Building Technologies & Urban Systems Division
Energy Technologies Area
Lawrence Berkeley National Laboratory

Quantifying and simulating the weather forecast uncertainty for advanced building control

Wanfu Zheng[1,2], Laura Zabala[3,4], Jesus Febres[3], David Blum[5], Zhe Wang[1,2]

[1]Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, [2]HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, [3]R2M Solution, [4]Faculty of Engineering of Bilbao, University of the Basque Country UPV/EHU, [5]Building Technology & Urban Systems Division, Lawrence Berkeley National Laboratory

Energy Technologies Area
Month 2025

doi.org/10.1080/19401493.2025.2453537

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Building Technologies Office, of the US Department of Energy under Contract No. DE-AC02-05CH11231.

Disclaimer:

Quantifying and simulating the weather forecast uncertainty for advanced building control

Wanfu Zheng,[1,2,#] Laura Zabala,[3,4,#] Jesus Febres,[3] David Blum,[5] Zhe Wang[1,2,*]

[1]Department of Civil and Environmental Engineering,
The Hong Kong University of Science and Technology, Hong Kong, China
[2]HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute
Futian, Shenzhen, China
[3] R2M Solution Spain, Bilbao, Spain
[4] Faculty of Engineering of Bilbao, University of the Basque Country UPV/EHU, Bilbao, Spain
[5] Building Technology & Urban Systems Division, Lawrence Berkeley National Laboratory,
Berkeley, CA, USA
#These two authors contribute equally to this paper
*Corresponding author: cezhewang@ust.hk

ABSTRACT

Weather forecast uncertainty is unavoidable despite technological advancements. Accurately quantifying and modeling this uncertainty is essential for developing and comparing advanced building controllers. In this study, we present a structured approach using a first-order autoregressive model (AR(1)) to model uncertainty in ambient temperature and global solar irradiation (GHI) forecasts. We analyzed weather data from four cities and employed Jensen-Shannon divergence (JSD) to evaluate the similarity between synthetic and actual forecast errors. The average JSD values for temperature are 0.027 (Berkeley), 0.021 (Leuven), 0.018 (Berlin), and 0.008 (Oslo), and for GHI, the average JSD values are 0.016 (Berkeley), 0.058 (Leuven), and 0.013 (Berlin). The low JSD values indicate a high similarity between the synthetic and real

forecast error distributions. Our approach successfully generates synthetic weather forecasts that mirror the statistical properties of actual forecasts. The implementation of our method for uncertain forecast generation is being added to the BOPTEST framework.

KEYWORDS

Weather forecast uncertainty; Autoregressive model; Ambient temperature; Global solar irradiation; Jensen-Shannon divergence

1 Introduction

Weather forecasts are essential inputs for prediction-based building control strategies, such as Model Predictive Control (MPC) (Joe and Karava 2019; Bianchini et al. 2019; Drgoňa et al. 2020) and Reinforcement Learning (RL) (Wei et al. 2017; Du et al. 2021). By incorporating weather forecasts, these control strategies can optimize pre-cooling or pre-heating system operation to enhance energy efficiency and cost savings (Wang and Hong 2020), or provide load flexibility to respond to dynamic electricity prices or carbon emissions factors (Chaturvedi and Rajasekar 2022), coordinate on-site storage usage, or maximize share of renewable generation (Drgoňa et al. 2020; Ma et al. 2021).

However, the weather forecast is inherently uncertain and prone to errors due to various factors (Dong et al. 2021), such as the chaotic nature of the atmosphere (Reichman and Dubowski 2021), inaccurate model formulations based on physical approximations (Meng et al. 2021) and a lack of knowledge about the initial state (Gagne et al. 2020; Flato et al. 2014). Considerable research effort has been devoted to modeling the uncertainty of weather forecasts and considering this uncertainty in control algorithms. These techniques include statistical models using time series models such as autoregressive (AR) (Bacher et al. 2009) and autoregressive integrated moving average (ARIMA) models (Naiqian et al. 2019) or the application of artificial neural

Nomenclature

Variables, parameters, and indices

A, B, P, Q	Probability distributions
τ	Lead time (number of steps ahead in the forecast)
k	Time step index
$\widetilde{X}(\tau\|k)$	Predicted value for the lead time τ at timestep k
$X(k+\tau)$	Measured value at time step $k+\tau$
$e(\tau\|k)$	Forecast error at lead time τ and time step k
y_t	Predicted forecast error at time t in AR model
ϕ_i	Coefficients of the AR model
ε_t	Residual term in AR model at time t
p	Order of the autoregressive model
h	Lag in partial autocorrelation function (PACF)
ϕ	Autoregressive coefficient in AR(1) model
ε	Residual term in AR(1) model
μ	Mean of a distribution
σ	Standard deviation of a distribution
u	Location parameter of Laplace distribution
b	Scale parameter of Laplace distribution
$\hat{e}(\tau\|k)$	Synthetic forecast error at lead time τ and time step k
$\hat{X}(\tau\|k)$	Synthetic forecast at lead time τ and time step k

Operators

PACF(h)	Partial autocorrelation function at lag h
Cov$(\cdot,\cdot)$	Covariance operator
Var$(\cdot)$	Variance operator
JSD$(A\|\|B)$	Jensen-Shannon divergence between distributions A and B
KLD$(A\|\|B)$	Kullback-Leibler divergence between distributions A and B

Abbreviations

GHI	Global Horizontal Irradiation
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
ANN	Artificial Neural Network
NWP	Numerical Weather Prediction
PACF	Partial Autocorrelation Function
RMSE	Root Mean Squared Error
KLD	Kullback-Leibler divergence
JSD	Jensen-Shannon divergence
MPC	Model Predictive Control
RMPC	Robust Model Predictive Control
SMPC	Stochastic Model Predictive Control

networks (ANNs) (Sorkun et al. 2017).

Numerical Weather Prediction (NWP) represents the forefront of weather forecasting technology. NWP applies partial differential equations to model atmospheric dynamics based on initial conditions, integrating small-scale processes through physical parameterizations (Waqas et al. 2024). Global NWP techniques provide high-quality forecasts up to 15 days ahead, which are widely used by weather services worldwide. However, even with sophisticated models and computational power, NWP cannot eliminate forecast errors entirely.

In the context of building simulations and advanced control applications, it is crucial to not only use weather forecasts but also to assess and incorporate their inherent uncertainties into simulation tools. Modeling weather forecast uncertainty allows researchers and practitioners to better replicate real-world conditions, enabling the development and testing of robust control strategies for such uncertainties. Simulation tools can more accurately predict building performance by accounting for forecast uncertainty, leading to more reliable and efficient control strategies.

Despite the importance of incorporating weather forecast uncertainty, many advanced building control strategies often neglect this aspect by using deterministic weather forecasts in simulation studies (Zheng et al. 2024) or in deterministic MPC implementations in real case studies (Wang et al. 2023). As a result, the controllers developed may risk violating operational constraints, such as thermal comfort limits, or exhibit undesirable behaviors like ineffective pre-cooling or pre-heating when actual weather conditions deviate from forecasts.

Accurately modeling weather forecast uncertainty serves two primary purposes. Firstly, it bridges the gap between idealized simulations and real-world implementations of building control systems by representing forecast uncertainty within building simulation tools. This approach enables more realistic assessments of controller performance under uncertain conditions, ensuring that strategies developed in simulation can perform reliably in practice. Secondly, it

is essential for implementing advanced control techniques such as Robust Model Predictive Control (RMPC) (An et al. 2024) and Stochastic Model Predictive Control (SMPC) (Wu et al. 2015). These control methods explicitly account for forecast uncertainty and can provide control decisions resilient to potential disturbances caused by uncertain weather forecasts.

Therefore, this study focuses on assessing and incorporating weather forecast uncertainty into building simulation tools, particularly for advanced control applications. By modeling this uncertainty, we aim to enhance the robustness and reliability of building control strategies, ensuring they perform effectively under real-world conditions where weather forecasts are imperfect.

1.1 Related studies

Accurately modeling weather forecast uncertainty is essential for developing robust building control strategies. In this context, Oldewurtel et al. (Oldewurtel et al. 2012) compared stochastic Model Predictive Control (MPC) with deterministic MPC and found that the stochastic approach led to fewer comfort violations, highlighting the importance of considering uncertainties in building control strategies. Hedegaard et al. (Hedegaard et al. 2018) found that forecast uncertainties can cause MPC to be up to 4% less effective in reducing heating costs. (Enríquez et al. 2016) demonstrated that not considering solar irradiation forecasts' uncertainty significantly increases thermal discomfort when using MPC for indoor temperature control. Researchers have employed various approaches, including statistical models, machine learning techniques, and methods that account for increasing uncertainty over the forecast horizon.

Several studies have utilized statistical models to simulate weather forecast uncertainties. Oldewurtel et al. (Oldewurtel et al. 2012) proposed an AR model with Gaussian residuals for forecasting temperature and solar radiation errors but did not validate them against real data. Kuijpers et al. (Kuijpers et al. 2022) developed an AR model with an additive term for variables

like temperature and humidity, using historical observations and assuming normally distributed residuals with time-varying standard deviations. Grant and Gehbauer (Grant and Gehbauer 2022) assessed the impact of forecast accuracy on MPC by generating synthetic forecasts with normally distributed errors but excluded extreme values, potentially underrepresenting extreme weather events.

Jiang et al. (Jiang et al. 2023) addressed forecast uncertainty in energy dispatch by Stochastic MPC strategy with deep learning forecasting models. They estimated prediction errors at each forecast horizon on the training data, assuming normally distributed errors, and generated multiple scenarios by adding Gaussian noise to the predicted values. Similarly, Halev et al. (Halev et al. 2024) added Gaussian noise with a standard deviation increasing logarithmically over time to model decreasing forecast accuracy, applying this to load and photovoltaic time series.

