Use of a numerical weather prediction model as a meteorological source for building thermal simulations

Javier López Gómez[a,*], Francisco Troncoso Pastoriza[a], Elena Arce Fariña[b], Pablo Eguía Oller[a], Enrique Granada Álvarez[a]

[a]GTE Research Group, School of Industrial Engineering, University of Vigo, Campus Lagoas-Marcosende, 36310, Vigo, Pontevedra, Spain

[b]Defense University Centre, Spanish Naval Academy, Plaza de España, s/n, 36920, Marín, Spain

*Corresponding author. e-mail address: javilopez@uvigo.es (J. López Gómez).

This version of the article has been accepted for publication, after peer review, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1016/j.scs.2020.102403

Abstract

Thermal simulations are a commonly used tool for energy efficiency analysis of buildings. Regional meteorological station networks are a prime source of weather data inputs, required for building thermal simulations. However, local measurements from weather stations are not always available, and when available, access to data may be expensive. This paper analysed a novel use of a numerical weather prediction mesoscale model, the Global Forecast System (GFS) sflux model, as a source of input data for transient thermal simulations. Two interpolation techniques (nearest neighbour and universal kriging) were used to generate local weather datasets from GFS outputs at 27 locations spread over an area of 29,574 km^2 in Galicia (northwest Spain). The performance of the GFS estimations was tested against weather measurements obtained from a government weather agency. A representative building with the most common features was selected for running thermal simulations in the TRNSYS environment, focused on heating demands, with estimated weather data as the input. The results highlighted that GFS-interpolated datasets consistently performed better than using measured data from the nearest weather station. GFS was found to be an appropriate weather source for building simulations and was able to provide good-quality, free and global-scale local weather inputs.

Keywords

Building simulation; Weather data; Numerical weather prediction; Global Forecast System (GFS); Kriging; Interpolation.

How to cite

López Gómez, J., Troncoso Pastoriza, F., Fariña, E. A., Eguía Oller, P., & Granada Álvarez, E. (2020). Use of a numerical weather prediction model as a meteorological source for the estimation of heating demand in building thermal simulations. *Sustainable Cities and Society*, *62*, 102403. https://doi.org/10.1016/j.scs.2020.102403

Nomenclature

AEMET	Spanish Meteorological Agency (*Agencia Estatal de Meteorología*)	MBE	Mean Bias Error
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers	MG	Galician Meteorological Agency (MeteoGalicia)
CDD	Cooling Degree Days	NCEP	National Centres for Environmental Prediction
CTE	Spanish Technical Building Code (*Código Técnico de Edificación*)	NN	Nearest Neighbour
CV(RMSE)	Root Mean Square Error Coefficient of Variation	NN GFS	Nearest Neighbour interpolation over GFS sflux outputs dataset
ECMWF	European Centre for Medium-Range Weather Forecasts	NN MG	Nearest Neighbour interpolation over MeteoGalicia measurements dataset
EPW	EnergyPlus Weather format file	NOAA	National Oceanic and Atmospheric Administration
FTP	File Transfer Protocol	NOMADS	National Oceanic and Atmospheric Administration Operational Model Archive and Distribution System
GFS	Global Forecast System	NWP	Numerical Weather Prediction
HDD	Heating Degree Days	Ref.	MeteoGalicia Reference dataset
HTTP	Hypertext Transfer Protocol		
HVAC	Heating, Ventilating and Air Conditioning	RMSE	Root Mean Square Error
IDAE	Spanish Institute for Energy Diversification and Saving (*Instituto para la Diversificación y Ahorro de Energía*)	TRNSYS	Transient System Simulation Tool
IFS	Integrated Forecast System	UK	Universal Kriging
IPMA	Portuguese Meteorological Agency (*Instituto Português do Mar e da Atmosfera*)	UK GFS	Universal Kriging interpolation over GFS sflux outputs dataset
MAE	Mean Absolute Error	UK MG	Universal Kriging interpolation over MeteoGalicia measurements dataset

1. Introduction

Thermal energy simulations of building models can be used to estimate energy demands when real measurements are not available [1]. Thermal simulations are widely used to evaluate the energy requirements of buildings in response to both actual and predicted environmental conditions [2]. For the last two decades, the building sector has made up between 20 and 40% of the final energy intake in industrial countries, with the shares of the EU and the USA being approximately 40% [3]. Of that percentage, maintaining the thermal comfort of occupants using heating, ventilating and air conditioning (HVAC) demands account for the largest share [4]. Thus, estimating the thermal energy demands beforehand may provide evidence to support decisions about construction and/or maintenance of existing and new buildings [5]. The estimates can also help to identify and predict the energetic impacts of installed efficiency-enhancing measures and retrofit decisions [6]. Guaranteeing the reliability of the simulation results is key to making decisions.

Weather data are environmental variables that often affect energy use or demand. External loads such as air temperature, solar radiation and wind flow are directly related to the energy consumption of buildings [7]. Previous studies have shown that a lack of accurate weather data has a large impact on energy demand estimations (e.g., the performance of thermal insulating elements is dependent on the surrounding temperature and humidity [8]). The estimated annual energy consumption of a building varies up to 7% depending on the weather dataset that is used

[9], [10]. Moreover, fluctuations in the surrounding weather conditions may result in an up to 5% variation in the actual energy consumption of a residential building [11], which is particularly significant for heating and cooling loads, for which the variation may be as large as 40% [5].

The two main causes that lead to the lack of precision in meteorological data are data gaps and an absence of on-site data. One way to solve these problems is to interpolate weather data from historical regional sources, either directly using geospatial methods, such as ordinary or universal kriging (UK) [9] or thin plate splines [10] interpolations, or using other available variables [12]. However, having access to regional weather data is not always possible.

In the Iberian Peninsula context, two national-level governmental weather agencies cover the territory: The Agencia Estatal de Meteorología (AEMET) in Spain and the Instituto Português do Mar e da Atmosfera (IPMA) in Portugal. Each of these agencies develops and maintains a network of automatic weather measurement stations for its respective country, alongside other attributions. In addition, three Spanish regions have their own regional weather agencies, with independent station networks that overlap the AEMET networks for such regions. Table 1 compiles information about all these agencies, obtained from their respective official websites.

Table 1: Governmental weather agencies for the Iberian Peninsula

Name	Short name	Operation environment	Operative stations	Station density [uds/km^2]	Free historical weather data
Agencia Estatal de Meteorología	AEMET	All Spain	807	0.0016	No*
MeteoGalicia	MeteoGalicia	Galician region	166	0.0056	Yes
Euskal Meteorologia Agentzia	Euskalmet	Basque region	50	0.0069	Yes
Servei Meteorològic de Catalunya	meteo.cat	Catalonian region	190	0.0059	Yes**
Instituto Português do Mar e da Atmosfera	IPMA	All Portugal	135	0.0015	No

*Free access available only for the latest 24 hours

**Free access only for personal use, students, research centres and public administration entities

Table 1 highlights that having access to historical weather data can be problematic for short time periods (more than a few hours) in some regions. When adequate observational data are not available, predictive outputs from numerical weather models may be used [13].

The National Centres for Environmental Prediction (NCEP), which are part of the United States National Oceanic and Atmospheric Administration (NOAA), constitute a governmental agency that carries out meteorological and climatic studies at both the national and global scale. The NCEP Global Forecast System (GFS) is one of the most extended global numerical weather prediction (NWP) models for mesoscale forecasting [14], [15], [16]. The closest competitor to the NCEP GFS is the Integrated Forecast System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF). Both provide global-scale predictions for several days over a discrete grid. However, unlike the IFS, the GFS outputs are freely licensed and available in the public domain.

The objective of this paper is to evaluate the use of one of the NCEP GFS models as a weather source for thermal heating demand simulations of buildings. This is a novel application of the GFS model that could potentially solve the previously introduced issue with the availability of regional weather data. As a free, globally available source of meteorological outputs, the potential use of the GFS outputs instead of a physical network of weather stations, where data from such networks may be scarce or not available, is worthy of in-depth analysis.

This study was carried out over the territory of the Spanish Galician autonomous region during 2018. The weather station network of the Galician regional meteorological agency was used to obtain meteorological data for six weather variables from 27 selected locations. These data were used as inputs for thermal heating simulations of a TRNSYS model of a representative building. These simulation results and weather inputs comprised the reference dataset, which was assumed to represent the true weather and heating demand conditions for the studied locations. Universal kriging and nearest neighbour interpolation methods were used to obtain localized data of the selected weather variables. The predicted heating demands obtained with TRNSYS were also generated. Weather interpolation and heating demand simulation error metrics were computed against the reference dataset previously obtained.

Previous studies performed by some of the authors have demonstrated the effectiveness of using meteorological data from government agencies to run building thermal simulations [17], [9], [10]. The interpolation methods used here have already been successfully applied in previous research work: the nearest neighbour algorithm was tested in [17], while the universal kriging algorithm was evaluated in [9], [10] and [13]. The simulated building model was also tested in [9]. The main novel aspect of the present study is the analysis of weather outputs of one GFS model used as meteorological inputs for thermal building simulations. A well-tested reference weather data, verified interpolation methods and thermal simulation models are used to evaluate the performance of this novel data source.

2. Methodology

Figure 1 shows a diagram that summarizes the methodology used and described in this section. The research started by defining the target area and selecting both the interpolation techniques and the meteorological data sources. Then, a building thermal demand model was defined. Finally, error metrics were used to evaluate the meteorological interpolation techniques and energy demand simulation results.

