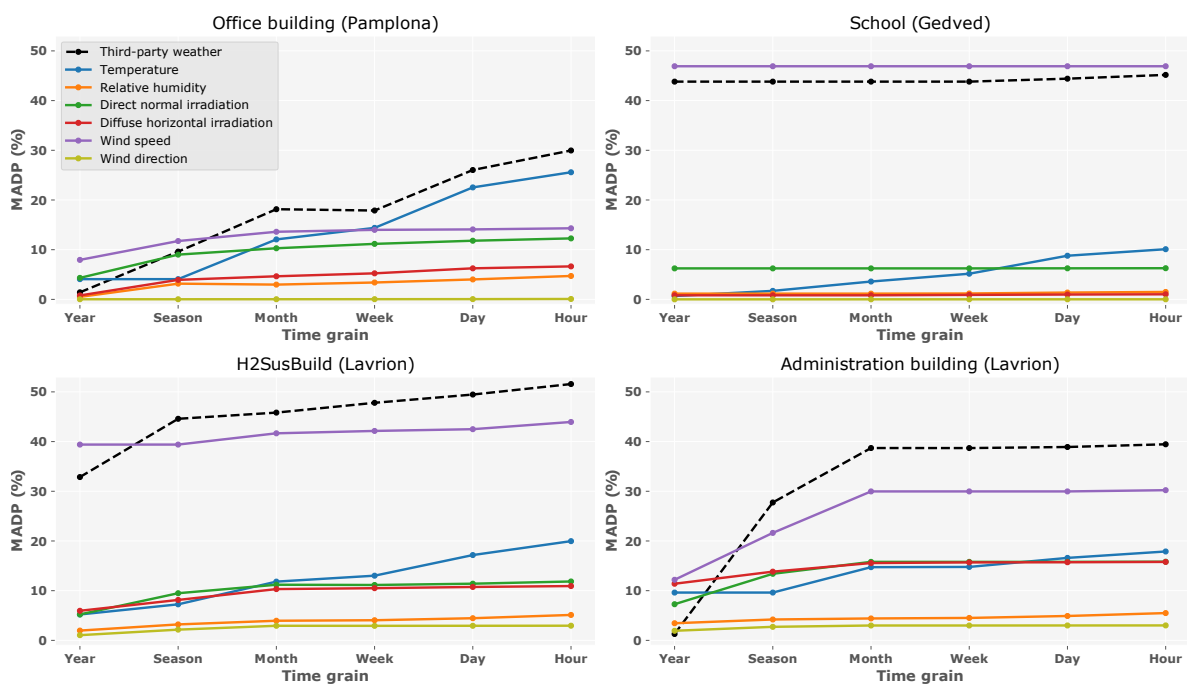
### 3.2.1. Analysis of the Buildings' Architecture Influence on the Energy Results

Since four test sites were available for this research, a complementary study was performed to analyze the influence of the building's architecture on the previous energy results. The four buildings, which were completely different in terms of the materials, construction systems, thermal mass, and window-to-wall ratio, were simulated with the same weather data (*on-site* and *third-party* weather files). For this study, we selected the most homogeneous weather when comparing the *third-party* to the *on-site* weather data: as shown in the previous analysis, energy demands are very sensitive to *WS*, so the Gedved weather was discarded for this analysis because its *WS* was the one with the worst fit to the reference (see Figure 3). *Temp* was also a sensitive parameter, and the three weathers had similar statistical metrics. Finally, for the solar radiation parameters *DHI* and *DNI*, Pamplona's

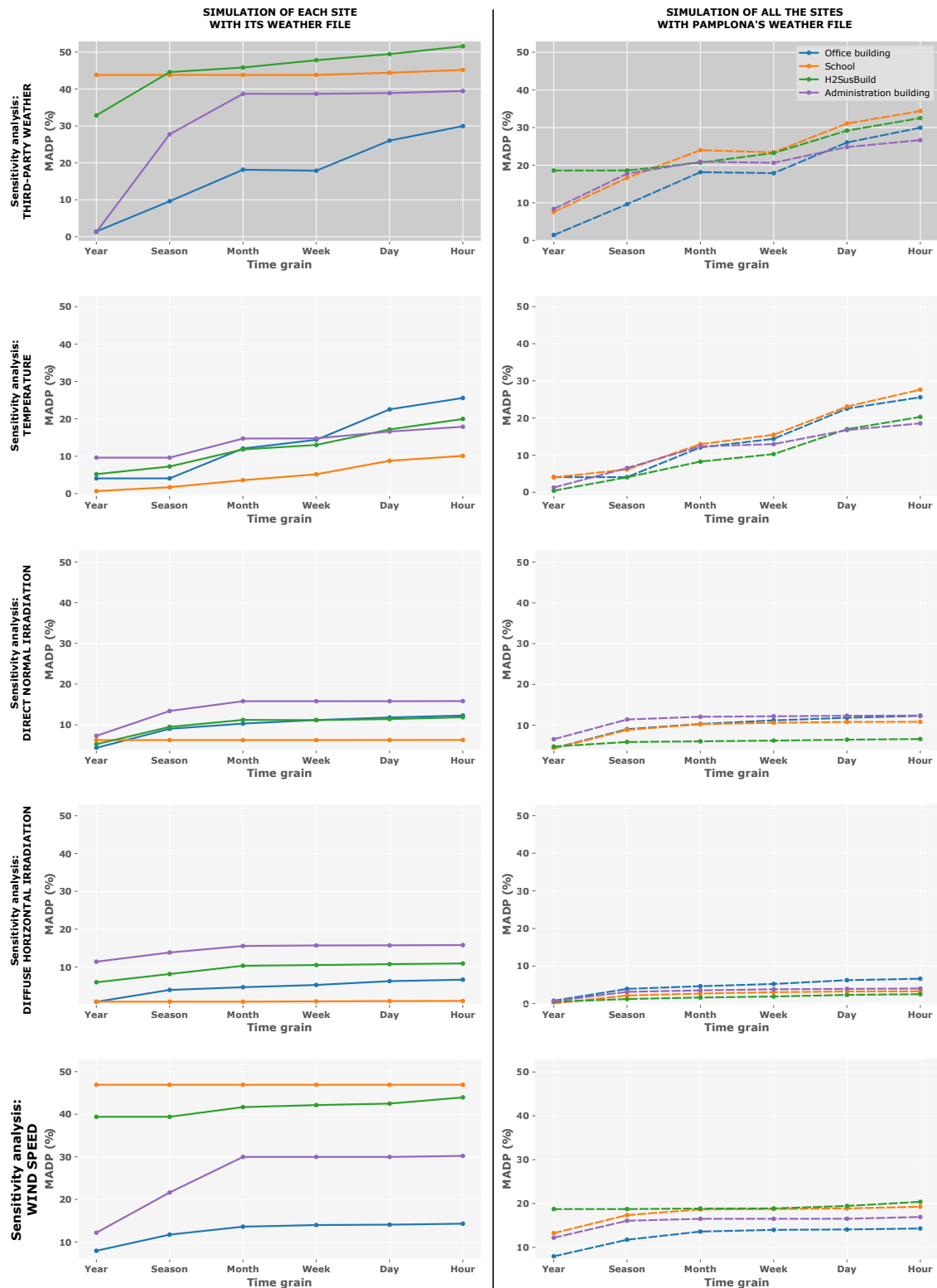
weather better matched the reference compared to Lavrion. Therefore, the Pamplona weather file was chosen to develop this theoretical study, and for this reason, the four models were configured to have the same internal loads, HVAC systems, and schedules as the Pamplona office building.