Machine learning techniques have been applied to better capture uncertainties and correlations in weather patterns. Liu et al. (Liu et al. 2018) focused on probabilistic forecasting of solar irradiance using a time-series AR model that incorporated sky-cover forecasts and modeled sky conditions as discrete random variables. They employed a nonlinear, non-Gaussian state-space model with Sequential Monte Carlo methods for parameter estimation, generating probabilistic forecasts that better capture uncertainties associated with varying sky conditions. (Scher and Messori 2018) outlines other machine learning techniques to model prediction uncertainty, including ensemble numerical models that assess forecast spread, convolutional neural networks trained on past forecast errors or spreads, and statistical methods such as clustering weather types or using persistence and local dimensions in phase space.

Recognizing that forecast accuracy decreases with the prediction horizon, some researchers have modeled uncertainty as a function of time. Gao et al. (Gao et al. 2023) proposed a tube-based robust MPC for HVAC systems, where the uncertainty space widens over time, enhancing the

robustness of control strategies by accounting for growing uncertainty in longer-term forecasts.

Simulation platforms like the Building Optimization Testing Framework (BOPTEST) (Blum et al. 2021) provide environments for testing advanced building control strategies. Currently, BOPTEST only provides deterministic weather forecasts to users for testing their control strategies. Incorporating stochastic weather models into such frameworks is essential for assessing and enhancing the robustness of control strategies under realistic uncertainty conditions.

Despite these developments, there is still a lack of practical tools for effectively simulating weather forecast uncertainties in building applications. Many existing models assume that residuals are normally distributed, which may be the case for temperature (Petersen and Bundgaard 2014), but which may not accurately reflect the actual distribution of forecast errors—particularly for variables like solar irradiance that are peaked, with narrower shoulders and longer tails, indicating non-normal behavior (Gayathry et al. 2024). Some models also inadequately represent extreme weather events due to practices like excluding large error values (Grant and Gehbauer 2022), potentially underestimating their impact on control strategies. The lack of validation against real forecast data in (Oldewurtel et al. 2012) undermines the reliability of the proposed method. When forecasts deviate from normality, simply comparing the mean and standard deviation may be insufficient; therefore, more comprehensive evaluation methods are necessary to compare the similarity between synthetic and historical forecasts effectively.

1.2 Scope and objectives

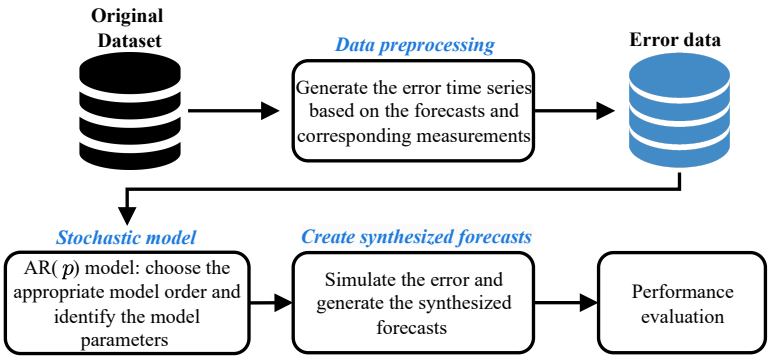
The objective of this study is twofold. Firstly, we aim to quantify the weather forecast uncertainty using datasets collected from four different locations, representing the common situation where an advanced control developer relies on weather forecasts provided by a service for a given location. We are particularly interested in investigating the uncertainty associated with ambient temperature and GHI forecasts. These inputs are critical for building thermal load and prediction-

based building control strategies. Secondly, we aim to develop a stochastic model to simulate weather forecast uncertainties accurately. This model will be used as an uncertain weather forecast generator in the BOPTEST framework to facilitate the development of RMPC and SMPC strategies. BOPTEST is a virtual building simulation framework that provides multiple physics-based building models with predefined baseline controllers, testing scenarios, and key performance indicator calculations, allowing control developers to develop and compare their building control strategies.

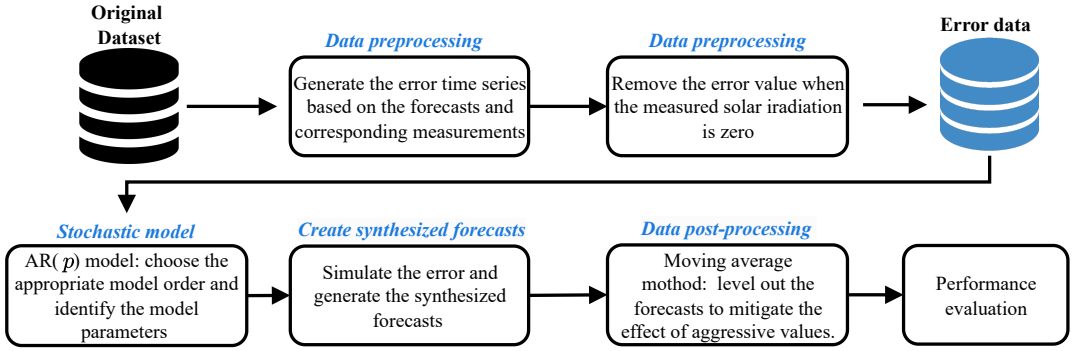
The remainder of this paper is organized as follows: In Section 2, we introduce the datasets and associated uncertainty quantification, the steps involved in generating synthetic forecasts, and the evaluation metrics to be used in this study. Next, we present the results, including the ambient temperature (Section 3.1) and the GHI (Section 3.2). We then discuss our contribution and limitations in Section 4, followed by a conclusion in Section 5.

2 Method

This section outlines the specific procedures for quantifying and simulating the weather forecast uncertainty, focusing on ambient temperature and GHI. The corresponding flow diagrams are illustrated in Fig.1. Notably, due to the unique characteristics of GHI forecasts, additional data preprocessing and post-processing steps are required, as depicted in Fig.1 (b). By following these procedures, we can effectively model the weather forecast uncertainties.



(a) Ambient temperature



(b) Global solar irradiation

Figure 1: Flowchart of the proposed method.

2.1 Datasets

This study analyzed historical weather forecast data alongside the corresponding measured weather data from four cities. The data ranges include: 1) Berkeley, USA: January 15, 2019, to December 31, 2019; 2) Leuven, Belgium: January 1, 2019, to December 31, 2019; 3)Berlin, Germany: October 1, 2020, to July 27, 2021, 4) Oslo, Norway: February 21, 2018, to September 17, 2018. Note that the forecasting and observation services are the same in the Berlin and Oslo cases. In contrast, in the Berkeley and Leuven cases, the forecasting service is separate from the observation service, which are weather stations located at building sites. These cases apply

to a practical situation confronting an advanced control developer. The datasets were obtained as follows:

- For Berkeley, the weather forecasts were obtained from the DarkSky API (Apple Inc) and stored for a field demonstration of MPC (Blum et al. 2022). During the same demonstration, the measured data were collected from a weather station on LBNL's campus and accessed through (The University of Utah). The DarkSky API provides ambient temperature and cloud cover forecasts but not GHI forecasts. Therefore, the GHI forecasts are generated using the k-Nearest Neighbors (kNN) machine learning algorithm, trained using 30 days of recent past measurements of GHI from the weather station in combination with past 1-step cloud cover, ambient temperature forecasts and predicted clear-sky irradiance. More details can be found in (Blum et al. 2022). In the MPC demonstration, the training process was run each time weather forecasts were collected, approximately once per hour, along with a validation of the model on the most recent previous day's data, to confirm the model's validity to be used for the MPC. To give a broad idea of error introduced by this kNN model, during the period of data used for this current paper's study and across all validation data sets, the mean Root Mean Square Error (RMSE) of the method was 175 W/m^2, and the standard deviation of the RMSE was 93.4 W/m^2.

- For Berlin, the weather forecasts were obtained from the MOSMIX model from Germany's Meteorological Services Deutscher Wetterdienst (DWD) (Deutscher Wetterdienst a). The DWD also provided the measured data via the OpenData server (Deutscher Wetterdienst b).

- For Leuven, the weather forecasts were purchased from the OpenWeather service (OpenWeather). The measured data were collected at thpdme Vliet Building in Leuven, operated

by the Building Physics Section of the KU Leuven.

- For Oslo, the weather forecasts were obtained from the Norwegian Meteorological Institute at (Norwegian Meteorological Institute). The measured data were collected from its API (Meteorologisk institutt).

Both ambient temperature and GHI are analyzed at the first three locations and just the ambient temperature for Oslo due to data availability. All forecasts offer a prediction horizon of 48 hours at 1-hour time intervals and all measurements are recorded at 1-hour time intervals. The Berkeley, Berlin and Oslo forecasts are updated hourly, while the Leuven forecasts are updated every 6 hours. For Leuven's GHI dataset, the first and second-hour prediction time intervals (i.e., 1- and 2-hour-ahead forecast) are excluded due to missing data, resulting in the prediction horizon extending from the 3rd to the 48th hour. To prepare the data for model development and validation, we split each city's dataset chronologically into two subsets: a training set comprising the first 70% of the data used to identify the model parameters, and a validation set consisting of the remaining 30% of the data used to validate the models.

2.2 Preprocessing

A forecast's lead time τ represents τ steps ahead forecast. For example, $\tau = 1$ means one-step ahead forecasts. The error for the lead time τth of the forecast at timestep k, denoted by $e(\tau|k)$, is calculated as the difference between the lead time τth of forecast at timestep k and the corresponding measurement at timestep $(k + \tau)$. The equation is defined as follows:

$$e(\tau|k) = \widetilde{X}(\tau|k) - X(k + \tau) \qquad \tau = 1, \ldots, 48 \tag{1}$$

where $\widetilde{X}(\tau|k)$ denotes the predicted value for the lead time τ at timestep k, $X(k + \tau)$ denotes the corresponding measured value.