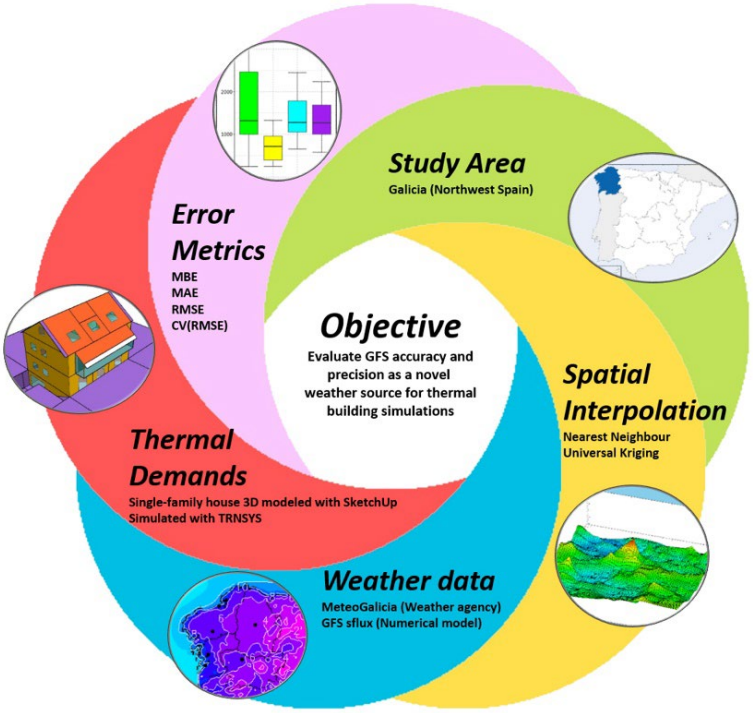

Figure 1: Methodology diagram

2.1. Study area

This research focused on the Galician autonomous region, located in north-western Spain. Galicia has a fractured orography of eroded mountains and an irregular coast. The region has a rainy, maritime climate that is strongly affected by the Atlantic Ocean, but it evolves into a more continental/mediterranean climate in the inland areas. According to the Köppen-Geiger classification [18], the Csb warm-summer mediterranean climate is predominant. The Cfb marine climate is present in some locations, especially over the northernmost area. Figure 2 shows a 5 km raster resolution of the Köppen-Geiger classification for the entire Iberian Peninsula. A red line denotes the Galician study area boundary. This map was generated using the R code provided in [19].

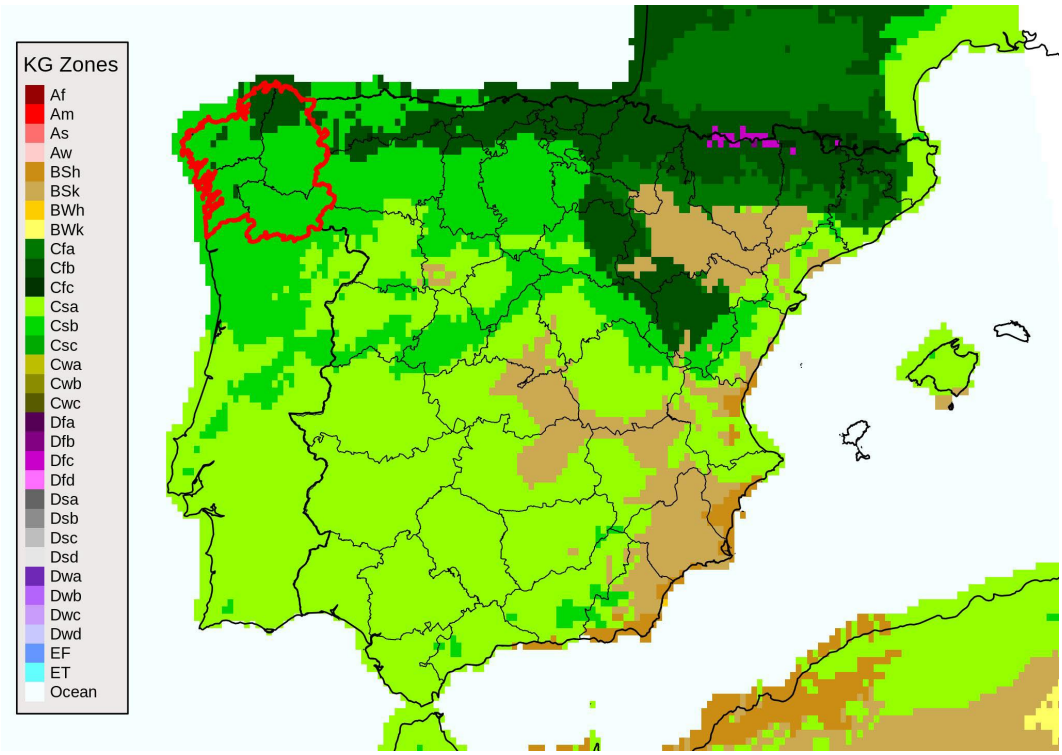

Figure 2: Köppen-Geiger climate classification over Iberian Peninsula

2.2. Interpolation techniques

Based on previous studies [17], [9], [10], two interpolation techniques were chosen for both weather data sources: nearest neighbour and universal kriging.

The NN technique assigns to the target location the closest source of information, selected from a pool of provided observation points. The data values to be estimated at the target location are assumed to be equal to the local measurements available at this selected nearest observation point. This method provides fast estimations with a negligible computation cost. The main drawback of this technique is its low accuracy and precision when the closest observation point is distant or when the estimated fields are irregular with local discontinuities.

Kriging interpolation techniques focus on spatial relationships between the values of a pool of observation points. Kriging interpolations compute an experimental variogram that relates distances between pairs of observation points and the relations for data values observed at said pairs of points. The technique fits afterwards a mathematical model to this experimental

variogram. Once the best model is selected, it is used to choose appropriate weighting parameters for interpolation at any new estimation point based on data from the weighted observation points. Kriging interpolations provide the best linear unbiased estimations (usually referred to as BLUE) from all possible interpolation techniques [20]. A review of 51 comparative studies on spatial interpolation techniques available at [21] notes that kriging methods are generally the best choice, with a few exceptions.

Universal kriging is a member of the kriging family that tries to estimate a global trend (or drift) in observation data, as opposed to ordinary kriging, in which the global trend is assumed to be null. Some authors indistinctly refer to this method as universal kriging, kriging with a trend, regression kriging or kriging with external drift. Other authors, however, prefer to reserve the two latter names for specifically referring to some slightly different versions of UK. In such versions, the trend is predicted using secondary variables, leaving the UK term for the kriging version where such a trend is estimated using only the spatial coordinates [21]. In the present study, UK refers to a kriging technique with an unknown (estimated) trend predicted using more than only spatial coordinates. Specifically, the four variables (estimators) that were used for the kriging trend prediction were geographical latitude and longitude, local ground elevation (measured from mean sea level), and distance to the nearest coastline point, following the conclusions of [17].

All interpolations were conducted using the R environment with the following external packages: FNN for nearest neighbour [22], and automap [23] and gstat [24] for universal kriging. For error analysis and plotting, Python scripts were used.

2.3. Weather data

2.3.1. Weather variables

Based on the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) guideline 14 [25], the weather inputs used for the building energy modelling were composed of six elements: dry bulb air temperature, air relative humidity, absolute pressure, global horizontal solar radiation (which may be sometimes referred to as "solar radiation" henceforth), wind speed and wind direction. These six main weather variables composed the meteorological input data passed to the thermal simulations.

In addition, three more auxiliary weather variables were used in this study: sea level pressure, northward wind component speed and eastward wind component speed. The sea level pressure variable was used for error evaluation instead of absolute pressure, so comparisons between locations with large altitude differences would not be biased by the effects of these large height differences. Sea level pressure was calculated using equation (1):

$$P_{SL} = P\left[\frac{T}{T+L\Delta z}\right]^{\frac{g_0 M}{RL}} \quad (1)$$

where P and T are the pressure and the temperature at local ground level height, respectively; Δz is the local altitude from sea level; $g_0 = 9.8067\ m/s^2$ is the gravitational acceleration; $M = 0.02894\ kg/kmol$ is the air molar mass; $R = 8.3145\ J/(mol \cdot K)$ is the universal gas constant; and $L = -6.5\ ℃/km$ is the standard temperature lapse rate taken from the Standard Atmosphere [26]. The equation itself was derived from the thermodynamic ideal gas law.

The wind speed and direction variables obtained from the different data sources were transformed into northward (blowing from the south towards the north) and eastward (blowing from the west toward the east) wind components, so the wind conditions could be interpolated using two scalar magnitudes rather than a vector. Wind scalar components were also used when computing estimation error metrics, which helped avoid false error results caused by the angular and cyclic nature of wind direction (e.g. a 359° estimation, when compared to a 1° observation, would return

an incorrect difference of 358° instead of 2°). Wind components were calculated using equations (2) and (3):

$$NW = -\cos(\pi WD/180) \cdot WS \quad (2)$$

$$EW = -\sin(\pi WD/180) \cdot WS \quad (3)$$

where WD is the wind angular direction (measured on sexagesimal degrees) and WS is the wind speed magnitude. The equations were taken from [27].