Figure 7 shows the *MADP* results for this study. The two graphs on the top present the results using the *third-party* weather file, which had all the weather parameters provided by the weather service. They show how when the test sites were simulated with their own weather files (graph on the left), the *MADP* results and the trend of the curve were very different for the four test sites (each colored line represents one test site). However, when the test sites were simulated with Pamplona’s weather file (graph on the right with dashed lines), the curves became very similar, reducing the differences in the *MADP* values and in the trend of the curve.

Thus, the main value responsible for the variation in the energy demand was the weather dataset employed in the simulations and not the building’s characteristics. This effect was also reflected in the results from the sensitivity analysis, which are also presented in Figure 7 for the *Temp*, *DNI*, *DHI*, and *WS* parameters. The curves from the graphs on the right, which are the simulations of all the test sites with Pamplona’s weather file, were very similar compared to the curves from the graphs on the left, especially in the case of wind speed. This study demonstrated the great influence of the weather parameters on the variation of the building’s energy demand, almost independently of the model, and this showed the importance of the selection of the weather dataset used in the BEM simulations.



**Figure 6.** Representation of the *MADP* (%) of the energy demand analysis for the four test sites and for the different time grains. The dashed black line represents the results using the weather file with all the *third-party* weather parameters. Each colored line represents each one of the weather parameter results from the sensitivity analysis.



**Figure 7.** The comparison between the simulation results when each test site was simulated with its weather (on the left with continuous lines) and when all the test sites were simulated with Pamplona’s weather (on the right with dashed lines). Each color represents a test site. The graphs show the MADP for the energy demand for the different temporal resolutions. From above to below are shown: results when the *third-party* weather file was used (all the parameters were changed in the weather file) and the sensitivity analysis results for temperature, direct normal irradiation, diffuse horizontal irradiation, and wind speed.

### 3.3. Indoor Temperature Analysis Results

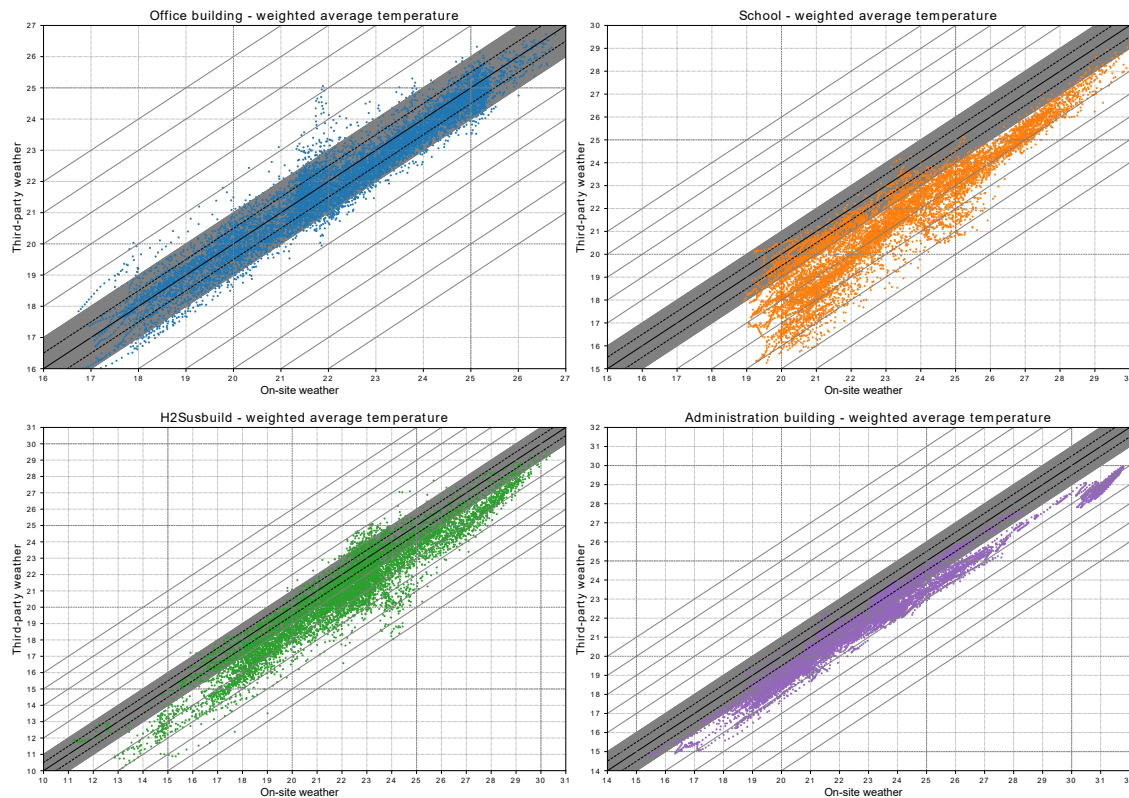
In the indoor temperature study, *MAE* and *RMSE* for the hourly criteria were used in the analysis in order to measure the quantitative variation in the temperature curve, and  $R^2$  was used to study the deviation in the temperature curve form. The results are shown in Table 4. In this case, the same energy was injected into the model for the two simulations, using the *on-site* and *third-party* weather datasets. The comparison of the thermal zones' temperature provided by the simulations provided the influence of the weather dataset employed in the indoor temperature conditions. In the table, the results are shown using three criteria: (1) for all the thermal zones (*All*), where the statistical metrics are calculated using the indoor temperature of all the thermal zones of the building; and (2, 3) for the maximum (*Max*) and minimum (*Min*) temperature, where the metrics are calculated using only the temperature of the thermal zone that provides the maximum/minimum temperature in each time step in order to compare the effect in the internal healthy conditions when using the two weather data sources. For this study, the statistical metrics were calculated only for the hourly criteria. The rest of the time grains were not considered since, for the temperature analysis, the data were not accumulated when longer periods were analyzed.