Figs. 2 and 3 display the distribution of forecast errors for temperature and GHI at one lead time, respectively. The temperature variable exhibits a normal distribution of forecast errors across all locations and all lead times, albeit with a mean that differs slightly from zero in most cases. The histograms of the GHI forecast errors exhibit a sharp peak and a high concentration near zero, which suggests that a significant proportion of the error values are equal to zero. Further analysis reveals that the cause of this phenomenon is the high accuracy of the predictions during nighttime when the measured GHI is zero. As shown in Fig. 4, which provides an example of the comparison between measured and forecast GHI data, the predictions during nighttime were consistently accurate. Given this situation, it has been determined no uncertainty will be added in cases where the measured GHI is equivalent to 0. In practice, we drop the errors in every prediction horizon when the measured GHI equals 0 and get new error data; the new error histogram of the GHI is shown in Fig. 5. We can see that the spikes in the histogram become flatter as we remove a large number of error values equal to 0, which also occur in all lead times.

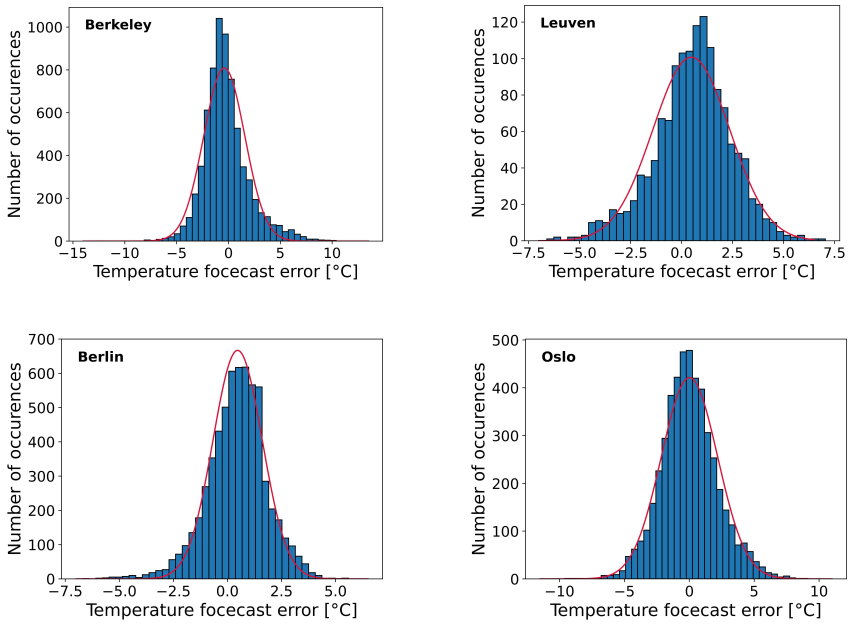


Figure 2: Histograms of the temperature forecast error distribution with a normal distribution fit (red) at lead time 13 of the prediction horizon.

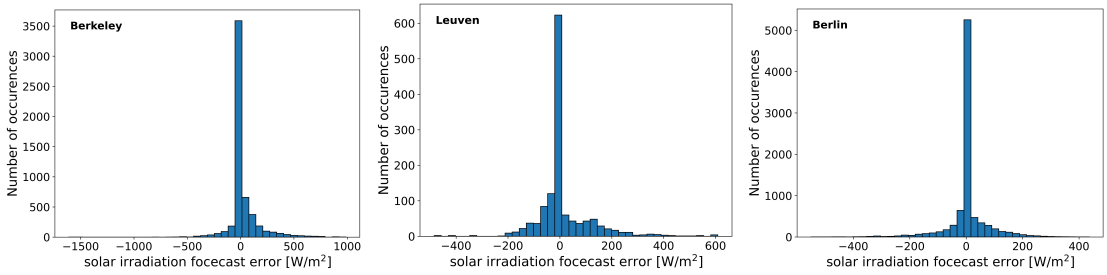


Figure 3: Histograms of the GHI forecast error distribution at lead time 13 of the prediction horizon.

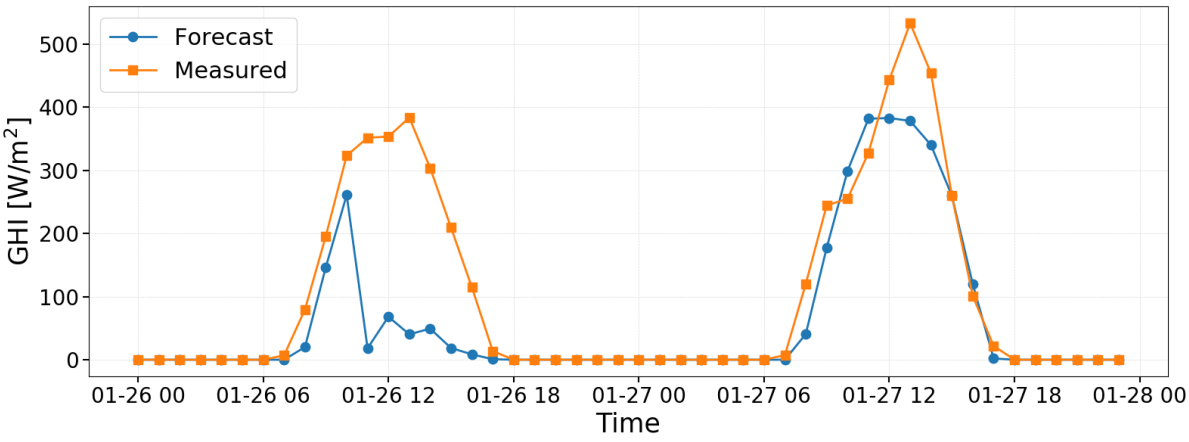


Figure 4: Comparison between the forecast and measured GHI data in Berkeley at 2019-01-25 23:00:00.

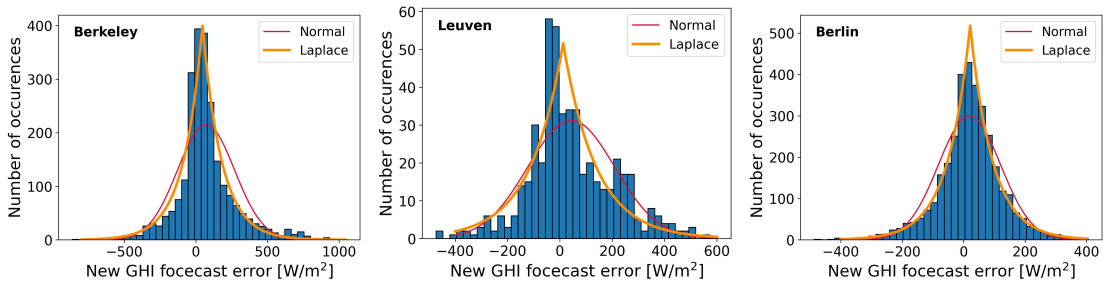


Figure 5: Histogram of the GHI forecast error distribution at lead time 13, exemplifying the distribution shape, which is similar across all lead times. Errors when the measured GHI is zero have been removed. The red and orange curves represent the Normal and Laplace distribution fits to the histograms, respectively.

To statistically characterize the distribution of the new GHI forecast errors, we considered Normal, Laplace, Gamma, and Logistic distributions as potential fits for the GHI forecast errors. To assess the goodness-of-fit for each distribution, we employed the statistical metric the Akaike Information Criterion (AIC) (Akaike 1974).

The AIC measures the relative quality of statistical models for a given data set. It is defined as:

$$\mathrm{AIC} = 2k - 2\ln(L), \tag{2}$$

where k is the number of parameters in the model, and L is the maximum value of the likelihood function for the model. Lower AIC values indicate a better fit to the data, balancing model complexity and goodness of fit. Our results indicate that the Laplace distribution consistently fits the GHI forecast error data better than the Normal and other distributions with lower AIC values. This finding aligns with the visual observations from the histograms of the new GHI forecast errors, which closely resemble the sharp peak and heavy tails characteristic of the Laplace distribution.

2.3 Stochastic model

This study applies the time series AR model to simulate the forecast error of the 48-hour prediction horizons. The AR(p) model can be represented as

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t \tag{3}$$

where y_t is the forecast error at lead time t, ϕ_i are the model's parameters, ε_t denotes the residual that captures the disturbances that affect the system, and p specifies the order of the AR model.

To apply the AR model, determining the appropriate order p is crucial, as it influences the model's ability to capture the underlying dynamics of the time series. Several methods exist for selecting the optimal order of an AR model, including the partial autocorrelation function (PACF) introduced by Box and Jenkins and information criteria like the AIC and Bayesian Information Criterion (BIC) (Chen et al. 2013). In this study, we utilize PACF analysis to identify the model order for our time series data. The PACF provides insights into past values' direct effect on the series' current value, excluding the indirect effects mediated through intermediate lags (Satrio et al. 2021). Specifically, the partial autocorrelation at lag h, denoted as PACF(h), measures the correlation between y_t and y_{t-h} after removing the influence of the intervening

lags $y_{t-1}, \ldots, y_{t-h+1}$. It is defined as:

$$\text{PACF}(h) = \frac{Cov\left(y_t, y_{t-h} \mid y_{t-1}, \ldots, y_{t-h+1}\right)}{\sqrt{Var\left(y_t \mid y_{t-1}, \ldots, y_{t-h+1}\right) Var\left(y_{t-h} \mid y_{t-1}, \ldots, y_{t-h+1}\right)}} \tag{4}$$

Analyzing the PACF is useful for identifying the order of an AR model because, for an AR(p) process, the PACF theoretically exhibits a sharp cutoff after lag p, which means that the partial autocorrelations are significantly different from zero for lags up to p and not significantly different from zero for lags greater than p. The property allows us to visually inspect the PACF plot to determine the appropriate order of the AR model.