2.3.2. Data sources

The weather data files used for this research were obtained from two sources: The MeteoGalicia historical database, available at the agency webpage, and outputs from one of the NCEP GFS models, available on one of the National Oceanic and Atmospheric Administration Operational Model Archive and Distribution System (NOMADS) real time data servers.

MeteoGalicia has regularly been used by some of the authors in previous studies [17], [9], [10]. This source includes a network of 167 meteorological stations scattered over the Galician territory. Weather stations have a higher density in coastal areas, and are scarcer in inland zones. All weather stations record and store data with a 10-minute temporal resolution [28]. Near-ground variables are measured at a height of 1.5 or 2 m, while wind-related variables are most commonly measured at a height of 10 m. Current conditions and latest measurements can be consulted on the MeteoGalicia webpage [29]. This website also provides free access to all historical recorded datasets and supports automatic, periodic data requests.

The GFS model chosen in this paper was the GFS sflux, which is solved on a semi-Lagrangian grid with 13 km grid point separation in the horizontal plane and 64 layers for vertical resolution, from ground level to 0.2 hPa (approximately 55 km above sea level). The technical identifier for this resolution grid is *T1534*. The temporal resolution is one hour, with predictions generated for 384 hours (16 days), though only three-hour resolution outputs are available after the first 120 forecast steps. The GFS sflux model is solved every six hours, with four executions run every day starting from 00:00:00 UTC. All currently operational GFS model outputs are stored on the NOMADS NCEP HTTP and FTP servers [30]. GFS sflux output results are stored as hourly grib2 files, containing 112 variables (for the 0-5 first hours) or 114 (for the 6-384 hours) for all the planetary grid points. Each grib2 file is publicly available for 10 days.

Air temperature and relative humidity measurements from MeteoGalicia stations were available at 2 m above the local ground surface. The GFS output files contained estimation values for both temperature and humidity at the same elevation from ground level, so they could be directly compared. The same was applicable to the wind variables, available at a local height of 10 m for both MeteoGalicia and GFS. Absolute pressure and global solar radiation GFS outputs were available at the local ground surface level, while measurements from MeteoGalicia were taken by sensors at 1-2 m above local ground. However, these measurement height differences were sufficiently small to cause a negligible effect on the results. All GFS variables were available at a 1-hour temporal resolution, so they were used without further processing. MeteoGalicia observations were available with a 10-minute resolution, so they were averaged to 1-hour values.

2.4. Thermal simulations

Figure 3 shows a representative building with the most common construction features in Galicia, which was used to run the thermal simulations. The building was an isolated single-family house, so no thermal contributions or shadows from other structures were included in the simulation environment. The building comprised three interconnected floors above the ground and one underground garage, with a net floor area of 292.4 m^2. The main façade had a southwest orientation. Trimble SketchUp software was used to define the building geometry.

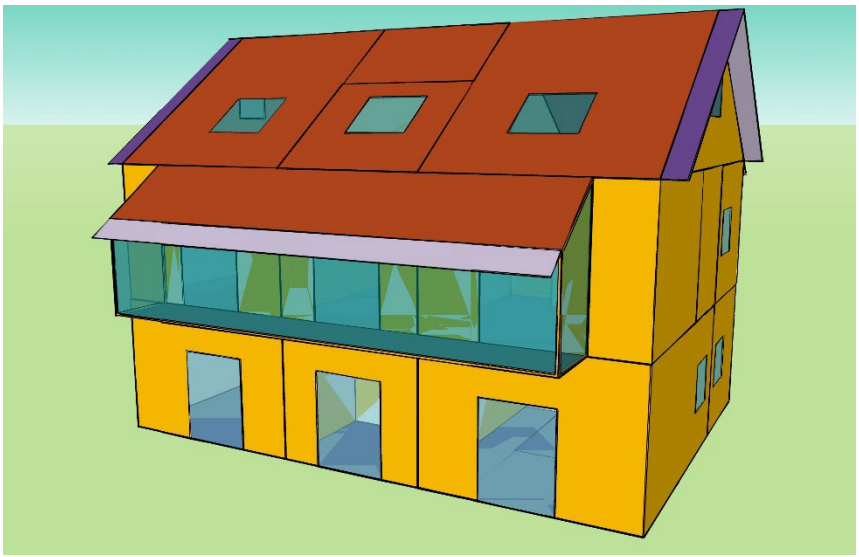

Figure 3: South-western facade of the target building model on SketchUp

Table 2 summarizes the enclosure and window data.

Table 2: Enclosure and window data

Category	Type	d [mm]	U-value [W/m^2 K]
External wall	Roof	320	0.709
	Outer facade	340	0.489
Adjacent wall	Forged floor	391	0.756
	Interior wall	100	2.609
	RadiantFloor_10	423	0.722
	RadiantFloor_30	483	0.699
Boundary wall	Basement wall	200	3.892
	Ground	400	2.314
Window	Glazing	4/16/4	1.060
	Frame	-	3.030

The TRNSYS simulation environment was used to run thermal simulations. The model was built following the main applicable regulations in Spain: The Spanish Technical Building Code (CTE) [31] and the Institute for Energy Diversification and Saving (IDAE) specifications [32]. The CTE instructions state that calculations of cooling resources are not mandatory for the Galician zone. Hence, refrigeration demands were not considered. A hypothesis of unlimited available thermal resources was assumed, which means that the results obtained from the TRNSYS simulations accounted for ideal demands and were not restricted by installed thermal power. The influences of atmospheric pressure and wind on the building thermal demands (along with air temperature and humidity) were included by modelling heat gains and losses due to infiltrations. A scheduled occupancy profile was imposed, with a 20 °C temperature control set for the afternoon/early night part of day (starting at 15:00 and stopping at 23:00 UTC+1). The heating seasons were established from January 1 to March 31 and from October 20 to December 31. Figure 4 shows the TRNSYS model. More information about the thermal model of the building can be found in [9].

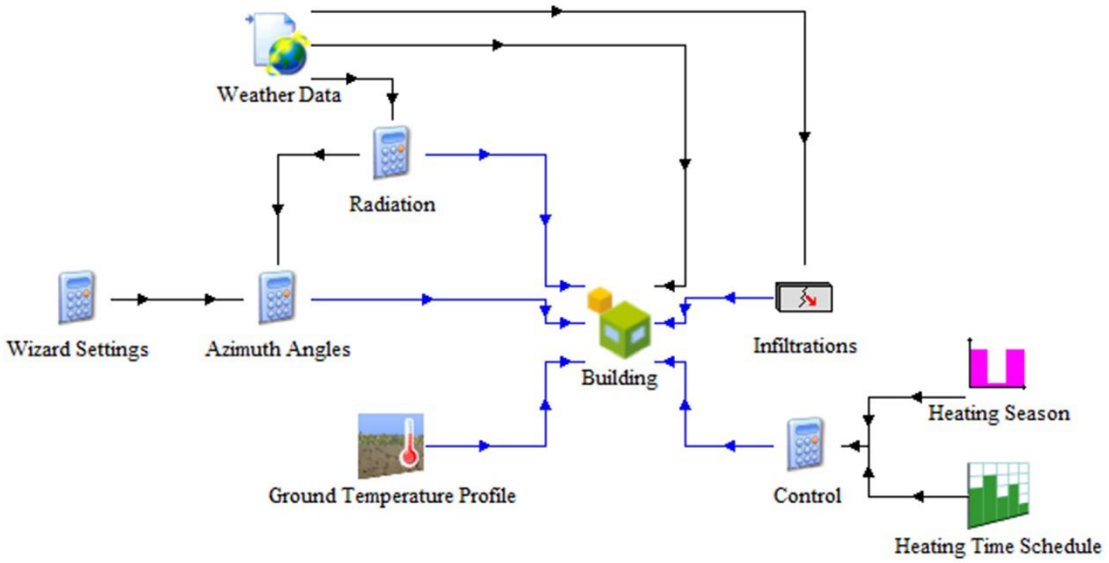

Figure 4: TRNSYS model

Raw weather data had to be adapted and filled before being used in the TRNSYS thermal simulations. This software requires the weather input values to be complete, with missing hours on the simulation temporal span not being allowed. Data gaps (caused by unavailable values in the local databases) were filled using universal kriging interpolations of MeteoGalicia weather data. Those filled hours were not computed when evaluating the results. Weather Converter, an auxiliary program provided by the EnergyPlus simulation suite, was used to generate weather format (EPW) files.

2.5. Operational conditions

2.5.1. Temporal interval

The present study was carried out using weather data covering all of 2018, which was due to the GFS sflux local database starting to operate during late October 2017, so 2018 was the first complete year available.

2.5.2. Locations

Multiple locations across the Galician study area were selected to perform weather interpolation and building simulation estimations. The use of a wide pool of locations allowed to consider the geographical influence on building energy demands, as stated in [33], [34].

The set of study locations comprised 40 points on an initial stage (the 10 most populated municipalities for each of the 4 Galician provinces). This first selection was then narrowed by assigning a MeteoGalicia station to each municipality. Locations whose corresponding station was more than 10 km away or had less than 75% valid data for the study time interval, were discarded. Locations whose corresponding station was already assigned to a larger location were also discarded. Thus, 27 locations were finally selected as weather interpolation and heating simulation scenarios.