This study provided similar results to the previous energy results. Regarding the results for the temperature of all the thermal zones (*All*), the high results in the  $R^2$  for the four test sites (from  $\pm 87$  to  $\pm 95\%$ ) showed that the shape of the indoor temperature curve was very similar when the two weather datasets were used in the simulation. However, the quantitative statistical metrics showed a significant impact on the indoor temperature. The Gedved school was the test site with the highest *MAE* ( $1.72\text{ }^\circ\text{C}$ ), in line with the energy analysis where this test site reached deviations in the energy demand of  $\pm 45\%$ . The main reasons why the Gedved school had a higher impact on the indoor temperature were the influence of the lack of correlation of the wind speed between the *on-site* and *third-party* weather datasets and the way infiltrations were simulated based on the leakage area. The office building in Pamplona was the site with a minor impact on the indoor temperature (*MAE* of  $0.55\text{ }^\circ\text{C}$ ) when the weather dataset was changed. When the maximum and minimum temperatures reached in the building were employed in the analysis (*Max* and *Min* in the table, respectively), it can be seen that the statistical metrics were similar to *All*. This means that the variation in the indoor temperature when using third-party weather data was stable and produced similar variations when the indoor conditions were minimum or maximum.

**Table 4.** The statistical metrics (*MAE*, *RMSE*, and  $R^2$ ) used in the indoor temperature analysis for the four test sites. *All*: metrics calculated with the temperature of all thermal zones; *Max/Min*: metrics calculated with the temperature of the thermal zone with the maximum/minimum temperature in each time step.

Statistic Index	Office Building—Pamplona	School—Gedved	H2SusBuild—Lavrion	Administration Building—Lavrion
<i>MAE</i> ( $^\circ\text{C}$ )—Min/All/Max	1.03/0.55/0.62	2.00/1.72/1.38	1.32/1.52/1.36	1.07/1.21/1.53
<i>RMSE</i> ( $^\circ\text{C}$ )—Min/All/Max	0.76/0.73/0.81	2.23/1.92/1.50	1.65/1.93/1.70	1.38/1.50/2.00
$R^2$ (%)—Min/All/Max	91.86/92.20/89.53	85.41/89.39/95.51	90.21/86.89/85.58	95.90/94.74/86.81

There was no high differences between the *MAE* and *RMSE* results for the four cases, which indicated that the variations in the indoor temperature were quite homogeneous with no significant outliers. Figure 8 shows, in a visual way, the results with a scatter plot for each test site where the temperature was weighted by a thermal zone volume of air. The office building in Pamplona (above left) had fewer scattered temperature points since they were closer to the black line than the rest of the cases, and most of the points had a difference of  $1\text{ }^\circ\text{C}$  or less. On the other hand, the Gedved school (above right) provided the worst correlation for almost all the temperature points with a difference bigger than  $1\text{ }^\circ\text{C}$  due to the difference between both weather files.



**Figure 8.** Scatter plots for each test site of the indoor temperature simulated using the *on-site* and *third-party* weather files.

#### 4. Discussion

This paper shows how to study the impact of using two different *actual* weather datasets on building energy model simulations (weather data for the year 2019): one weather dataset with data measured in the building's surroundings (*on-site*), which was considered the reference weather data; and the other supplied by a weather service provider (*third-party*). Four test sites with different uses and architecture characteristics, located in three different locations, were employed in the studMor diagram, (2) an energy demand comparison, and (3) an indoor temperature comparison. In the case of the energy approach, a sensitivity analysis of the main weather parameters was also performed to study the influence of each parameter in the energy simulation. The energy results were provided with a different temporal resolution from the annual to hourly criteria in order to highlight the differences in the results.

The results for the energy analysis showed that as the time grain decreased, the impact of using different weather datasets grew, which agrees with ASHRAE [44]. The differences between the annual and hourly *MADP* until 38% are shown in the results. This was because when the energy demand was accumulated in periods of time longer than an hour, the variation in the results was minimized due to the compensation effect of the underestimated and overestimated energy use. This must be taken into account when a weather data source is chosen according to its purpose. For instance, if the weather data will be used for model calibration purposes, it is important to take into consideration the monthly or hourly criteria, as the most used standards (ASHRAE [44], *FEMP* [58], and *IPMVP* [59]) employ these time grains for their recommendations, and as the study showed, the use of different weather datasets had a significant impact on the *CV(RMSE)* results. Another application of BEM where the time grain of the analysis is relevant is model predictive control, where the hourly criteria are required.

The sensitivity analysis of the main weather parameters showed the different influence that each parameter had on the energy demand variation of each test site. In this regard, the relative humidity and wind direction had little influence on the models. In the case of the wind direction, the low



influence was due to these test sites using mechanical ventilation instead of natural. On the other hand, the results also showed that the two parameters that produced a higher impact in the energy use were the wind speed and temperature. The high influence in the energy demand due to wind speed was explained by the *third-party* wind speed data having a low correspondence to the *on-site* data and because the models employed in the study used dynamic infiltrations that took into account the wind speed, instead of other models with constant values in the infiltration parameters. Therefore, the energy results for the wind speed showed that particular attention should be paid to this parameter when BEMs use dynamic infiltrations, as it has a great influence on the model's energy performance.

With the available data, the results obtained in this study suggested that for this models, some of the weather parameter data could be obtained from *third-party* weather sources, avoiding the installation of on-site sensors, as they had a low influence on the simulation results. This is the case of the relative humidity, wind direction, and even diffuse horizontal irradiation, the sensor being very expensive. On the other hand, to obtain the wind speed and outdoor temperature data, which are the weather parameters that were shown to be the most influential in the models' energy performance, we recommend the installation of an anemometer and a temperature sensor near the building. Having both *on-site* and *third-party* weather data sources would allow the verification of the data. An *on-site* sensor would also provide information regarding the micro-climate generated due to the surrounding characteristics of the building, which could be difficult to see reflected in the calculated weather data from a *third-party*.