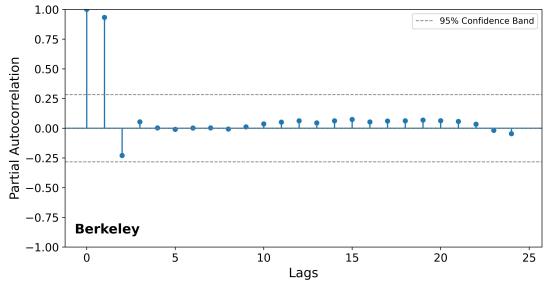
Fig. 6 presents the PACF plots for the temperature and GHI forecast errors. In Fig. 6 (a), the PACF of the temperature variable shows a significant spike at lag 1 that exceeds the 95% confidence band, indicating a strong partial autocorrelation at this lag. The confidence bands are calculated as $\pm 1.96/\sqrt{N}$, where N is the sample size, providing a 95% Beyond lag 1, the partial autocorrelations drop to near zero and remain within the confidence bands, suggesting that higher-order lags do not contribute significantly to the model. Similarly, Fig. 6 (b) depicts the PACF for the GHI variable, which also displays a significant partial autocorrelation at lag 1 and values within the 95% confidence band at higher lags. Based on these observations, we conclude that the time series of forecast errors for temperature and GHI can be adequately modeled using an AR(1) process. The significant partial autocorrelation at lag 1 indicates that the immediate past error directly influences the current forecast error, while the insignificance of higher lags (as they fall within the 95% confidence band) implies that further past errors do not have a direct impact.

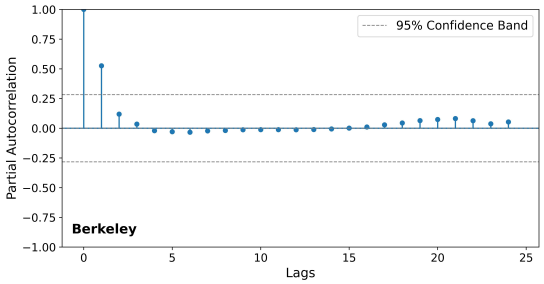
Therefore, the AR(1) model used in this study is defined as follows:

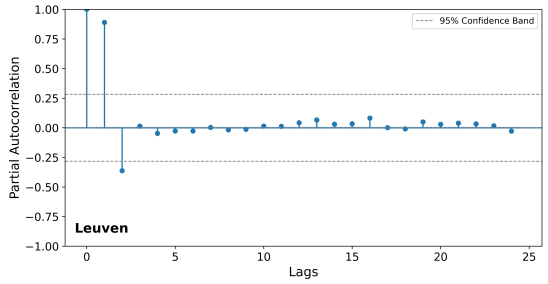
$$e\left(\tau + 1 \mid k\right) = \phi e\left(\tau \mid k\right) + \varepsilon \tag{5}$$

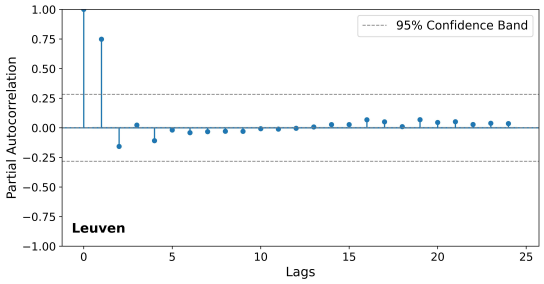
where e is the time series of forecast error and ε is the model residual. The known condition k means to predict a specific time k and τ represents the lead time, ranging from 1 to 47. We assume

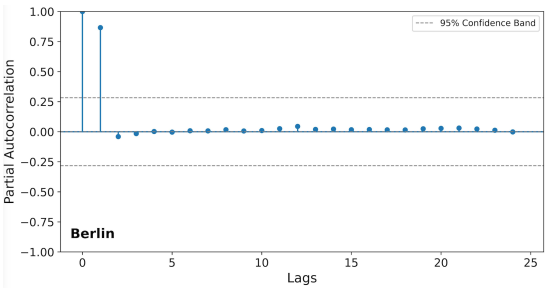
that the model of temperature variable is subject to normally distributed residual, denoted as $\varepsilon \sim \mathcal{N}(\mu, \sigma)$, where μ and σ represent the mean and standard deviation of the distribution. The residual associated with the GHI variable is modeled using a Laplace distribution, represented as $\varepsilon \sim Laplace(u, b)$, with u and b denoting the location and scale parameters. We employ the least-squares approach to estimate the parameters ϕ and ε using the AR(1) model.

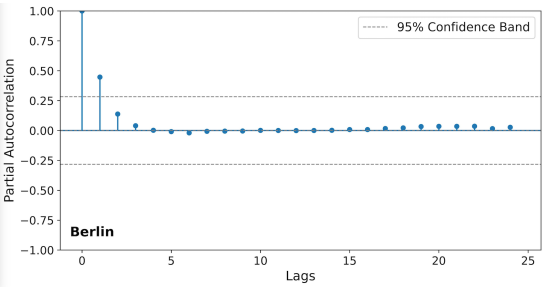


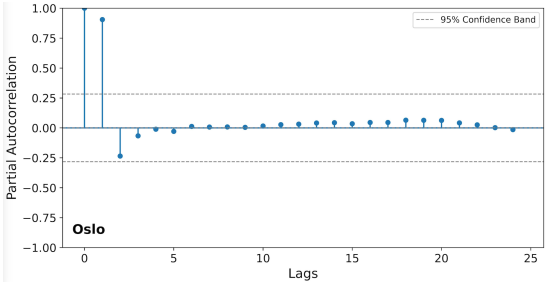












(a) PACF for ambient temperature

(b) PACF for GHI

Figure 6: Partial autocorrelation function for forecast temperature and GHI errors with 95% confidence interval.

2.4 Generate synthetic forecasts

To generate synthetic forecasts that show stochastic properties similar to historical forecasts, we utilize the propagation property of the AR model and add the synthetic forecast error to the measured data.

In the context of time series propagation, for ambient temperature, we obtain the initial error $\widehat{e}(\tau = 1|k)$ by sampling from the distribution of the historical error $e(\tau = 1)$, which is characterized by $e(\tau = 1) \sim \mathcal{N}(\mu_1, \sigma_1)$. Similarly, for GHI, $\widehat{e}(\tau = 1|k)$ is sampled from $e(\tau = 1)$ with $e(\tau = 1) \sim Laplace(u_1, b_1)$. Then we recursively substitute $\widehat{e}(\tau|k)$ into Equation (5) to obtain $\widehat{e}(\cdot|k)$. It is worth noting that for GHI, instances where the corresponding measurement $X(k + \tau)$ is zero lead to a resetting of $\widehat{e}(\tau|k)$ to zero.

After generating the synthetic forecast errors, we calculate the synthetic forecasts $\widehat{X}(\tau|k)$ by adding the synthetic forecast error to the corresponding measurement:

$$\widehat{X}(\tau|k) = X(k + \tau) + \widehat{e}(\tau|k) \tag{6}$$

Furthermore, GHI forecasts can exhibit large prediction errors with a relatively high standard deviation, as depicted in Fig. 5. During the synthetic forecast generation process, the stochastic model may produce excessively large errors in rapid succession, leading to unrealistic high-frequency fluctuations in the synthetic GHI forecasts, especially when the measured GHI is small. The fluctuations are artifacts of the modeling process and are characterized by abrupt alternations between large overestimations and underestimations in consecutive time steps, which are not commonly observed in actual forecast error patterns. To mitigate the impact of these unrealistic fluctuations, we apply a moving average filter with a window size of 3 samples to the synthetic GHI forecasts. The moving average filter is well-known for its ability to smooth out high-frequency variations in time series data (Martins et al. 2019), making it a suitable tool for refining synthetic GHI forecasts. The post-processing step helps the synthetic forecasts maintain

realistic temporal characteristics and avoid introducing artificial noise, thereby ensuring that the control performance testing outcomes represent real operating conditions. The moving average filter does not eliminate realistic large forecast errors due to genuine weather events like sudden cloud cover. The filter simply reduces unrealistic erratic behavior from the stochastic modeling process and is not reflective of actual weather dynamics.

2.5 Evaluation metrics

To evaluate the stochastic properties of the generated and historical (the ground-truth) forecasts, the Jensen-Shannon divergence (JSD) (Lin 1991) is used to quantify the similarity between the probability distribution of actual and synthetic errors. JSD is a statistical measure of the similarity between two probability distributions and is defined as:

$$\mathrm{JSD}(A\|B) = \frac{1}{2}\,\mathrm{KLD}\left(A\|\frac{1}{2}\,(A+B)\right) + \frac{1}{2}\,\mathrm{KLD}\left(B\|\frac{1}{2}\,(A+B)\right) \tag{7}$$

where $\mathrm{KLD}(\cdot)$ denotes the Kullback-Leibler divergence (KLD). JSD is symmetric and bounded, which means $\mathrm{JSD}(A\|B) = \mathrm{JSD}(B\|A)$ and ranges between [0, 1]. JSD=0 indicates that probability distribution A is identical to B. A higher JSD indicates a greater dissimilarity between the two probability distributions. JSD $\geq$ 0.4 is generally considered to indicate significant disparities between the two distributions (Netzel and Stepinski 2015) and JSD $\leq$ 0.1 indicates the two distributions being closely similar to each other (Baasch et al. 2021).

We choose JSD over KLD because JSD is symmetric (i.e., not affected by the distribution order), whereas KLD is not symmetric. The symmetry ensures that the divergence measure treats both distributions equally, providing an unbiased assessment of their similarity. Additionally, JSD is always finite and bounded between 0 and 1, enhancing interpretability and making meaningful comparisons across different scenarios. In contrast, KLD can be unbounded and produce infinite values if the probability distributions have zero probabilities in places where the other distribution does not, making JSD a more robust choice for our analysis.