Table 3 presents a summary of the studied locations, including geographical estimators for UK interpolations, Köppen-Geiger climate classification, and heating and cooling degree days (HDD and CDD). These three last variables were not used as inputs in the study. However, these variables were included in the table to provide more information about the locations where the study was performed.

Table 3: Locations description

Study location	Latitude	Longitude	Altitude [m]	Coast distance [km]	Köppen climate	HDD_{18} [°C·days/year]	CDD_{18} [°C·days/year]
A Coruña	43.36	-8.44	131	1.2	Csb	1644	340
A Veiga	42.36	-7.01	1229	125.1	Csb	3307	96
Avión	42.41	-8.24	553	31.3	Csb	2220	246
Burela	43.64	-7.37	421	3.2	Cfb	2188	25
Celanova	42.17	-7.97	623	54.1	Csb	2177	264
Chantada	42.61	-7.72	391	78.9	Csb	2083	232
Ferrol	43.49	-8.25	37	0.5	Cfb	1497	224
Foz	43.56	-7.28	73	1.3	Cfb	1781	85
Lalín	42.61	-8.14	500	48.1	Csb	2367	119
Lugo	42.99	-7.54	419	62.5	Csb	2209	163
Monforte de Lemos	42.47	-7.50	645	92.1	Csb	2300	187
Ourense	42.35	-7.88	139	60.1	Csb	1561	579
Ponteareas	42.18	-8.52	41	13.7	Csb	1661	362
Pontevedra	42.41	-8.66	57	0.4	Csb	1406	353
Redondela	42.32	-8.60	260	0.6	Csb	1803	327
Ribadeo	43.54	-7.08	51	1.7	Csb	1527	116
Ribeira	42.56	-9.03	30	0.1	Csb	1387	266
Santiago de Compostela	42.88	-8.56	255	26.5	Csb	1841	229
Santiago de Compostela	42.89	-8.52	305	29.8	Csb	1929	182
Sarria	42.80	-7.38	416	80.7	Csb	2254	182
Vedra	42.78	-8.43	225	29.1	Csb	1736	279
Verín	41.97	-7.40	546	105.7	Csb	2141	397
Vigo	42.17	-8.69	460	7.6	Csb	1942	295
Vigo	42.24	-8.73	7	0.0	Csb	1122	351
Vilalba	43.23	-7.78	684	34.8	Csb	2773	54
Vilanova de Arousa	42.58	-8.80	3	0.0	Csb	1372	244
Viveiro	43.63	-7.63	59	4.0	Cfb	1738	138

2.5.3. Analysed datasets

Five datasets, shown in Figure 5, were collected to perform the analysis: reference (Ref.), nearest neighbour applied to MeteoGalicia data (NN MG), universal kriging applied to MeteoGalicia data (UK MG), nearest neighbour applied to GFS data (NN GFS) and universal kriging applied to GFS data (UK GFS).

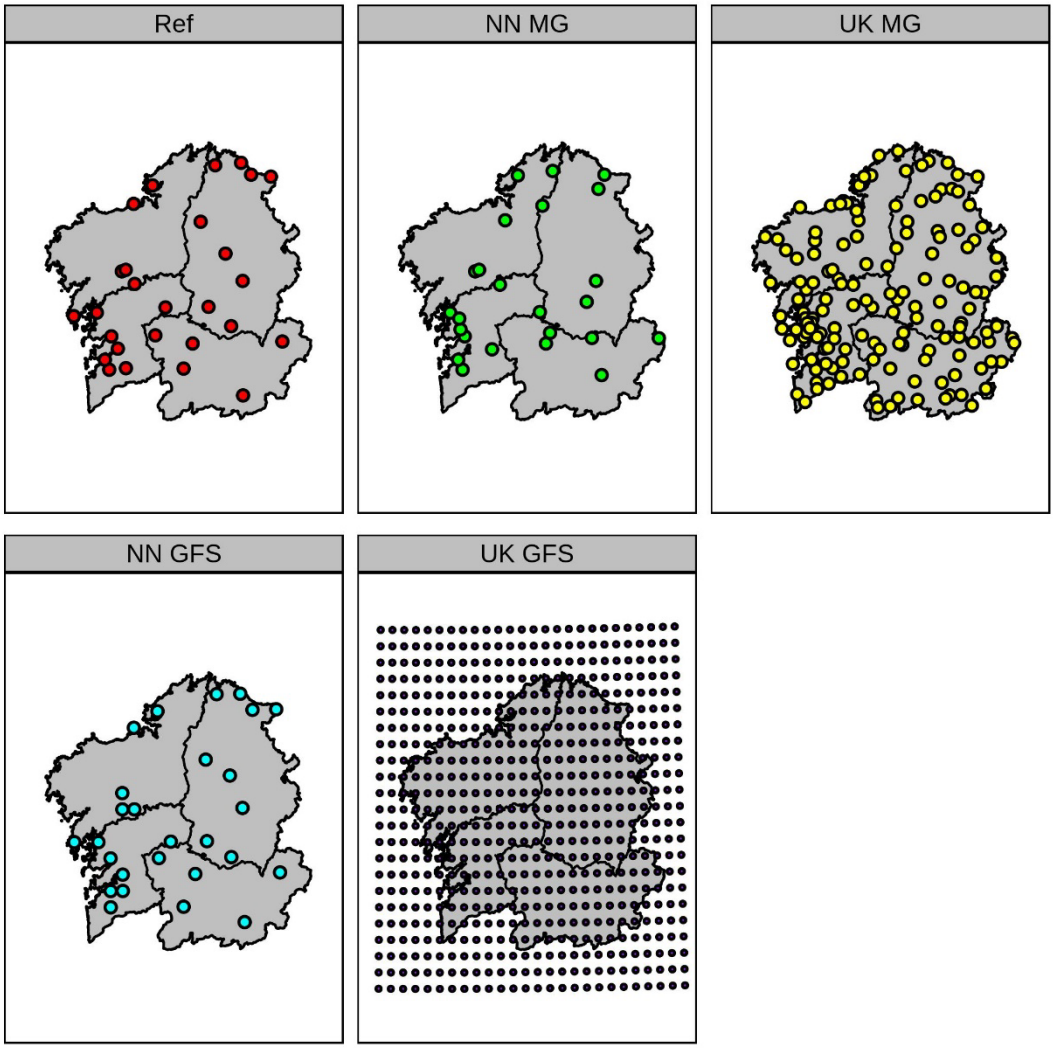

Figure 5: Datasets locations

The Ref. dataset consisted of the 27 MeteoGalicia weather stations assigned to the corresponding weather and simulation scenarios, and the geographical coordinates of these points were used as estimation targets for the interpolation techniques. As real measured data, the Ref. dataset was used as a baseline against which the accuracy and precision of the estimations of the other four datasets were tested.

To implement the nearest neighbour spatial interpolation technique over the MeteoGalicia weather data, it was necessary to apply a selection algorithm. This algorithm was designed to ensure that no station from the Ref. dataset could act as its own nearest neighbour. The criterion of only using stations with at least 75% of available data hours for the studied temporal span was also applied. Nevertheless, a single station would be allowed to be the nearest neighbour for more than one weather scenario. Once the algorithm was applied, the NN MG dataset consisted of 24 locations to be linked to the 27 studied scenarios by means of geographical proximity. Hence, each studied location would have a single nearest neighbour MG station, but each MG station could act as a nearest neighbour of more than one studied location.

The UK MG dataset consisted of all the MeteoGalicia weather stations with at least one measurement for the study temporal interval. A total of 149 stations comprised this dataset and were used for each interpolation, although the farthest points had a smaller influence on the interpolation results than closer points based on the mathematical foundations of kriging interpolation techniques. Each station belonging to the Ref. dataset was temporally removed from

this UK MG dataset when estimating weather data at such location, so falsely perfect estimation scores would not be achieved.

The NN GFS dataset consisted of the sum of the GFS points closest to each of the studied locations, similar to that of NN MG. 27 GFS points were included in this dataset, each linked to one studied location. Unlike with the NN MG dataset, no NN GFS point ended up acting as a nearest neighbour to more than one study location, due to the particular distribution of the GFS points and studied locations.

The last dataset, UK GFS, was composed of 598 points. These points were selected after applying a spatial filter to all the GFS points stored in the local database: a rectangular shape containing all the weather scenarios was defined, and then an outer buffer of 0.5° was applied. There was no need to remove points in this case, because the Ref. and UK GFS datasets had no common elements.

Although both the MG and the GFS nearest neighbours were assigned to each studied location using the same methodology, the differences in the underlying sets of available points caused a notable difference in the resulting distances between pairs of points: the GFS points were in a regular, relatively dense grid, so the distance between each reference location and its nearest GFS point had a limited maximum value. The maximum distance corresponded to half the diagonal of the smallest square defined by the four surrounding grid points. For the GFS sflux model grid, whose points were separated by approximately 13 km, this maximum NN distance was close to 18 km. However, the MG weather stations were scattered in the study region rather than placed on a regular lattice, which meant that the theoretical maximum distance limit present on the GFS neighbours was not applicable to NN MG. The distances between each location and its corresponding MG and GFS nearest neighbours are presented in Table 4.