The energy study showed that the weather dataset selected for the dynamic energy simulations had a great impact on the buildings energy performance, especially for short temporal resolutions. To emphasize the impact of the weather datasets in the building energy models, a theoretical study was performed, simulating all the test sites with the Pamplona weather file. The results showed that all the weather parameters produced similar variations in the energy demand and also a similar trend of the curve for the different time grains, independently of the model. This demonstrated the significant role played by the weather data and the importance of their correct selection when performing the building energy simulations.

In the case of the indoor temperature study, the significant impact of using different weather datasets was also shown. Although the high  $R^2$  results for the four test sites showed that the shape of the indoor temperature curves was similar when both the *on-site* and *third-party* weather files were used, the quantitative metrics demonstrated a significant influence on the indoor temperature of the test sites with a MAE higher than 1.5 °C in some cases.

This paper showed the variation in the simulation results when two different *actual* weather datasets (*on-site* and *third-party*) were employed. In future research, it would be interesting to collect the empirical energy and temperature data from the test sites to study which of the weather datasets is closer to reality when comparing the simulation results using both weather datasets and the actual energy and temperature measurements.

## 5. Conclusions

The aim of this research is to show how to study the variations in energy demand and indoor temperature when using two different *actual* weather data sources with the purpose of analyzing if the data from the sensors of an *on-site* weather station (with high economic cost and maintenance) could be replaced by the data provided by a *third-party*. In this regard, it can be concluded that this research is not enough to make a general evaluation of the impact of using *third-party actual* weather data, but it has shown relevant variations in the energy demand and indoor temperature in the four test sites when both weather datasets are used in the simulations, especially for hourly criteria (used in calibration processes and in other applications such as model predictive control). The study also showed that for these types of building energy models, which employ dynamic infiltrations, wind speed's influence on the energy demand is relevant. The significant variations in the results lead us to make the recommendation for researchers to analyze in detail the impact on the building

energy models using *third-party actual* weather data before employing it. For example, this analysis methodology could be performed before making an investment into a weather station by installing a rented provisional one for a period of time. This way, the most influential weather parameters for a specific building energy model and a *third-party* weather provider could be determined, and with this information, an informed choice about which sensors are worth being purchased and installed can be made.

**Author Contributions:** E.L.S. supervised the methodology used in the article, performed the simulations and the analysis, and wrote the manuscript. G.R.R. and C.F.B. developed the methodology and participated in the data analysis. V.G.G. and G.R.R. developed the EnergyPlus models. A.P. provided resources for the study. All the authors revised and verified the manuscript before sending it to the journal. All authors read and agreed to the published version of the manuscript.

**Funding:** This project received funding from the European Union’s Horizon 2020 research and innovation program under Grant Agreement No. 731211, project SABINA.

**Acknowledgments:** We would like to thank the National Technical University of Athens for providing the data of the H2SusBuild and administration building test sites located in Lavrion (Greece) and also Insero in the case of the Gedved school test site in Denmark.

**Conflicts of Interest:** The authors declare no conflicts of interest.

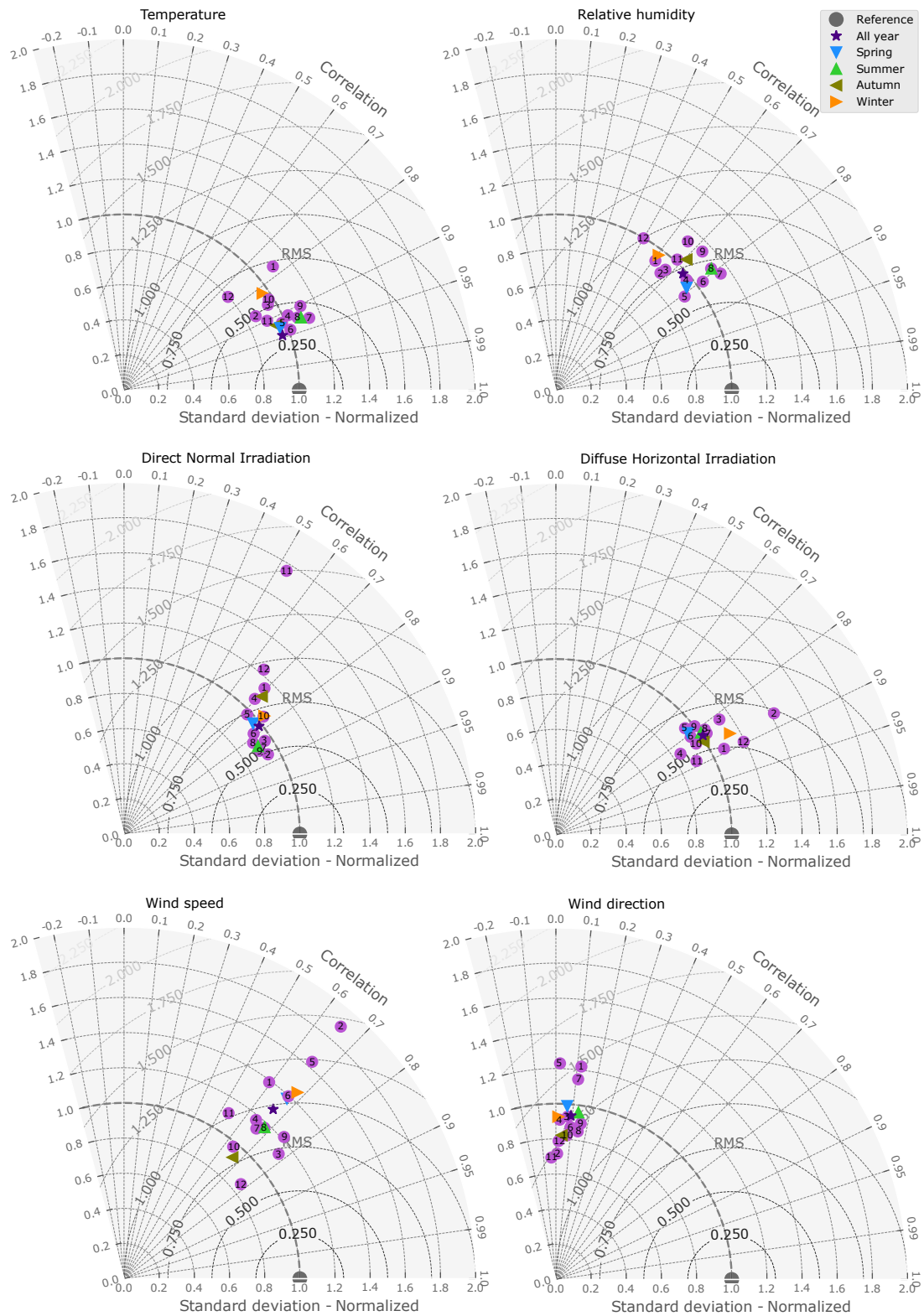
## Abbreviations

The following abbreviations are used in this manuscript:

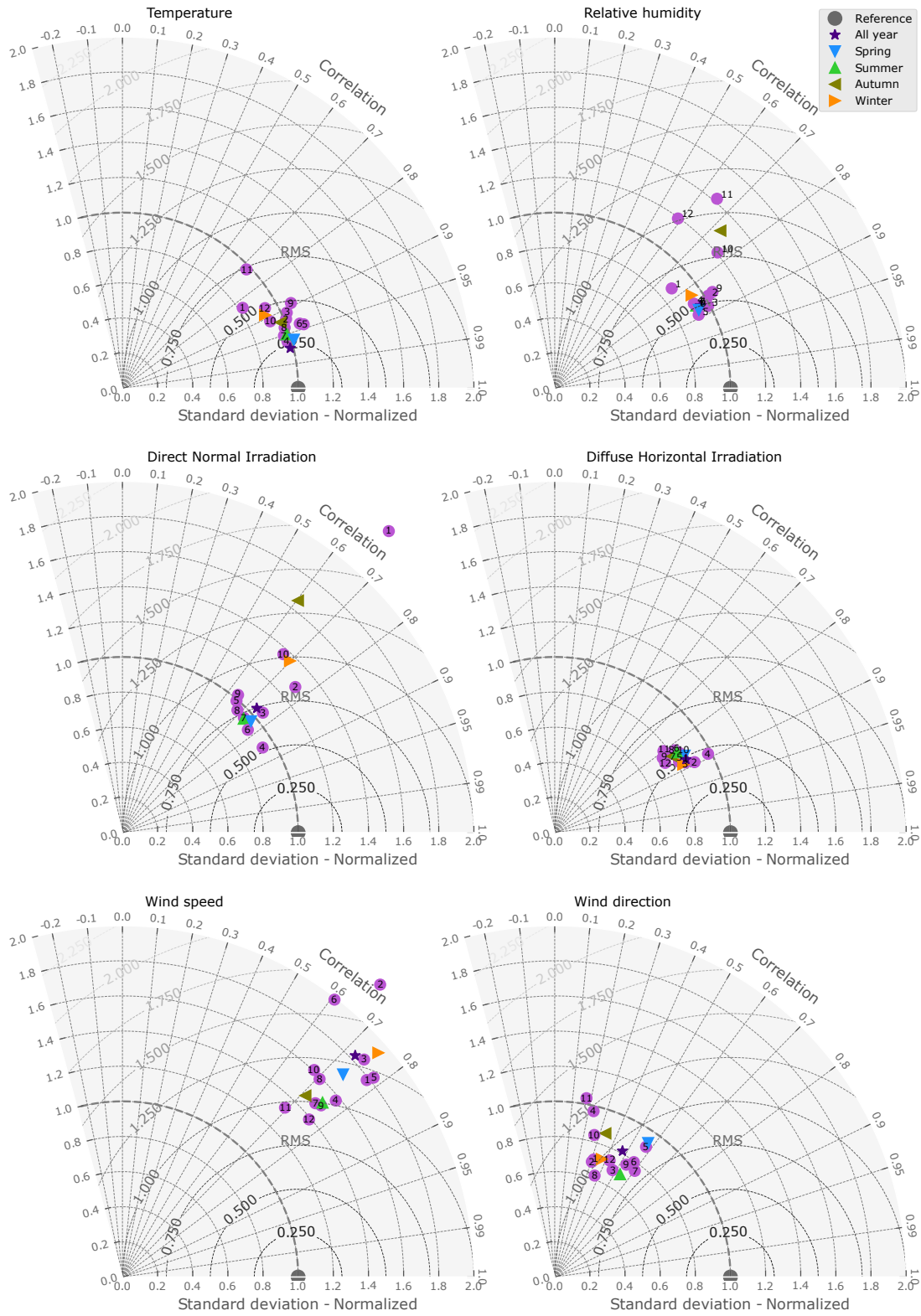
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BEM	Building energy model
CV(RMSE)	Coefficient of variance RMSE
DHI	Diffuse horizontal irradiation
DNI	Direct normal irradiation
EU	European Union
FEMP	Federal Energy Management Program
HVAC	Heating, ventilation, and air-conditioning
IPMVP	International Performance Measurement and Verification Protocol
LTCP	Lavrion Technological and Cultural Park
MAE	Mean absolute error
MADP	Mean absolute deviation percentage
$R^2$	Coefficient of determination
RH	Relative humidity
$RMS_{diff}$	Centered root-mean-squared difference
RMSE	Root-mean-squared error
Temp	Temperature
TPW	<i>Third-party</i> weather data
WD	Wind direction
WS	Wind speed