3 Result

In this section, we utilize the statistical parameter JSD to analyze the similarity between the synthetic and historical forecast errors and then create the synthetic forecast.

3.1 Ambient temperature

Table 1 presents the AR(1) parameters for four cities: Berkeley, Leuven, Berlin and Oslo. The similarity between the historical error and the synthetically generated error in the ambient temperature is evaluated by calculating the JSD. A synthetic error is generated for all points with the historical forecast collected for each location. This evaluation metric represents the four analyzed locations in Fig. 7.

Table 1: Summary of AR(1) Parameters for Ambient Temperature in Four Cities. ϕ represents the autoregressive coefficient, μ and σ are the mean and standard deviation of the residuals ε, where $\varepsilon \sim \mathcal{N}(\mu, \sigma)$.

City	**Berkeley**	**Leuven**	**Berlin**	**Oslo**
$\boldsymbol{\phi}$	0.939	0.918	0.903	0.924
$\boldsymbol{\mu}$	0.008	0.045	0.047	0.017
$\boldsymbol{\sigma}$	0.711	0.771	0.575	0.850

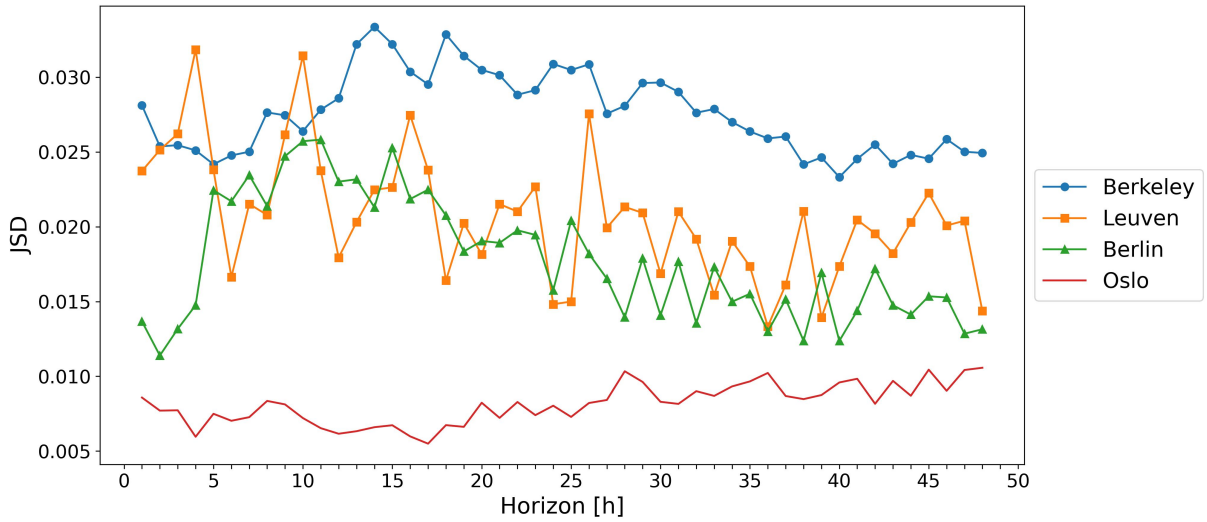


Figure 7: JSD analysis between historical and synthetic ambient temperature forecast errors across prediction horizons for four cities.

The JSD shows a very high similarity in the distributions for Oslo. Meanwhile, the similarity is notably high for the Leuven and Berlin locations and slightly lower for the Berkeley locations. The ambient temperature error model is assessed by analyzing additional statistical parameters. As the ambient temperature error has a Gaussian distribution, the evolution of the mean error and the standard deviation through the prediction horizon is also analyzed in Fig. 8 and Fig. 9, respectively. The location of Leuven has been left out of the figures for clarity of the results, as the predictions are updated every 6 hours in Leuven, and the results cannot be analyzed hourly in the same way as the rest of the locations. All the available historical forecasts are used to calculate the statistical indicators for the real error, and the emulated forecasts are generated for all the points with historical forecasts available for each location.

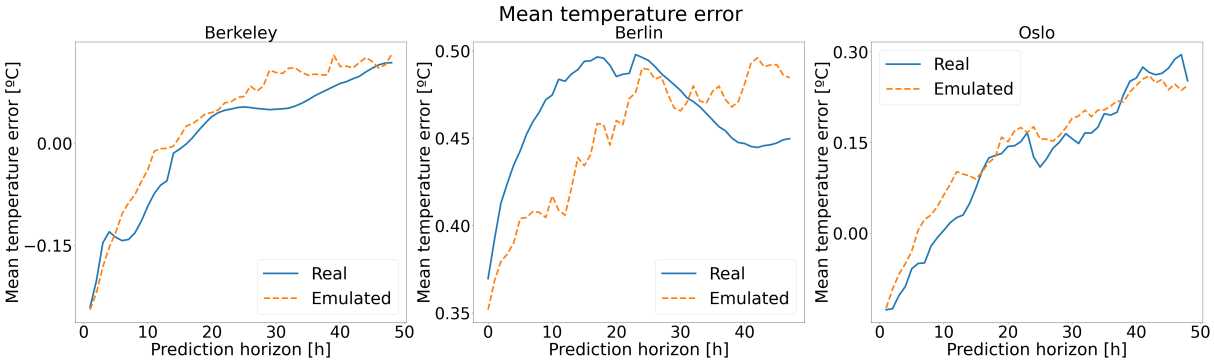


Figure 8: Mean temperature error evolution through the prediction horizon for the real error and the synthetically generated error for Berkeley, Berlin and Oslo.

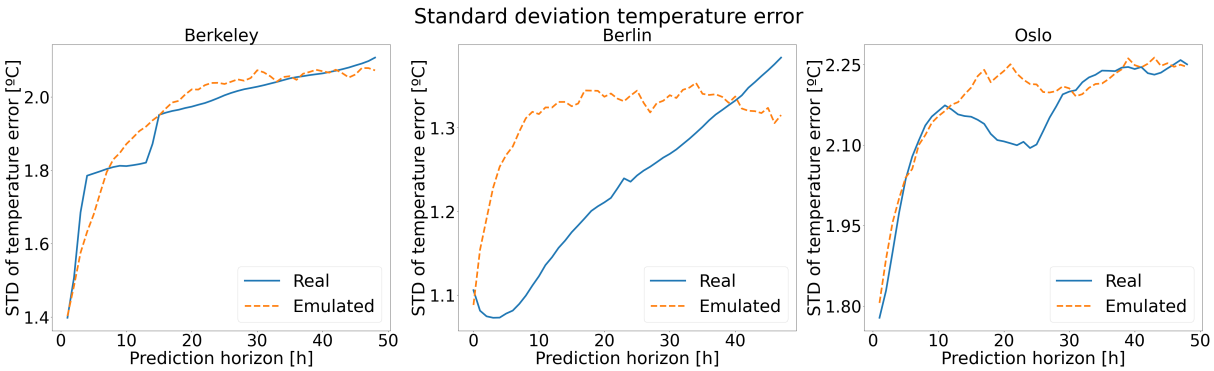


Figure 9: Standard deviation temperature error evolution through the prediction horizon for the real error and the synthetically generated error for Berkeley, Berlin and Oslo.

The mean and standard deviation of the synthetic errors evolve along the prediction horizon, following the same tendency as the real errors at the Berkeley and Oslo locations. In the case of Berlin, the tendency differs slightly. Fig. 10 shows an example of the real ambient temperature for a specific time period, the historic forecast for the temperature, and the synthetically generated forecast with the proposed model.

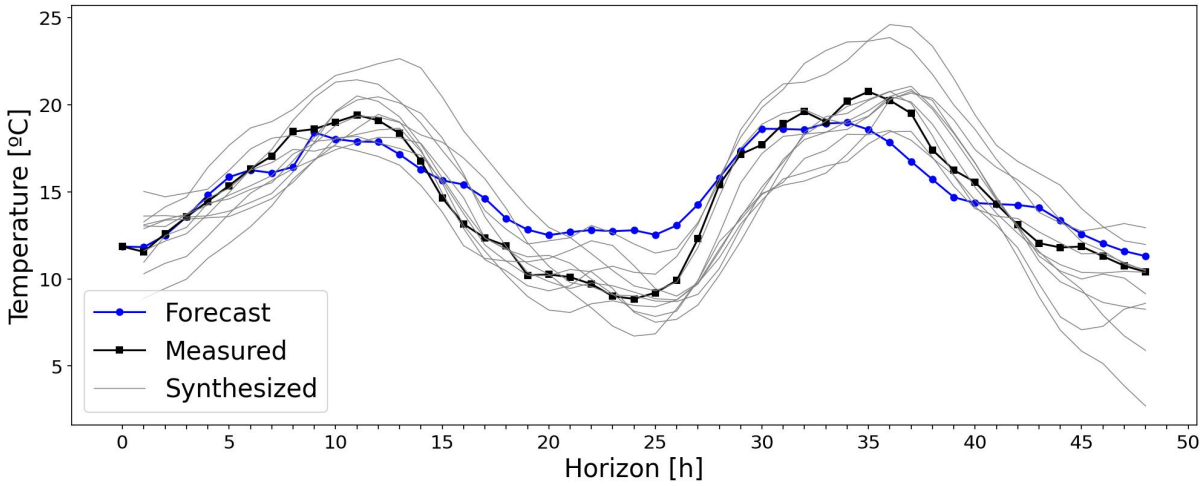


Figure 10: Comparison of actual ambient temperature forecast with 10 synthetic ambient temperature forecasts at 2018-02-21 00:00:00 from Oslo dataset.