Table 4: Nearest Neighbour distances to locations

Study location	Distance to NN MG [km]	Distance to NN GFS [km]	Study location	Distance to NN MG [km]	Distance to NN GFS [km]
A Coruña	19,54	4,88	Redondela	11,11	5,04
A Veiga	18,65	1,97	Ribadeo	15,87	4,72
Avión	17,31	6,47	Ribeira	18,56	3,21
Burela	12,40	0,99	Santiago de Compostela	3,33	6,46
Celanova	21,31	6,50	Santiago de Compostela	3,34	6,06
Chantada	18,02	3,65	Sarria	18,31	2,38
Ferrol	12,47	5,15	Vedra	9,50	3,71
Foz	12,34	4,26	Verín	16,78	2,95
Lalín	17,17	5,42	Vigo	8,69	6,20
Lugo	25,76	7,35	Vigo	8,69	4,71
Monforte de Lemos	11,13	0,75	Vilalba	15,79	7,48
Ourense	9,17	2,18	Vilanova de Arousa	8,70	1,34
Ponteareas	13,35	6,28	Viveiro	13,45	1,36
Pontevedra	6,47	6,42	-	-	-

2.6. Error measurements

To assess the performance of the interpolation techniques and data sources, several evaluation measures were considered: mean bias error (MBE), mean absolute error (MAE), root mean square error (RMSE) and root mean square error coefficient of variation (CV(RMSE)). These measures were computed using the equations shown below:

$$MBE = \sum_{i=1}^{n} \frac{X_i - Y_i}{N} \quad (4)$$

$$MAE = \sum_{i=1}^{n} \frac{|X_i - Y_i|}{N} \quad (5)$$

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(X_i - Y_i)^2}{N}} \quad (6)$$

$$CV(RMSE) = RMSE / \left(\frac{\sum_{i=1}^{n} Y_i}{N}\right) \quad (7)$$

where X_i is a particular estimated value, Y_i is the corresponding observed value, and N is the sample size. Differences, absolute differences and squared differences between the hourly estimation values and the observation values were computed for all locations and all datasets. As already stated, some hourly weather estimations were missing due to not having available data for interpolation and were filled with kriging estimations during the generation of EPW files. Only rows with no gaps (for the corresponding weather variable) were considered when computing error results. Simulated hourly results calculated using such filled values (for any meteorological variable) were also ignored for heating demand error evaluation.

3. Results and Discussion

3.1. Weather variables results

A statistical study was carried out for the six main variables used as inputs for the TRNSYS simulations. The number of valid estimated hours generated varied slightly between datasets and variables due to gaps in both the estimation and the reference datasets. The percentage of available results ranged between 91.25% and 96.96% of the maximum possible (i.e., 8760 hours per year for each location, with a total of 27 locations).

Figure 6 shows a comparison of the MBE and MAE results with box-and-whisker plots for all weather variables. The RMSE and CV(RMSE) plots showed the same pattern as the MAE plots, so those plots were omitted. Boxes show the interquartile range (distance between the first and third quartiles) and the median value. Whiskers extend from the lower (and higher) quartiles to the lowest (and highest) values within 1.5 times of the interquartile range. Black diamonds represent data outliers. For the surface level pressure boxplots, some extreme outliers were removed to allow a better visual comparison. Two outliers were removed for MBE (-9.8 hPa for NN MG and -5.9 hPa for UK MG) and two for MAE (11.1 hPa for NN MG and 6.5 hPa for UK MG). For the MBE plots only, an auxiliary line was drawn cutting the ordinate axis at 0 to help distinguish positive from negative errors.

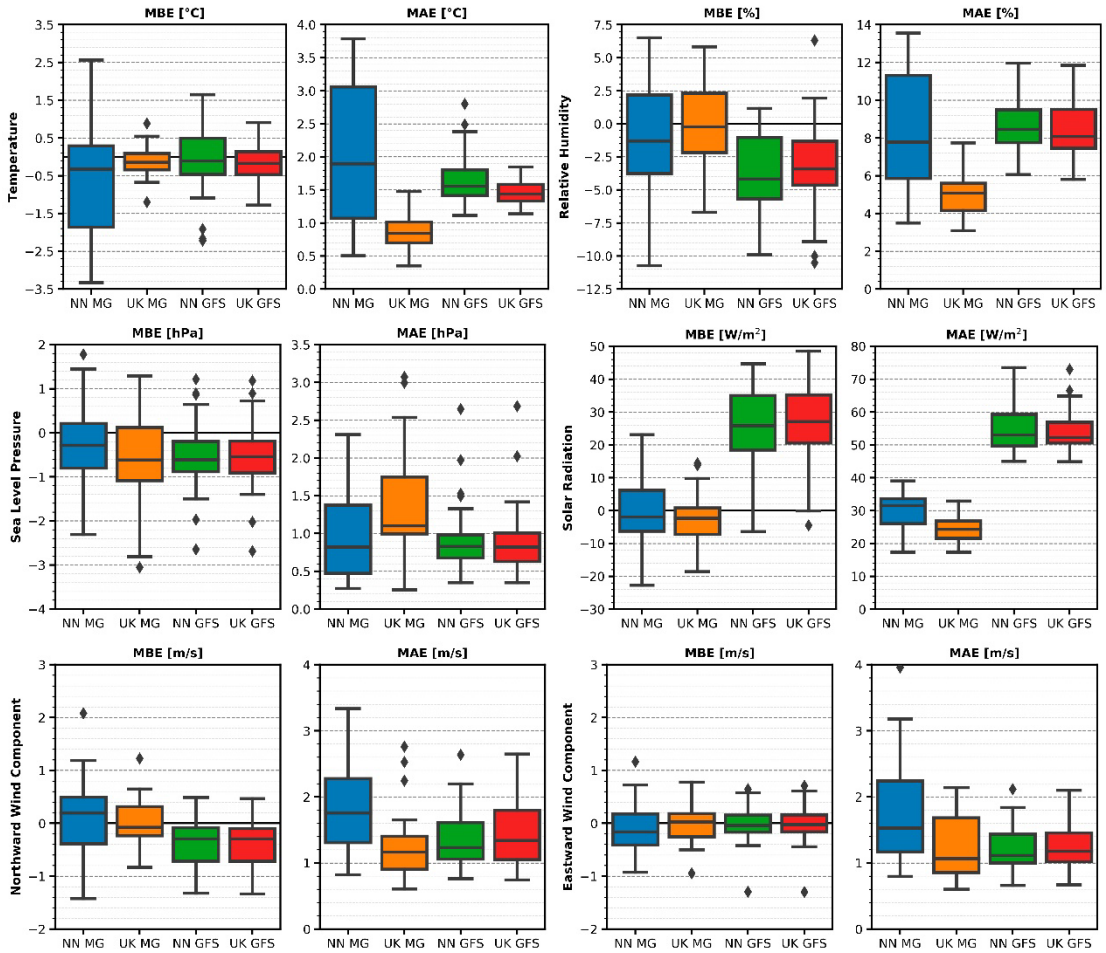

Figure 6: MBE and MAE box-and-whisker plots for weather variables

Table 5 summarizes the main statistics regarding the estimation values calculated for all the weather variables.

Table 5: Statistical summary of weather estimation values

Variable	Statistic	Datasets				
		Ref.	NN MG	UK MG	NN GFS	UK GFS
Temperature [°C]	Mean	13.38	12.83	13.28	13.13	13.12
	Median	12.91	12.49	12.83	12.42	12.41
	Standard deviation	6.34	6.54	6.28	6.56	6.49
Relative humidity [%]	Mean	81.12	80.09	80.91	77.43	77.75
	Median	85.00	84.00	85.33	82.32	82.55
	Standard deviation	16.19	17.02	14.99	17.52	17.26
Sea level pressure [hPa]	Mean	1017.57	1016.96	1016.84	1017.21	1017.21
	Median	1018.15	1017.89	1017.65	1017.64	1017.64
	Standard deviation	8.28	11.39	9.80	8.22	8.22
Horizontal global solar radiation [W/m²]	Mean	152.37	151.86	149.86	176.20	176.03
	Median	6.00	6.00	5.91	0.00	0.00
	Standard deviation	242.06	242.46	233.86	268.76	268.32
	Mean	0.51	0.64	0.52	0.14	0.16

Wind northward component [m/s]	Median	0.33	0.36	0.20	-0.02	-0.01
	Standard deviation	2.48	2.85	1.97	2.29	2.41
Wind eastward component [m/s]	Mean	0.23	0.16	0.24	0.20	0.22
	Median	0.08	0.07	0.11	-0.01	0.01
	Standard deviation	2.18	2.72	1.83	1.88	1.96

UK MG was the best local weather estimation dataset, as it provided the most accurate results (MAE errors) for temperature, relative humidity, horizontal global solar radiation and northward wind component. UK MG was close to GFS datasets for providing the best accuracy on the eastward wind component variable and to UK GFS for pressure. The nature of the weather values provided by MeteoGalicia was arguably the reason for this better performance, as they were real measurements, as opposed to the GFS forecast values.

NN MG had a much worse general performance than its kriging counterpart, with the greatest dispersion levels in all but sea level pressure and solar radiation. The differences between the NN GFS and UK GFS datasets were less relevant, with UK GFS performing better than its nearest neighbour counterpart on temperature and solar radiation, worse on wind components, and similar in relative humidity and sea level pressure.