## Appendix A

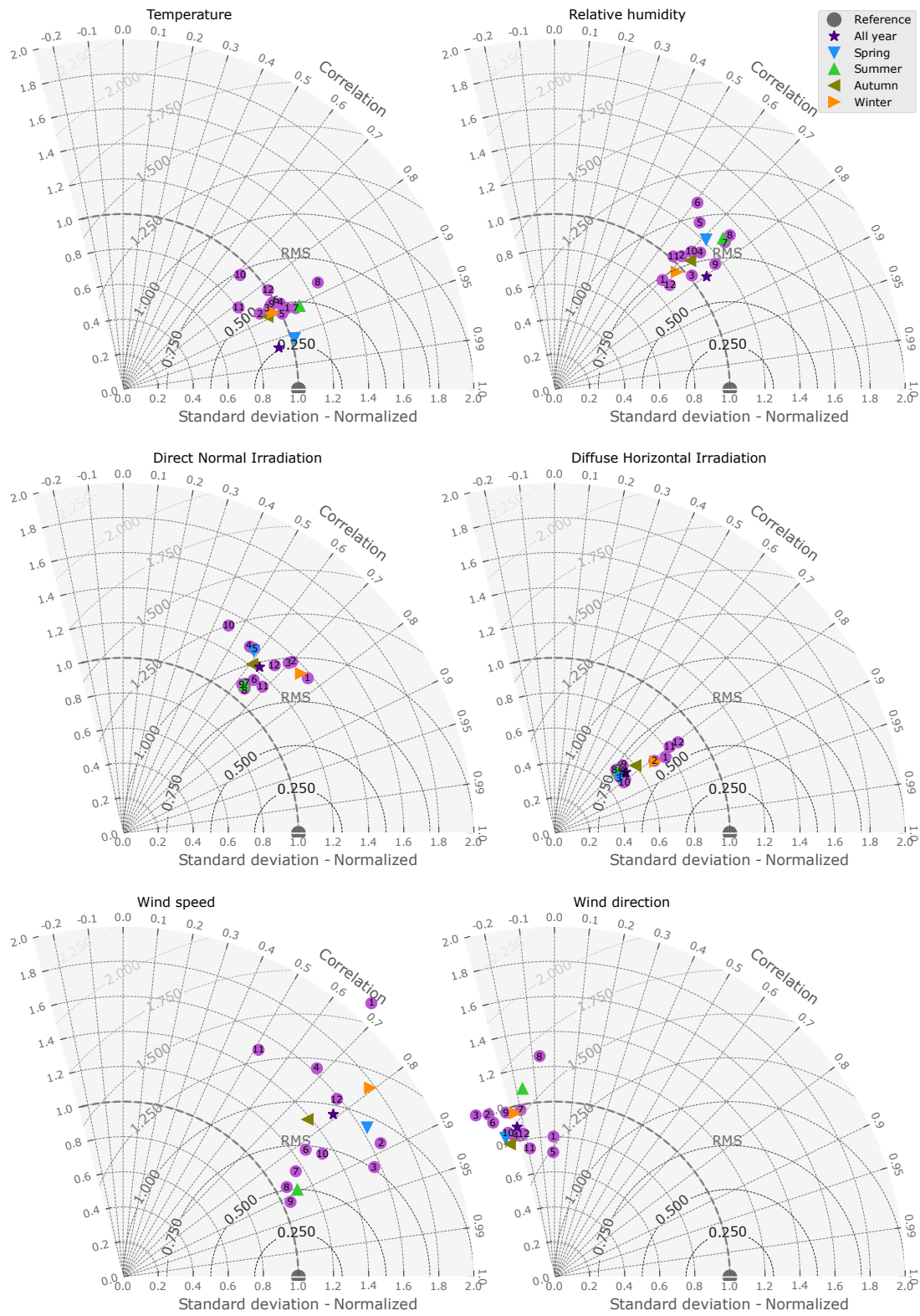
The following figures show detailed Taylor diagrams for each one of the weathers analyzed in the study to complement Figure 3. Figure A1 shows Pamplona’s weather, Figure A2 Gedved’s weather, and Figure A3 Lavrion’s weather. For each location, six diagrams are shown, one for each weather parameter (*Temp*, *RH*, *DNI*, *DHI*, *WS*, and *WD*). The diagrams show the statistical indexes (correlation  $R$ , centered root-mean-squared difference  $RMS_{diff}$ , and standard deviation  $\sigma_f$ ) calculated on an hourly basis for annual, season, and monthly data.



**Figure A1.** Normalized Taylor diagrams of Pamplona (Spain) comparing *on-site* and *third-party* weather data, showing temperature, relative humidity, direct normal irradiation, diffuse horizontal irradiation, wind speed, and wind direction.



**Figure A2.** Normalized Taylor diagrams of Gedved (Denmark) comparing *on-site* and *third-party* weather data, showing temperature, relative humidity, direct normal irradiation, diffuse horizontal irradiation, wind speed, and wind direction.



**Figure A3.** Normalized Taylor diagrams of Lavrion (Greece) comparing *on-site* and *third-party* weather data, showing temperature, relative humidity, direct normal irradiation, diffuse horizontal irradiation, wind speed, and wind direction.