3.2 GHI

Table 2 presents the AR(1) parameters for three cities: Berkeley, Leuven, and Berlin. Fig. 11 illustrates a comparison between Berlin's historical and synthetic forecast errors across selected prediction horizons of 1, 12, 24, and 48 hours. The data is generated using the same method as in the ambient temperature section, which utilizes historical forecasts and calculated AR(1) parameters. The results reveal a high degree of similarity between the historical and synthetic forecast errors, underscoring the effectiveness of this method. The datasets for Berkeley and Leuven exhibit a similar behavior. To further quantify this similarity, we use the statistical parameter JSD. Fig. 12 presents the results of JSD analysis for the three cities. As we can see from the figure, the JSD of GHI across prediction horizons is relatively small for all three cities, indicating that the landscapes are similar overall but with some differences. Fig. 13 compares an example of 10 synthetic GHI forecasts with the actual forecast from the Berlin dataset. Synthetic forecasts capture the basic shape of the profile measurement profile, exhibit a range of variability around it, and cover the actual forecasts.

Table 2: Summary of AR(1) Parameters for GHI in Three Cities. The autoregressive coefficient is denoted by ϕ. Residuals, ε, are characterized by a Laplace distribution with location u and scale b: $\varepsilon \sim \text{Laplace}(u, b)$.

City	Berkeley	Leuven	Berlin
$\boldsymbol{\phi}$	0.67	0.63	0.62
$\boldsymbol{u}$	10.63	4.44	1.86
$\boldsymbol{b}$	87.44	91.97	45.64

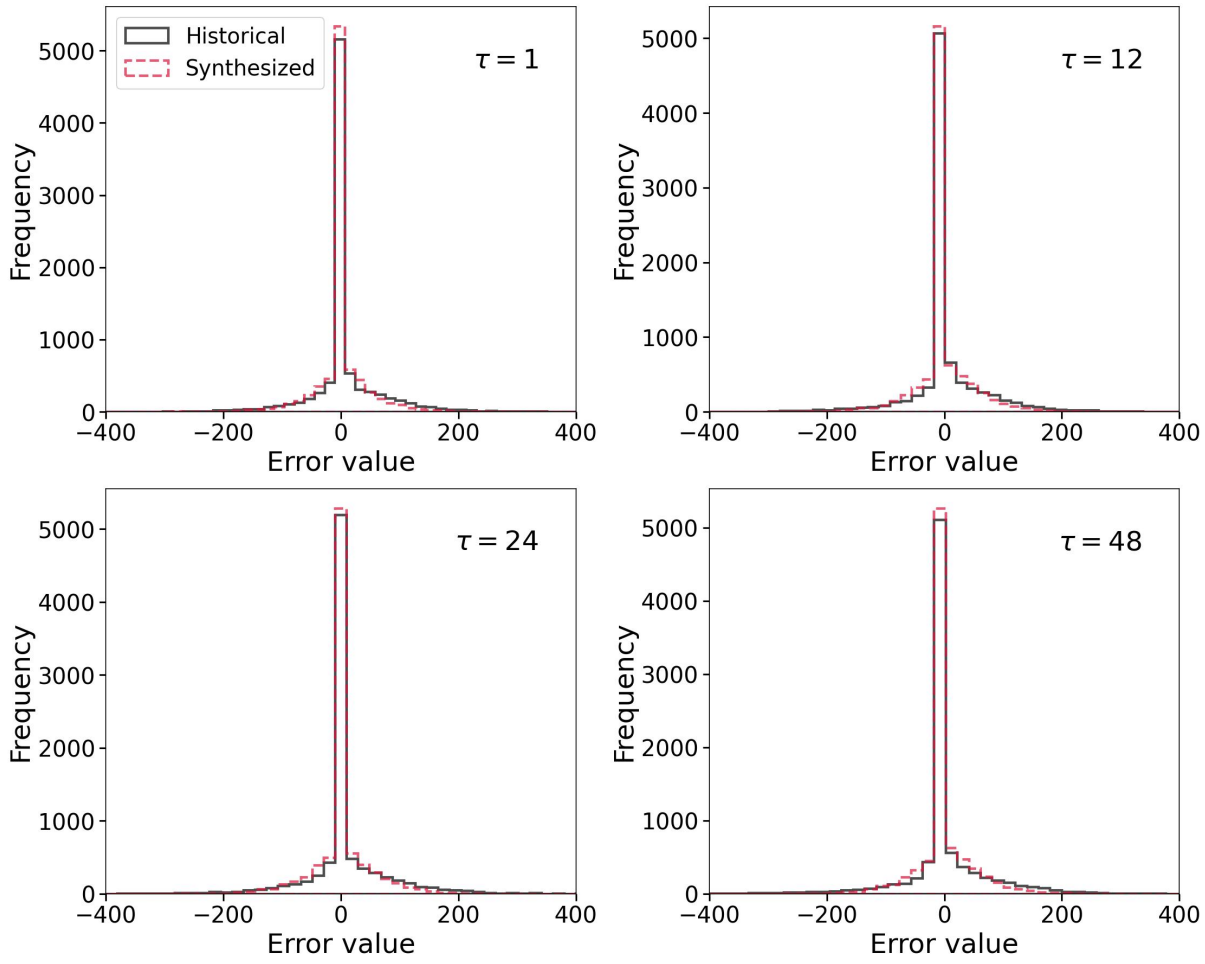


Figure 11: Comparison of forecast error distributions for historical and synthetic forecasts at selected lead times ($\tau = 1, 12, 24, 48$ hours) for the Berlin data. Each subplot represents a different lead time τ.

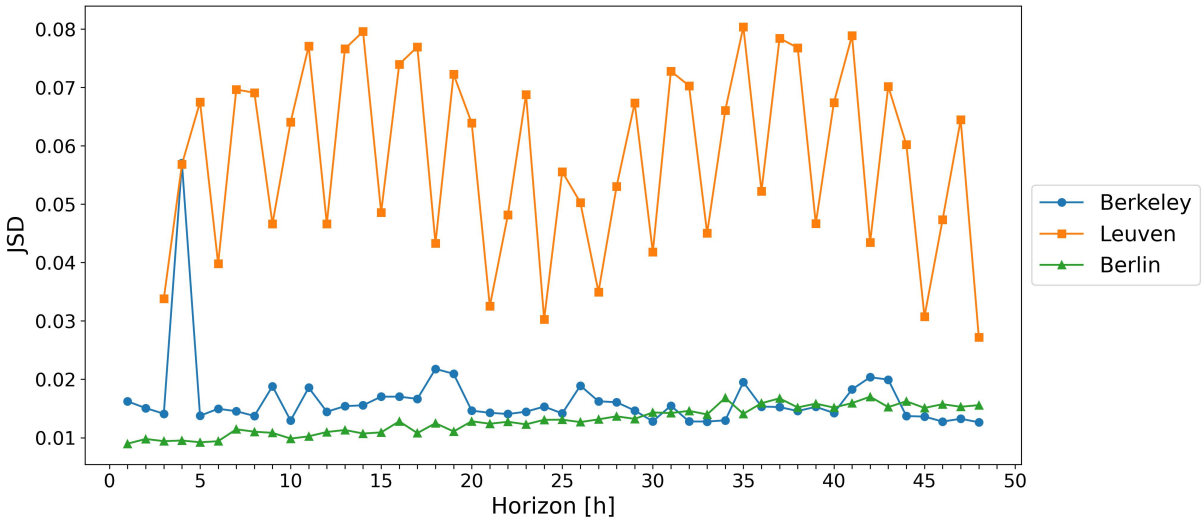


Figure 12: JSD analysis between historical and synthetic GHI forecast errors across the prediction horizon for three cities.

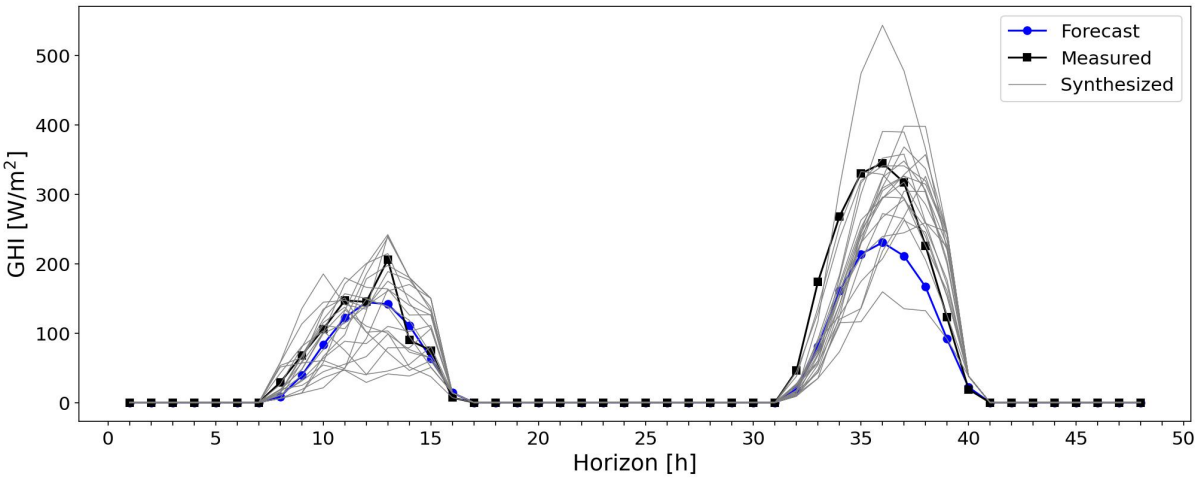


Figure 13: Comparison of actual GHI forecast with ten synthetic GHI forecasts at 2020-04-13 00:00:00 from Berlin dataset.