The smaller gap between the two GFS interpolations (when compared with their MG counterparts) can be explained by means of the distribution of the GFS points. As already shown in Table 4, the actual NN GFS distances to each studied location were smaller than the corresponding NN MG points. Only for two scenarios were the NN GFS distances larger than the NN MG distances. These GFS distances were also far less than their maximum theoretical value (~18 km). This large difference in distances is a plausible explanation for the overall better performance of the NN GFS interpolation compared to that of the NN MG dataset.

Precision errors (RMSE and CV(RMSE)) followed the same general pattern as MAE errors, with similar relative performance scores between the different datasets for the weather variables. Thus, plots for those two precision error metrics are not included.

Overall, the four estimation datasets tended to underestimate the weather variables, according to the MBE median values. However, the horizontal global solar radiation bias errors shown in Figure 6 were dependent on the weather data source used. That is, the MG datasets underestimated and the GFS datasets overestimated the reference values. NN MG overestimated the northward wind component, while its kriging counterpart slightly underestimated it. The reverse situation was found for the eastward component, although the differences were smaller. The two GFS datasets behaved similarly when underestimating or overestimating each variable. Excluding relative humidity and solar radiation, the median MBE values were similar for all the estimation datasets, while the first and third quartiles and the values outside them showed greater differences.

In most cases, the differences between the reference values and the interpolation results for each of the four estimated datasets were not severe. From the CV(RMSE) results, which allowed direct comparisons between different variables, solar radiation was the worst estimated variable for all the datasets, while pressure was variable with the most accurate estimations.

The analysis performed in this section highlighted the best fit with the reference data provided by the UK MG dataset, arguably due to the combination of the MeteoGalicia dense network of observation stations and the use of the universal kriging interpolation technique. Although the GFS sflux grid had more points in the study area than the MeteoGalicia network, the nature of the weather outputs at each point was quite different: each station provided observational measurements, while each GFS grid point provided an estimated value.

The comparison between NN MG and UK GFS and between the two GFS datasets, however, highlighted a different perspective: Even if the GFS outputs had an inherently higher uncertainty than that of the MeteoGalicia observations, both the NN and UK GFS datasets provided better estimations than NN MG for almost all variables (being worse only for solar radiation). The differences between NN GFS and UK GFS were far less than those between their MeteoGalicia counterparts. As already stated, the GFS grid was regular, which means that any selected location would have a nearest point available at a determined maximum distance of approximately 18 km. In contrast, the MeteoGalicia network was non-regular and scarcer in some areas. Moreover, the nearest station may not be available at a given moment, which means NN would potentially pick an input value farther away when using the MeteoGalicia database instead of GFS, as already indicated by the results presented in Table 4.

3.2. Thermal simulation results

The same analysis pattern used for the weather variables was applied to the heating demand results. For this variable, the amount of valid available estimation results was compulsorily lower than that of the meteorological variables because any gap in any of the six TRNSYS weather inputs would cause the heating demand output to be considered invalid. The percentage of available results ranged between 86.37% and 96.16% for the different studied scenarios. Figure 7 shows a heating demand comparison for all the error metrics. The same considerations used for the plots of meteorological variables were kept here. By design, the heating demand estimations were necessarily zero outside the heating season months and zero for the non-heating hourly interval for each day of the heating season. If the error metrics were to be computed considering the non-heating days and hours, the results would be strongly skewed towards zero, artificially reducing the MBE, MAE and RMSE values and increasing the CV(RMSE) values (due to having a mean observation value skewed towards zero). To avoid this issue, only the hours inside the 15:00-22:00 UTC+1 (included) interval from January 1st to March 31th and from October 20th to December 31th were considered when computing the heating demand error metrics.

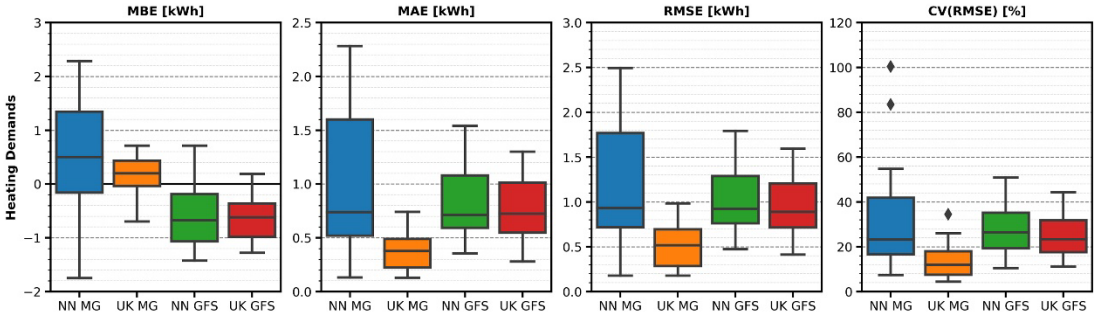

Figure 7: MBE, MAE, RMSE and CV(RMSE) box-and-whisker plots for heating demands

Table 6 summarizes the main statistics regarding the estimation values calculated for this variable. As with the previous plots, non-heating days and hours were removed before computing the statistics.

Table 6: Statistical summary of heating demand estimation values

Variable	Statistic	Datasets				
		Ref.	NN MG	UK MG	NN GFS	UK GFS
Heating demand [kWh]	Mean	3.97	4.47	4.13	3.39	3.38
	Median	4.01	4.41	4.21	3.38	3.34
	Standard deviation	2.35	2.62	2.31	2.15	2.15

For all the meteorological results presented in the previous section, a sensible order of magnitude can be intuitively determined, allowing the fast-checking of flawed values (e.g., an air temperature of 80 °C or an average wind speed of 200 m/s). As heating demands are substantially dependent on the type of building model, a reference order of magnitude may be hard to establish without more information. For this reason, the yearly total heating demands were computed and are provided in Table 7 for the 27 locations using TRNSYS outputs from the Ref. dataset.

Table 7: Annual heating demands for the studied locations

Study location	Latitude	Longitude	Heating demands [kWh/m²year]	Study location	Latitude	Longitude	Heating demands [kWh/m²year]
A Coruña	43.36	-8.44	11.8	Redondela	42.32	-8.60	16.6
A Veiga	42.36	-7.01	31.4	Ribadeo	43.54	-7.08	11.9
Avión	42.41	-8.24	19.9	Ribeira	42.56	-9.03	10.3
Burela	43.64	-7.37	17.6	Santiago de Compostela	42.88	-8.56	16.6
Celanova	42.17	-7.97	19.4	Santiago de Compostela	42.89	-8.52	17.7
Chantada	42.61	-7.72	19.8	Sarria	42.80	-7.38	23.4
Ferrol	43.49	-8.25	13.0	Vedra	42.78	-8.43	15.9
Foz	43.56	-7.28	14.7	Verín	41.97	-7.40	20.5
Lalín	42.61	-8.14	21.8	Vigo	42.17	-8.69	16.0
Lugo	42.99	-7.54	21.2	Vigo	42.24	-8.73	8.4
Monforte de Lemos	42.47	-7.50	21.5	Vilalba	43.23	-7.78	23.6
Ourense	42.35	-7.88	15.1	Vilanova de Arousa	42.58	-8.80	10.7
Ponteareas	42.18	-8.52	15.3	Viveiro	43.63	-7.63	16.2
Pontevedra	42.41	-8.66	10.9	-	-	-	-

The mean heating demand results showed absolute differences for NN MG, UK MG, NN GFS and UK GFS of 11.9%, 3.4%, 13.6% and 11.9% respectively. As expected, the use of UK MG resulted in consistently lower errors for heating loads. UK MG had the lowest bias and the highest accuracy and precision and was also the dataset with the least dispersion. NN MG was the worst dataset. Although the median values of the MAE, RMSE and CV(RMSE) metrics were not far from those of the GFS datasets, they showed a greater dispersion. The differences between NN GFS and UK GFS were again of a lesser order, with UK GFS offering slightly better results in all but the CV(RMSE) metrics.

The comparison of the MBE results showed that both MeteoGalicia estimation datasets tended to overestimate heating demands, while both GFS datasets tended to underestimate heating demands. This finding was in contrast to the MeteoGalicia underestimation of weather conditions. As a colder environment was estimated, it was natural that simulations fed with MeteoGalicia data would predict higher energy consumptions, thus overestimating heating needs. GFS underestimated most weather conditions, but horizontal global solar radiation was consistently overestimated at greater scale than the underestimation of other weather variables. Following the same logic applied to the MeteoGalicia results, the large solar radiation overprediction may have caused the estimation of a warmer environment, thus causing the GFS-fed simulations to compute a lower energy consumption, underestimating heating demands.

When comparing the CV(RMSE) behaviour patterns of the different weather and simulation variables, the heating demand results were similar to those of temperature and relative humidity.

That is, UK MG has a prominent performance, with both GFS datasets being ranked between UK MG and NN MG. This was not the case for solar radiation and the pressure and wind variables, as already stated in the previous section.

The high dispersion for the heating demands obtained by all the estimation datasets was reflected both in the standard deviation values presented in Table 6 and in the amplitudes of the whisker plots presented in Figure 7. To highlight this feature, errors were computed by applying the MAE metric formula over the estimation results grouped by two criteria: by day, on a first test, and by day and location, on a second test. The results are shown in Figure 8. In the first individual plot, MAE was computed for each day, combining the results from the whole set of 27 locations. A single line represents the temporal evolution of the MAE value for each estimation dataset. In the second set of plots, MAE was computed separately for each location and each day. The thick lines represent the evolution of the mean values for the MAE results on each day, while the lightly coloured buffer band around those lines represents the 95% certainty interval. Minor ticks on the x-axis were added to mark odd month days. All plots were cut to avoid representing non-heating months. Gaps during the last four days of December caused these days to not be plotted. The same assumption regarding non-heating days and hours from Figure 7 was used.