## References

1. United Nations. *The Sustainable Development Goals Report 2019*; United Nations: New York, NY, USA, 2019. Available online: <https://unstats.un.org/sdgs/report/2019/The-Sustainable-Development-Goals-Report-2019.pdf> (accessed on 20 April 2020).
2. UNFCCC Secretariat. *Paris Agreement*; Report of the Conference of the Parties on Its Twenty-First Session, Held in Paris from 30 November to 13 December 2015; United Nations Framework Convention on Climate Change, 2015. Available online: [https://en.wikipedia.org/wiki/Paris\\_Agreement](https://en.wikipedia.org/wiki/Paris_Agreement) (accessed on 20 April 2020).
3. IEA. *Global Status Report for Buildings and Construction, towards a Zero-Emission Efficient and Resilient Buildings and Construction Sector*; IEA: Paris, France, 2019.
4. IBPSA-USA—International Building Performance Simulation Association. Building Energy Software Tools (BEST Directory). Available online: <https://www.buildingenergysoftwaretools.com/software-listing> (accessed on 23 April 2020).
5. Nguyen, A.T.; Reiter, S.; Rigo, P. A review on simulation-based optimization methods applied to building performance analysis. *Appl. Energy* **2014**, *113*, 1043–1058.
6. Bhandari, M.; Shrestha, S.; New, J. Evaluation of weather datasets for building energy simulation. *Energy Build.* **2012**, *49*, 109–118.
7. Marion, W.; Urban, K. *Users manual for TMY2s: Derived from the 1961–1990 National Solar Radiation Data Base*; Technical Report; National Renewable Energy Laboratory: Golden, CO, USA, 1995.
8. Wilcox, S.; Marion, W. *Users Manual for TMY3 Data Sets*; National Renewable Energy Laboratory: Golden, CO, USA, 2008.
9. Thevenard, D.J.; Brunger, A.P. The development of typical weather years for international locations: Part I, algorithms. *Ashrae Trans.* **2002**, *108*, 376–383.
10. Thevenard, D.J.; Brunger, A.P. The development of typical weather years for international locations: Part II, production/Discussion. *Ashrae Trans.* **2002**, *108*, 480.
11. Joe, Y.; Fenxian, H.; Seo, D.; Krarti, M. Development of 3012 IWEC2 Weather Files for International Locations (RP-1477). *Ashrae Trans.* **2014**, *120*, 340–355.
12. Henze, G.P.; Pfafferott, J.; Herkel, S.; Felsmann, C. Impact of adaptive comfort criteria and heat waves on optimal building thermal mass control. *Energy Build.* **2007**, *39*, 221–235.
13. Sandels, C.; Widén, J.; Nordström, L.; Andersson, E. Day-ahead predictions of electricity consumption in a Swedish office building from weather, occupancy, and temporal data. *Energy Build.* **2015**, *108*, 279–290.
14. Ramos Ruiz, G.; Lucas Segarra, E.; Fernández Bandera, C. Model predictive control optimization via genetic algorithm using a detailed building energy model. *Energies* **2019**, *12*, 34.
15. Fernández Bandera, C.; Pachano, J.; Salom, J.; Peppas, A.; Ramos Ruiz, G. Photovoltaic Plant Optimization to Leverage Electric Self Consumption by Harnessing Building Thermal Mass. *Sustainability* **2020**, *12*, 553.
16. Lazos, D.; Sproul, A.B.; Kay, M. Optimisation of energy management in commercial buildings with weather forecasting inputs: A review. *Renew. Sustain. Energy Rev.* **2014**, *39*, 587–603.
17. Crawley, D.B.; Huang, Y.J. Does it matter which weather data you use in energy simulations. *User News* **1997**, *18*, 2–12.
18. Crawley, D.B. Which weather data should you use for energy simulations of commercial buildings? *Trans. Am. Soc. Heat. Refrig. Air Cond. Eng.* **1998**, *104*, 498–515.
19. Erba, S.; Causone, F.; Armani, R. The effect of weather datasets on building energy simulation outputs. *Energy Procedia* **2017**, *134*, 545–554.
20. Voisin, J.; Darnon, M.; Jaouad, A.; Volatier, M.; Aimez, V.; Trovão, J.P. Climate impact analysis on the optimal sizing of a stand-alone hybrid building. *Energy Build.* **2020**, *210*, 109676.
21. González, V.G.; Ruiz, G.R.; Segarra, E.L.; Gordillo, G.C.; Bandera, C.F. Characterization of building foundation in building energy models. In Proceedings of the Building Simulation 2019: 16th Conference of IBPSA, Rome, Italy, 2–4 September 2019.
22. Jentsch, M.F.; James, P.A.; Bourikas, L.; Bahaj, A.S. Transforming existing weather data for worldwide locations to enable energy and building performance simulation under future climates. *Renew. Energy* **2013**, *55*, 514–524.
23. Jentsch, M.F.; Bahaj, A.S.; James, P.A. Climate change future proofing of buildings—Generation and assessment of building simulation weather files. *Energy Build.* **2008**, *40*, 2148–2168.

24. Dickinson, R.; Brannon, B. Generating future weather files for resilience. In Proceedings of the International Conference on Passive and Low Energy Architecture, Los Angeles, CA, USA, 11–13 July 2016; pp. 11–13.
25. Chow, T.T.; Chan, A.L.; Fong, K.; Lin, Z. Some perceptions on typical weather year—From the observations of Hong Kong and Macau. *Sol. Energy* **2006**, *80*, 459–467.
26. Wang, L.; Mathew, P.; Pang, X. Uncertainties in energy consumption introduced by building operations and weather for a medium-size office building. *Energy Build.* **2012**, *53*, 152–158.
27. Song, S.; Haberl, J.S. Analysis of the impact of using synthetic data correlated with measured data on the calibrated as-built simulation of a commercial building. *Energy Build.* **2013**, *67*, 97–107.
28. Silvero, F.; Lops, C.; Montelpare, S.; Rodrigues, F. Generation and assessment of local climatic data from numerical meteorological codes for calibration of building energy models. *Energy Build.* **2019**, *188*, 25–45.
29. Ciobanu, D.; Eftimie, E.; Jaliu, C. The influence of measured/simulated weather data on evaluating the energy need in buildings. *Energy Procedia* **2014**, *48*, 796–805.
30. Cuerda, E.; Guerra-Santin, O.; Sendra, J.J.; Neila, F.J. Understanding the performance gap in energy retrofiting: Measured input data for adjusting building simulation models. *Energy Build.* **2020**, *209*, 109688.
31. Du, H.; Barclay, M.; Jones, P.J. Generating high resolution near-future weather forecasts for urban scale building performance modelling. In Proceedings of the 15th IBPSA Conference, San Francisco, CA, USA, 7–9, August 2017.
32. Du, H.; Jones, P.; Segarra, E.L.; Bandera, C.F. Development of a REST API for obtaining site-specific historical and near-future weather data in EPW format. Presented at Building Simulation and Optimization, Cambridge, UK, 11–12 September 2018.
33. Du, H.; Bandera, C.F.; Chen, L. Nowcasting methods for optimising building performance. In Proceedings of the 16th IBPSA Conference, Rome, Italy, 2–4 September 2019.
34. Henze, G.P.; Kalz, D.E.; Felsmann, C.; Knabe, G. Impact of forecasting accuracy on predictive optimal control of active and passive building thermal storage inventory. *HVAC&R Res.* **2004**, *10*, 153–178.
35. Agüera-Pérez, A.; Palomares-Salas, J.C.; de la Rosa, J.J.G.; Florencias-Oliveros, O. Weather forecasts for microgrid energy management: Review, discussion and recommendations. *Appl. Energy* **2018**, *228*, 265–278.
36. Yan, X.; Abbas, D.; Francois, B. Uncertainty analysis for day ahead power reserve quantification in an urban microgrid including PV generators. *Renew. Energy* **2017**, *106*, 288–297.
37. Powell, K.M.; Sriprasad, A.; Cole, W.J.; Edgar, T.F. Heating, cooling, and electrical load forecasting for a large-scale district energy system. *Energy* **2014**, *74*, 877–885.
38. Petersen, S.; Bundgaard, K.W. The effect of weather forecast uncertainty on a predictive control concept for building systems operation. *Appl. Energy* **2014**, *116*, 311–321.
39. Zhao, J.; Duan, Y.; Liu, X. Uncertainty analysis of weather forecast data for cooling load forecasting based on the Monte Carlo method. *Energies* **2018**, *11*, 1900.
40. Haben, S.; Giasemidis, G.; Ziel, F.; Arora, S. Short term load forecasting and the effect of temperature at the low voltage level. *Int. J. Forecast.* **2019**, *35*, 1469–1484.
41. Seo, D.; Huang, Y.J.; Krarti, M. Impact of typical weather year selection approaches on energy analysis of buildings. *ASHRAE Trans.* **2010**, *116*, 416–427.
42. Radhi, H. A comparison of the accuracy of building energy analysis in Bahrain using data from different weather periods. *Renew. Energy* **2009**, *34*, 869–875.
43. SABINA SmArt BI-Directional Multi eNergy gAteway. Available online: <https://sabina-project.eu/> (accessed on 20 April 2020).
44. Guideline, A. *Guideline 14-2002, Measurement of Energy and Demand Savings*; American Society of Heating, Ventilating, and Air Conditioning Engineers: Atlanta, GA, USA, 2002.
45. Taylor, K.E. Taylor Diagram Primer. 2005. Available online: [http://www.pcmdi.llnl.gov/about/staff/Taylor/CV/Taylor\\_diagram\\_primer.pdf](http://www.pcmdi.llnl.gov/about/staff/Taylor/CV/Taylor_diagram_primer.pdf) (accessed on 20 April 2020).
46. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **2001**, *106*, 7183–7192.
47. Dong, T.Y.; Dong, W.J.; Guo, Y.; Chou, J.M.; Yang, S.L.; Tian, D.; Yan, D.D. Future temperature changes over the critical Belt and Road region based on CMIP5 models. *Adv. Clim. Chang. Res.* **2018**, *9*, 57–65.
48. Chen, G.; Zhang, X.; Chen, P.; Yu, H.; Wan, R. Performance of tropical cyclone forecast in Western North Pacific in 2016. *Trop. Cyclone Res. Rev.* **2017**, *6*, 13–25.