4 Discussion

4.1 Contribution

In this work, we model the weather forecast uncertainty in scenarios where weather forecasts are provided by external services to advanced control developers. We focus on ambient temperature and GHI, two critical boundary conditions for a building's thermal load. The studied datasets comprise data collected from different locations with varying climate regions, including cases where the forecasts and measurements are provided by the same service and cases where they are not. We found that the weather forecasts are biased, meaning the mean of the forecast residuals is non-zero. Additionally, as expected, the standard deviations of the forecast residuals increase with the prediction horizon.

Our primary contribution involves applying the AR(1) model to emulate weather forecast data. We discovered that the AR(1) model effectively represents the tendencies of these statistical parameters. The residuals of the temperature forecast follow a normal distribution, while those of the GHI forecast do not. The residuals of the GHI forecast are centered around zero because the GHI forecasts at nighttime are always zero, with no forecast errors. After excluding the GHI residuals during nighttime, we found that the remaining forecast residuals follow a Laplace distribution.

We also propose a workflow for modeling and generating synthetic weather forecast data using the AR(1) model. The temperature forecasting process begins with pre-processing the raw data by comparing historical forecasts and measured data to obtain the forecast residuals. Then, we identify the parameters of the AR(1) model using the least-squares approach and generate the synthetic forecast data using the identified AR(1) model. For the solar irradiation forecasts, additional pre-processing and post-processing steps are necessary. In the pre-processing stage, we remove the nighttime data before fitting the Laplace model. In the post-processing stage, we

apply a moving average filter to the generated forecast data since the standard deviation of the forecast error for solar irradiation is larger than that for ambient temperature forecasts.

Finally, we introduce Jensen-Shannon divergence as a metric to quantify the statistical similarity between the generated and real weather forecast errors. A smaller Jensen-Shannon divergence value indicates a higher similarity between the distributions of the forecast residuals. We believe the Jensen-Shannon divergence is a more comprehensive and suitable statistical measure because it compares the entire distribution rather than just two statistics, such as the mean and standard deviation.

Though we focused on ambient temperature and GHI forecasts in this study, the proposed approach has the potential to be applied to other weather forecast variables. Moreover, the findings of this study can be used to develop building controllers that consider the presence of uncertain weather forecasts using techniques such as stochastic Model Predictive Control (MPC) or robust MPC.

4.2 Limitation and future work

Our study has several limitations. First, our empirical analysis is based on the data collected from four cities. Whether these four cities are adequately representative remains unknown. It is necessary to validate the generalizability of the method as more site datasets are available. Second, we only consider the self-dependence in our auto-regressive model, ignoring other exogenous uncertainty sources, such as cloudiness information for GHI variables. There is a trade-off between model accuracy and complexity. We hope our model can be as simple as possible until it cannot be simpler. The other uncertain sources (such as cloudiness) are less accessible. Therefore, we did not include them in our models. In the case of the ambient temperature, in some locations, the proposed model presents a higher error in the first hour of the prediction horizon than in the rest of the time points due to a worse fitting to the proposed normal

distribution. More location analysis will be needed to investigate this observation. Finally, this work did not consider any potential correlation of forecast errors among them, for example, where a forecast that under-predicts global horizontal irradiation may be more likely to also under-predict ambient air temperature for the same horizon. Instead, the synthetic forecasts are generated individually for each variable.

In the next step, we propose developing an emulator of weather forecast uncertainty, which will be integrated into the BOPTEST forecast API, which currently provides test controllers with deterministic forecasts of disturbances in a given test case. By incorporating our weather forecast uncertainty emulator into BOPTEST, building controllers can be tested and optimized in the simulated environment that considers the uncertainty of weather forecasts.

Additionally, future studies could extend our approach to model the uncertainties in diffuse and direct solar irradiation components separately. Incorporating these components would enhance the applicability of our model for applications that specifically rely on direct or diffuse solar irradiation data.

5 Conclusion

Although with the improvement of weather forecast technology, the uncertainty of weather forecast is unavoidable. Quantifying and modeling weather forecast uncertainty is important because it is necessary to consider weather forecast uncertainty when developing and comparing advanced building controllers. To address this problem, we present a workflow to model the weather forecast uncertainty using the AR model. We are especially interested in the ambient temperature and GHI forecast, two essential boundary conditions for building thermal dynamics. Four ambient temperatures and three GHI datasets from different locations were analyzed. We applied the AR(1) model to simulate the weather forecast, a normal distribution to fit the temperature forecast residuals and a Laplace distribution to fit the GHI forecast residuals. We

used the JSD to evaluate the similarity between synthetic and actual forecast errors. Our results show that the average JSD temperature over the horizons of Berkeley, Leuven, Berlin and Oslo are approximately 0.027, 0.021, 0.018 and 0.008, respectively, and the average JSD of GHI over the horizons of Berkeley, Leuven, and Berlin are approximately 0.016, 0.058 and 0.013, respectively. The low values of JSD indicate a high similarity between the distributions of generated and actual forecast errors. We also found the forecast errors vary a lot in different locations. We then generate several synthetic forecasts and compare them to the original forecasts. Our results show that the synthetic predictions cover the same area as the original forecasts, showing that our approach can generate synthetic weather forecasts that are statistically similar to the actual ones.

Acknowledgments

The authors acknowledge Harald Taxt Walnum, Laura Maier, Javier Arroyo and Ettore Zanetti for their support in providing data used for this study. The authors would also like to thank the following institutions for making this data available: Lawrence Berkeley National Laboratory, Building Physics Section of the KU Leuven, Deutscher Wetterdienstand Norwegian Meteorological Institute. The authors would like to thank Tekniker research and technology centre for supporting and contributing to this study during Project-1.

This work emerged from IBPSA Project 1 and Project 2, international projects conducted under the umbrella of the International Building Performance Simulation Association (IBPSA). Project 1 developed and demonstrated a BIM/GIS and Modelica Framework for building and community energy system design and operation. Project 2 will develop and demonstrate the Building Optimization Testing Framework (BOPTEST) for the testing, evaluating, and benchmarking of building and community energy system controls.

Funding

Laura's work has been partially funded by the Next Generation EU program from the European Union. The work described in this paper was substantially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. C6003-22Y), and by the Hong Kong University Grants Committee (UGC) Grant 26209323. This work was partly supported by the Project of Hetao ShenzhenHong Kong Science and Technology Innovation Cooperation Zone (HZQB-KCZYB-2020083). This research was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Building Technologies of the U.S. Department of Energy, under Contract No. DE-AC02-05CH11231. Weather forecast and measured data collection at the Berkeley site was supported by the U.S.-China Clean Energy Research Center for Building Energy Efficiency (CERC-BEE).

Data Availability Statement

The data supporting this study's findings are available from the corresponding author upon reasonable request.

Declaration of interests

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

Hirotugu Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974. doi: 10.1109/TAC.1974.1100705.

Zhuoer An, Xinghua Liu, Gaoxi Xiao, Guangyu Song, and Peng Wang. Tube-based mpc strategy

for load frequency control of multi-area interconnected power system with hess. *Journal of Energy Storage*, 99:113340, 2024. doi: 10.1016/j.est.2023.113340.

Apple Inc. Dark sky api. URL https://darksky.net/dev. Last accessed 2022.

Gaby Baasch, Guillaume Rousseau, and Ralph Evins. A conditional generative adversarial network for energy use in multiple buildings using scarce data. *Energy and AI*, 5:100087, 2021. doi: 10.1016/j.egyai.2021.100087.

Peder Bacher, Henrik Madsen, and Henrik Aalborg Nielsen. Online short-term solar power forecasting. *Solar Energy*, 83(10):1772–1783, 2009. doi: 10.1016/j.solener.2009.05.016.

Gianni Bianchini, Marco Casini, Daniele Pepe, Antonio Vicino, and Giovanni Gino Zanvettor. An integrated model predictive control approach for optimal hvac and energy storage operation in large-scale buildings. *Applied Energy*, 240:327–340, 2019. doi: 10.1016/j.apenergy.2019.01.187.

David Blum, Javier Arroyo, Sen Huang, Ján Drgoňa, Filip Jorissen, Harald Taxt Walnum, Yan Chen, Kyle Benne, Draguna Vrabie, Michael Wetter, et al. Building optimization testing framework (boptest) for simulation-based benchmarking of control strategies in buildings. *Journal of Building Performance Simulation*, 14(5):586–610, 2021. doi: 10.1080/19401493.2021.1986574.

David Blum, Zhe Wang, Chris Weyandt, Donghun Kim, Michael Wetter, Tianzhen Hong, and Mary Ann Piette. Field demonstration and implementation analysis of model predictive control in an office hvac system. *Applied Energy*, 318:119104, 2022. doi: 10.1016/j.apenergy.2022.119104.

Shobhit Chaturvedi and Elangovan Rajasekar. Application of a probabilistic lhs-pawn approach

to assess building cooling energy demand uncertainties. *Building Simulation*, 15(3):373–387, 2022. doi: 10.1007/s12273-021-0815-6.

Shuhao Chen, Wanli Min, and Rong Chen. Model identification for time series with dependent innovations. *Statistica Sinica*, 23(2):873–899, 2013. doi: 10.5705/ss.2010.219.

Deutscher Wetterdienst. Statistically optimised weather forecasts (mosmix), a. URL https://www.dwd.de/EN/research/weatherforecasting/met_applications/nwp_applications/mosmix_application_node.html. Last accessed 17 April 2023.

Deutscher Wetterdienst. Opendata, dwd, b. URL https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/. Last accessed 17 April 2023.

Bing Dong, Reisa Widjaja, Wenbo Wu, and Zhi Zhou. Review of onsite temperature and solar forecasting models to enable better building design and operations. *Building Simulation*, 14: 885–907, 2021. doi: 10.1007/s12273-020-0759-2.