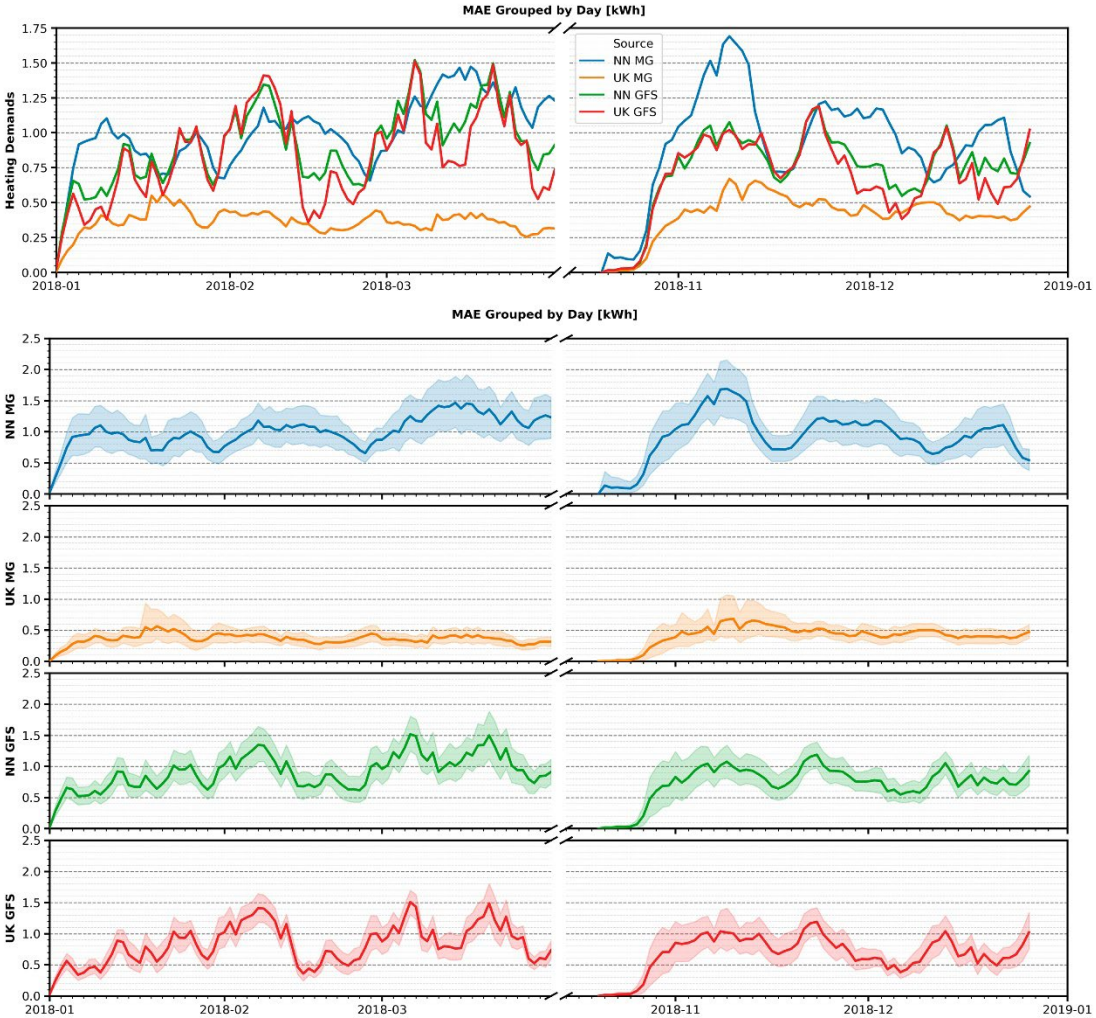

Figure 8: MAE results for heating demands grouped by day: combined and split with buffer band

Figure 8 shows that the TRNSYS daily heating demand errors had a strong temporal dependency with differences up to 150%. The UK MG dataset displayed a regular behaviour with the less error all year long (except for one day in January and two days in December, where it was outperformed by UK GFS). The UK MG buffer band with 95% confidence was also much smaller

than those of the other datasets, showing notable differences only during the second half of January, the last days of October and the first half of November, when the remaining datasets also showed this behaviour. NN MG had the worst performance during the October-December period, but had much fewer differences with the GFS datasets during January-March, even performing better on several days. The NN MG dataset displayed the greatest buffer band, with relatively large differences during almost the entire study span. Both GFS datasets had a more irregular pattern than those of their MG counterparts and were more rugged but had a smaller buffer band than that of NN MG. UK GFS performed slightly worse than NN GFS during 34 days and was quite similar for most of the year.

As with the weather variables, UK MG provided the best fit with the reference data for heating demands. NN MG, NN GFS and UK GFS were ranked as the fourth, third and second best estimations, respectively, similar to temperature and relative humidity. The CV(RMSE) values fell within the ASHRAE acceptance limits (from 10% to 30%) for the UK MG method in almost all locations. Median and the 1st and 3rd quartiles stayed well below 20% for this dataset, as already shown in Figure 7. For UK GFS, the 3rd quartile reached a value slightly over 30%, with the median and 1st quartile values close to 20%.

3.3. Study limitations and future research

After showing and discussing the main results obtained, an acknowledgement of the key weaknesses and limitations of this study is essential, along with proposed future works that could amend these limitations.

The temporal span of the present study was a full year, which provided information about seasonal variations on the weather variables used in the thermal simulations. However, inter-annual variability was not captured, which means that the particular results presented in this study could only be directly applicable to the particular meteorological conditions developed during that year. However, the general methodology of this study can still be applied over wider time spans. Future work comprising more than one year of meteorological data will help to extend the obtained conclusions to a wider set of potential weather conditions.

The 27 selected scenarios offered adequate coverage over the studied region. However, the studied region as a whole shared a few common traits in regard to its geographical and meteorological features: low and mid-low altitudes, relative proximity to the open ocean and similar external weather influences. A subsequent step would be to extend this study to a larger area with more climatic and orographic variability. A different kind of meteorological observation network with a lower density of stations scattered over a more heterogeneous region is already planned to be used as a test comparison in future work.

In addition to the regional-scale geographical variability issue, the selection of those studied locations was limited by the presence of meteorological weather stations. As stations are usually located on cleared, open spaces, there was an intrinsic limitation on the kind of local microclimates that could be covered. The reference dataset comprised weather stations located in both rural and urban areas. However, the urban stations were located in open spaces (such as parks, suburbs and large avenues) rather than on narrow streets, which means that the specific study results could not be directly applied to buildings located on these latter kind of streets, despite the general methodology still being valid.

Finally, the thermal simulations of the building model were affected by the specifications of the previously mentioned building regulations and the meteorological conditions of the studied area. The discarding of cooling resources in the study simulations was a direct result of these constraints and was directly linked with the aforementioned issue relating to the variability in the regional orographic and meteorological features. A further study including this point could be projected to

ensure that the results and conclusions obtained for heating demand simulations with GFS weather inputs are also applicable to cooling demands.

4. Conclusions

The main objective of this paper was to evaluate the use of a GFS model as a local weather data source for the thermal demand simulations of buildings. Four estimated weather datasets were elaborated for the complete year of 2018 using two data sources (observations from the regional weather network of MeteoGalicia and forecasts from the GFS sflux model) and two interpolation techniques (nearest neighbour and universal kriging), and were compared against a reference dataset of 27 MeteoGalicia weather stations.

For the main objective, the results highlighted the viability of GFS sflux as a weather source for building thermal simulations. Overall, the performance of GFS sflux was better than that of the nearest MeteoGalicia station used as a data source, although it was still not better than using the whole MeteoGalicia network for universal kriging interpolation of meteorological data (except for sea level pressure). The primary conclusions related to the main objective are as follows:

- GFS datasets overestimated horizontal global solar radiation, while MeteoGalicia underestimated it. All other weather variables were underestimated to some extent by both sources.

- MeteoGalicia overestimated heating demands due to its colder weather estimations. GFS underestimated heating requirements due to its large overestimations of solar radiation and the subsequent estimation of a warmer environment.

- Horizontal global solar radiation was the worst estimated weather variable for both the MeteoGalicia and GFS datasets. Sea level pressure was the best estimated variable for all datasets.

- All the errors in the weather and simulation variables were sensible for both the MeteoGalicia and GFS datasets.

- The MeteoGalicia dense network of meteorological stations and the observational nature of its data allowed UK MG to provide the best estimations of both the weather and the heating demand variables.

- Despite using forecasted data as inputs for interpolation, both GFS datasets performed better than NN MG for heating demands and all weather variables except solar radiation.

- The GFS regular grid of relatively close forecast points allowed a greater performance of NN GFS than NN MG. The same reason caused the smaller differences between UK GFS and NN GFS compared to that of their MeteoGalicia counterparts.

- The CV(RMSE) errors for UK MG fell within ASHRAE acceptance limits for almost all locations. For UK GFS, the errors fell slightly over the acceptance threshold for the 3rd quartile and inside for the 1st quartile and median.