49. Yan, G.; Wen-Jie, D.; Fu-Min, R.; Zong-Ci, Z.; Jian-Bin, H. Surface air temperature simulations over China with CMIP5 and CMIP3. *Adv. Clim. Chang. Res.* **2013**, *4*, 145–152.
50. Nabeel, A.; Athar, H. Stochastic projection of precipitation and wet and dry spells over Pakistan using IPCC AR5 based AOGCMs. *Atmos. Res.* **2020**, *234*, 104742.
51. de Assis Tavares, L.F.; Shadman, M.; de Freitas Assad, L.P.; Silva, C.; Landau, L.; Estefen, S.F. Assessment of the offshore wind technical potential for the Brazilian Southeast and South regions. *Energy* **2020**, *196*, 117097.
52. Crawley, D.B.; Lawrie, L.K.; Winkelmann, F.C.; Buhl, W.F.; Huang, Y.J.; Pedersen, C.O.; Strand, R.K.; Liesen, R.J.; Fisher, D.E.; Witte, M.J.; et al. EnergyPlus: Creating a new-generation building energy simulation program. *Energy Build.* **2001**, *33*, 319–331.
53. Crawley, D.B.; Lawrie, L.K.; Pedersen, C.O.; Winkelmann, F.C.; Witte, M.J.; Strand, R.K.; Liesen, R.J.; Buhl, W.F.; Huang, Y.J.; Henninger, R.H.; et al. EnergyPlus: An update. In Proceedings of the SimBuild, Boulder, CO, USA, 4–6 August 2004; Volume 1.
54. EnergyPlus. *Auxiliary Programs: EnergyPlus™ version 8.9.0 Documentation*; U.S. Department of Energy: Washington, DC, USA, 2018.
55. Kolassa, S.; Schütz, W. Advantages of the MAD/MEAN ratio over the MAPE. *Foresight Int. J. Appl. Forecast.* **2007**, *6*, 40–43.
56. Petojević, Z.; Gospavić, R.; Todorović, G. Estimation of thermal impulse response of a multi-layer building wall through in-situ experimental measurements in a dynamic regime with applications. *Appl. Energy* **2018**, *228*, 468–486.
57. Lucas Segarra, E.; Du, H.; Ramos Ruiz, G.; Fernández Bandera, C. Methodology for the quantification of the impact of weather forecasts in predictive simulation models. *Energies* **2019**, *12*, 1309.
58. U.S. Department of Energy. *M&V Guidelines: Measurement and Verification for Federal Energy Projects Version 3.0*; U.S. Department of Energy: Washington, DC, USA 2008.
59. IPMVP Committee. *International Performance Measurement and Verification Protocol: Concepts and Options for Determining Energy and Water Savings, Volume I*; Technical Report; National Renewable Energy Laboratory: Golden, CO, USA, 2001.
60. González, V.G.; Colmenares, L.Á.; Fidalgo, J.F.L.; Ruiz, G.R.; Bandera, C.F. Uncertainty's Indices Assessment for Calibrated Energy Models. *Energies* **2019**, *12*, 2096.
61. Zhao, J.; Liu, X. A hybrid method of dynamic cooling and heating load forecasting for office buildings based on artificial intelligence and regression analysis. *Energy Build.* **2018**, *174*, 293–308.
62. Perez, R.; Cebecauer, T.; Šúri, M. Semi-empirical satellite models. In *Solar Energy Forecasting and Resource Assessment*; Academic Press: Boston, MA, USA, 2013; pp. 21–48.
63. Kato, T. Prediction of photovoltaic power generation output and network operation. In *Integration of Distributed Energy Resources in Power Systems*; Elsevier: Amsterdam, The Netherlands, 2016; pp. 77–108.
64. Willmott, C.J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82.
65. Meteoblue. Available online: <https://meteoblue.com/> (accessed on 20 April 2020).
66. Meteoblue Weather Simulation Data. Available online: <https://content.meteoblue.com/en/specifications/data-sources/weather-simulation-data/> (accessed on 6 July 2020).
67. Guglielmetti, R.; Macumber, D.; Long, N. *OpenStudio: An Open Source Integrated Analysis Platform*; Technical Report; National Renewable Energy Laboratory (NREL): Golden, CO, USA, 2011.
68. EnergyPlus. *EnergyPlus Input Output Reference*; US Department of Energy: Washington, DC, USA, 2018.
69. Ruiz, G.R.; Bandera, C.F.; Temes, T.G.A.; Gutierrez, A.S.O. Genetic algorithm for building envelope calibration. *Appl. Energy* **2016**, *168*, 691–705.
70. Ruiz, G.R.; Bandera, C.F. Analysis of uncertainty indices used for building envelope calibration. *Appl. Energy* **2017**, *185*, 82–94.
71. Fernández Bandera, C.; Ramos Ruiz, G. Towards a new generation of building envelope calibration. *Energies* **2017**, *10*, 2102.