Ján Drgoňa, Javier Arroyo, Iago Cupeiro Figueroa, David Blum, Krzysztof Arendt, Donghun Kim, Enric Perarnau Ollé, Juraj Oravec, Michael Wetter, Draguna L Vrabie, and Lieve Helsen. All you need to know about model predictive control for buildings. *Annual Reviews in Control*, 50:190–232, 2020. doi: 10.1016/j.arcontrol.2020.09.001.

Yan Du, Helia Zandi, Olivera Kotevska, Kuldeep Kurte, Jeffery Munk, Kadir Amasyali, Evan McKee, and Fangxing Li. Intelligent multi-zone residential hvac control strategy based on deep reinforcement learning. *Applied Energy*, 281:116117, 2021. doi: 10.1016/j.apenergy.2020.116117.

Ricardo Enríquez, María José Jiménez, and M del Rosario Heras. Solar forecasting requirements for buildings mpc. *Energy Procedia*, 91:1024–1032, 2016. ISSN 1876-6102. doi: https://doi.

org/10.1016/j.egypro.2016.06.271. URL https://www.sciencedirect.com/science/article/pii/S187661021630371X. Proceedings of the 4th International Conference on Solar Heating and Cooling for Buildings and Industry (SHC 2015).

Gregory Flato, Jochem Marotzke, Babatunde Abiodun, Pascale Braconnot, Sin Chan Chou, William Collins, Peter Cox, Fatima Driouech, Seita Emori, Veronika Eyring, et al. Evaluation of climate models. In *Climate Change 2013–The Physical Science Basis: Contribution of Working Group I to the Fifth Assessment Report of the IPCC*, pages 741–866. Cambridge University Press, 2014. doi: 10.1017/CBO9781107415324.

David John Gagne, Hannah M Christensen, Aneesh C Subramanian, and Adam H Monahan. Machine learning for stochastic parameterization: Generative adversarial networks in the lorenz'96 model. *Journal of Advances in Modeling Earth Systems*, 12(3):e2019MS001896, 2020. doi: 10.1029/2019MS001896.

Yuan Gao, Shohei Miyata, and Yasunori Akashi. Energy saving and indoor temperature control for an office building using tube-based robust model predictive control. *Applied Energy*, 341: 121106, 2023. doi: 10.1016/j.apenergy.2023.121106.

V Gayathry, Deepa Kaliyaperumal, and Surender Reddy Salkuti. Seasonal solar irradiance forecasting using artificial intelligence techniques with uncertainty analysis. *Scientific Reports*, 14(1):17945, 2024. doi: https://doi.org/10.1038/s41598-024-68531-3.

Peter Grant and Christoph Gehbauer. Evaluating the impacts of weather forecast inaccuracy on performance of model predictive control for dynamic facades. In *ASHRAE/IBPSA-USA Building Simulation Conference Proceedings of the 2022 Building Performance Analysis Conference and SimBuild*. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2022. doi: 10.26868/25746308.2022.C026.

Avishai Halev, Yongshuai Liu, and Xin Liu. Microgrid control under uncertainty. *Engineering Applications of Artificial Intelligence*, 138:109360, 2024. doi: 10.1016/j.engappai.2023.109360.

Rasmus Elbæk Hedegaard, Theis Heidmann Pedersen, Michael Dahl Knudsen, and Steffen Petersen. Towards practical model predictive control of residential space heating: Eliminating the need for weather measurements. *Energy and Buildings*, 170:206–216, 2018. doi: 10.1016/j.enbuild.2018.04.014.

Wei Jiang, Zhongkai Yi, Li Wang, Hanwei Zhang, Jihai Zhang, Fangquan Lin, and Cheng Yang. A stochastic online forecast-and-optimize framework for real-time energy dispatch in virtual power plants under uncertainty. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pages 4646–4652, 2023. doi: 10.1145/3583780.3614653.

Jaewan Joe and Panagiota Karava. A model predictive control strategy to optimize the performance of radiant floor heating and cooling systems in office buildings. *Applied Energy*, 245:65–77, 2019. doi: 10.1016/j.apenergy.2019.03.209.

Wouter JP Kuijpers, Duarte J Antunes, Simon van Mourik, Eldert J van Henten, and Marinus JG van de Molengraft. Weather forecast error modelling and performance analysis of automatic greenhouse climate control. *Biosystems Engineering*, 214:207–229, 2022. doi: 10.1016/j.biosystemseng.2021.12.014.

Jianhua Lin. Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, 1991. doi: 10.1109/18.61115.

Xiaoqi Liu, Parth Paritosh, Nimish M Awalgaonkar, Ilias Bilionis, and Panagiota Karava. Model

predictive control under forecast uncertainty for optimal operation of buildings with integrated solar systems. *Solar Energy*, 171:953–970, 2018. doi: 10.1016/j.solener.2018.06.038.

Li Ma, Hua Ge, Lin Wang, and Liangzhu Wang. Optimization of passive solar design and integration of building integrated photovoltaic/thermal (bipv/t) system in northern housing. *Building Simulation*, 14:1467–1486, 2021. doi: 10.1007/s12273-021-0763-1.

João Martins, Sergiu Spataru, Dezso Sera, Daniel-Ioan Stroe, and Abderezak Lashab. Comparative study of ramp-rate control algorithms for pv with energy storage systems. *Energies*, 12 (7):1342, 2019. doi: 10.3390/en12071342.

Qinglong Meng, Yuan Xi, Xingxing Zhang, Monjur Mourshed, and Yue Hui. Evaluating multiple parameters dependency of base temperature for heating degree-days in building energy prediction. *Building Simulation*, 14:969–985, 2021. doi: 10.1007/s12273-020-0752-9.

Meteorologisk institutt. What is frost. URL https://frost.met.no/index.html. Last accessed 17 April 2023.

Guan Naiqian, Guan Yuxin, and Sun Xuelian. Global temperature forecast based on arima model. In *2019 4th International Conference on Communication and Information Systems (ICCIS)*, pages 108–112. IEEE, 2019. doi: 10.1109/ICCIS49662.2019.00026.

Pawel Netzel and Tomasz F. Stepinski. Pattern-based assessment of land cover change on continental scale with application to nlcd 2001–2006. *IEEE Transactions on Geoscience and Remote Sensing*, 53(4):1773–1781, 2015. doi: 10.1109/TGRS.2014.2348715.

Norwegian Meteorological Institute. Catalog. URL https://thredds.met.no/thredds/metno.html. Last accessed 17 April 2023.

Frauke Oldewurtel, Alessandra Parisio, Colin N Jones, Dimitrios Gyalistras, Markus Gwerder, Vanessa Stauch, Beat Lehmann, and Manfred Morari. Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings*, 45: 15–27, 2012. doi: 10.1016/j.enbuild.2011.09.022.

OpenWeather. Current weather and forecast. URL https://openweathermap.org/. Last accessed 17 April 2023.

Steffen Petersen and Katrine Wieck Bundgaard. The effect of weather forecast uncertainty on a predictive control concept for building systems operation. *Applied Energy*, 116:311–321, 2014. ISSN 0306-2619. doi: https://doi.org/10.1016/j.apenergy.2013.11.060. URL https://www.sciencedirect.com/science/article/pii/S0306261913009677.

Rivka Reichman and Yael Dubowski. Gaseous pollutant transport from an underground parking garage in a mediterranean multi-story building—effect of temporal resolution under varying weather conditions. *Building Simulation*, 14:1511–1523, 2021. doi: 10.1007/s12273-020-0757-4.

Christophorus Beneditto Aditya Satrio, William Darmawan, Bellatasya Unrica Nadia, and Novita Hanafiah. Time series analysis and forecasting of coronavirus disease in indonesia using arima model and prophet. *Procedia Computer Science*, 179:524–532, 2021. doi: 10.1016/j.procs.2021.01.036.

Sebastian Scher and Gabriele Messori. Predicting weather forecast uncertainty with machine learning. *Quarterly Journal of the Royal Meteorological Society*, 144(717):2830–2841, 2018. doi: https://doi.org/10.1002/qj.3410. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3410.

Murat Cihan Sorkun, Christophe Paoli, and Özlem Durmaz Incel. Time series forecasting on solar irradiation using deep learning. In *2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*, pages 151–155. IEEE, 2017. URL https://api.semanticscholar.org/CorpusID:23986537.

The University of Utah. Mesowest data. URL https://mesowest.utah.edu/. Last accessed 17 April 2023.

Dan Wang, Yangzhe Chen, Wei Wang, Cheng Gao, and Zhe Wang. Field test of model predictive control in residential buildings for utility cost savings. *Energy and Buildings*, 288:113026, 2023. doi: 10.1016/j.enbuild.2023.113026.

Zhe Wang and Tianzhen Hong. Reinforcement learning for building controls: The opportunities and challenges. *Applied Energy*, 269:115036, 2020. doi: 10.1016/j.apenergy.2020.115036.

Muhammad Waqas, Usa Wannasingha Humphries, Bunthid Chueasa, and Angkool Wangwongchai. Artificial intelligence and numerical weather prediction models: A technical survey. *Natural Hazards Research*, 2024. doi: 10.1016/j.nhres.2024.11.004.

Tianshu Wei, Yanzhi Wang, and Qi Zhu. Deep reinforcement learning for building hvac control. In *Proceedings of the 54th Annual Design Automation Conference 2017*, pages 1–6, 2017. doi: 10.1145/3061639.3062224.

Hongyu Wu, Annabelle Pratt, and Sudipta Chakraborty. Stochastic optimal scheduling of residential appliances with renewable energy sources. In *2015 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2015. doi: 10.1109/PESGM.2015.7286584.

Wanfu Zheng, Dan Wang, and Zhe Wang. Economic model predictive control for building hvac system: A comparative analysis of model-based and data-driven approaches using the boptest framework. *Applied Energy*, 374:123969, 2024. doi: 10.1016/j.apenergy.2024.123969.