In addition to these conclusions related to the performance of GFS, two more points were noted:

- The Thermal simulation error results for the studied building model were more similar to those obtained for temperature and relative humidity and less similar to those obtained for solar radiation, wind and pressure.

- Universal kriging interpolation largely performed better than nearest neighbour when both were used with a non-regular grid of observation points. The differences were far less when combined with a regular grid of input points.

Altogether, GFS was shown to be an appropriate contender when compared against a high-density meteorological station network over a relatively small study area. Both GFS interpolation estimations had an overall better performance than using real measurements from the nearest weather station. The main limitations of the specific results obtained have already been acknowledged, and future studies have been proposed. However, regarding the main scope of the present study, the GFS sflux model has already been revealed as a global-scale, free, good-quality source of local weather data for building heating demand simulations.

Data Availability

The meteorological datasets provided by MeteoGalicia used in this article can be found at https://www.meteogalicia.gal/observacion/estacions/estacions.action, on its web page [29].

The meteorological datasets provided by NOAA used in this article are not available due to being hosted in a short-term storage server. Data for the last ten days can be found at https://www.nco.ncep.noaa.gov/pmb/products/gfs/, on its web page [30], but not for the time span analysed in this article.

Author Contribution

J.L.G.: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization and Writing - Original draft. **F.T.P.:** Conceptualization, Formal analysis, Methodology, Software, Validation and Writing - Review & editing. **E.A.F.:** Formal analysis, Methodology, Validation, Writing - Original draft and Writing - Review & editing. **P.E.O.:** Conceptualization, Resources, Supervision, Visualization and Writing - Review & editing. **E.G.A.:** Funding acquisition, Project administration, Resources and Writing - Review & editing.

Acknowledgments and Funding

This investigation article was partially supported by the University of Vigo through the grant *Convocatoria de Axudas á Investigación 2018: Axudas Predoutorais UVigo 2018* (grant number 00VI 131H 641.02). This investigation article was also partially supported by the Galician Government by means of the Xunta de Galicia CONECTA PEME 2018 (Project: TOPACIO IN852A 2018/37).

This paper was carried out in the framework of the GIS-Based Infrastructure Management System for Optimized Response to Extreme Events of Terrestrial Transport Networks (SAFEWAY) project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 769255. Neither the Innovation and Networks Executive Agency (INEA) nor the European Commission is in any way responsible for any use that may be made of the information it contains.

The authors would like to thank to the Galician weather agency MeteoGalicia for providing free, unlimited public access to their historical weather registers. The authors would like to thank to the National Centres for Environmental Prediction (NCEP) for providing free public access through their NOMADS servers to their real time forecast datasets.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M. Elci, B. Manrique Delgado, H.M. Henning, G.P. Henze, S. Herkel, Aggregation of residential buildings for thermal building simulations on an urban district scale, Sustainable Cities Soc., 39 (2018) 537-547.
[2] H. Kikumoto, R. Ooka, Y. Arima, T. Yamanaka, Study on the future weather data considering the global and local climate change for building energy simulation, Sustainable Cities Soc., 14 (1) (2015) 404-413.
[3] L. Pérez-Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, Energy and Buildings, 40 (3) (2008) 394-398.
[4] D. Qian, Y. Li, F. Niu, Z. O'Neill, Nationwide savings analysis of energy conservation measures in buildings, Energy Conversion and Management, 188 (2019) 1-18.
[5] A.A. Kasam, B.D. Lee, C.J.J. Paredis, Statistical methods for interpolating missing meteorological data for use in building simulation, Building Simulation, 7 (5) (2014) 455-465.
[6] A. Thomas, C.C. Menassa, V.R. Kamat, A systems simulation framework to realize net-zero building energy retrofits, Sustainable Cities Soc., 41 (2018) 405-420.
[7] M. Vuckovic, K. Kiesel, A. Mahdavi, Toward advanced representations of the urban microclimate in building performance simulation, Sustainable Cities Soc., 27 (2016) 356-366.
[8] U. Berardi, M. Naldi, The impact of the temperature dependent thermal conductivity of insulating materials on the effective building envelope performance, Energy and Buildings, 144 (2017) 262-275.
[9] P. Eguía Oller, J.M. Alonso Rodríguez, Á. Saavedra González, E. Arce Fariña, E. Granada Álvarez, Improving transient thermal simulations of single dwellings using interpolated weather data, Energy and Buildings, 135 (2017) 212-224.
[10] P. Eguía Oller, J.M. Alonso Rodríguez, Á. Saavedra González, E. Arce Fariña, E. Granada Álvarez, Improving the calibration of building simulation with interpolated weather datasets, Renewable Energy, 122 (2018) 608-618.
[11] M. Bhandari, S. Shrestha, J. New, Evaluation of weather datasets for building energy simulation, Energy and Buildings, 49 (2012) 109-118.
[12] J. Antonanzas, R. Urraca, F.J. Martinez-de-Pison, F. Antonanzas-Torres, Solar irradiation mapping with exogenous data from support vector regression machines estimations, Energy Conversion and Management, 100 (2015) 380-390.
[13] J. López Gómez, F. Troncoso Pastoriza, E. Granada Álvarez, P. Eguía Oller, Comparison between Geostatistical Interpolation and Numerical Weather Model Predictions for Meteorological Conditions Mapping, Infrastructures, 5 (2) (2020)
[14] D. Carvalho, A. Rocha, M. Gómez-Gesteira, C. Silva Santos, WRF wind simulation and wind energy production estimates forced by different reanalyses: Comparison with observed data for Portugal, Applied Energy, 117 (2014) 116-126.
[15] S. Fernández-González, M.L. Martín, E. García-Ortega, A. Merino, J. Lorenzana, J.L. Sánchez, F. Valero, J.S. Rodrigo, Sensitivity Analysis of the WRF Model: Wind-Resource Assessment for Complex Terrain, Journal of Applied Meteorology and Climatology, 57 (3) (2017) 733-753.

[16] P. Zhao, J. Wang, J. Xia, Y. Dai, Y. Sheng, J. Yue, Performance evaluation and accuracy enhancement of a day-ahead wind power forecasting system in China, Renewable Energy, 43 (2012) 234-241.
[17] P. Eguía, E. Granada, J.M. Alonso, E. Arce, A. Saavedra, Weather datasets generated using kriging techniques to calibrate building thermal simulations with TRNSYS, Journal of Building Engineering, 7 (2016) 78-91.
[18] M. Kottek, J. Grieser, C. Beck, B. Rudolf, F. Rubel, World map of the Köppen-Geiger climate classification updated, Meteorologische Zeitschrifft, 15 (3) (2006) 259-263.
[19] M. Kottek, F. Rubel, World maps of Köppen-Geiger climate classification. Present climate., in, 2019.
[20] N. Hofstra, M. Haylock, M. New, P. Jones, C. Frei, Comparison of six methods for the interpolation of daily, European climate data, Journal of Geophysical Research: Atmospheres, 113 (D21) (2008).
[21] J. Li, A.D. Heap, A review of spatial interpolation methods for environmental scientists, Geoscience Australia, Record 2008/23, Canberra, Australia, 2008.
[22] A. Beygelzimer, S. Kakadet, J. Langford, S. Arya, D. Mount, L. Shengqiao, FNN: Fast Nearest Neighbor Search Algorithms and Applications, in, 2019.
[23] P.H. Hiemstra, E.J. Pebesma, C.J.W. Twenhöfel, G.B.M. Heuvelink, Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network, Computers & Geosciences, 35 (8) (2009) 1711-1721.
[24] B. Gräler, E. Pebesma, G. Heuvelink, Spatio-Temporal Interpolation using gstat, The R Journal, 8 (1) (2016) 204-218.
[25] R.a.A.-c.E. American Society of Heating, ASHRAE Guideline 14-2002. Measurement of energy and demand savings, in, American Society of Heating, Refrigerating and Air-conditioning Engineers (ASHRAE), Atlanta, USA, 2002.
[26] I.O.f. Standardization, Standard Atmosphere, International Standard ISO 2533:1975, in, International Organization for Standardization, Switzerland, 1975, pp. 108.
[27] I.S.F.S.U.C.f.A. Research, ISFS Guides. Wind direction quick reference, in, 3090 Center Green Dr. Boulder, Colorado 80301, 2019.
[28] X.d. Galicia, Anuario climatológico de Galicia 2010, in, Xunta de Galicia, Santiago de Compostela, Spain, 2013.
[29] X.d. Galicia, MeteoGalicia. Observación. Rede meteorolóxica., in, 2019.
[30] NOAA, NCEP Products Inventory, in, National Oceanic and Atmospheric Administration (NOAA), 2018.
[31] M.d. Fomento, Código Técnico de la Edificación. Documento Básico DB-HE "Ahorro de Energía", in, Ministerio de Fomento, Madrid, Spain, 2017.
[32] I. Instituto para la Divesificación y Ahorro de la Energía, Condiciones de aceptación de procedimientos alternativos a LIDER y CALENER, in, Madrid, Spain, 2009.
[33] A.-T. Nguyen, S. Reiter, P. Rigo, A review on simulation-based optimization methods applied to building performance analysis, Applied Energy, 113 (2014) 1043-1058.
[34] M. Wetter, J. Wright, Comparison of a generalized pattern search and a genetic algorithm optimization method, Eighth International IBPSA Conference, 3 (2009).