

## Physics-Informed Kalman Filtering for Multi-Step Bias Correction in Indoor Temperature Forecasting

---

Georgios Daoutis<sup>1</sup>, Othon Tomoutzoglou<sup>2</sup>, George Kornaros<sup>3</sup>,  
and Marcello Coppola<sup>4</sup>

<sup>1</sup>Hellenic Mediterranean University, Greece

<sup>2</sup>Neotera, Italy

<sup>3</sup>Hellenic Mediterranean University, Greece

<sup>4</sup>STMicroelectronics, France

### Abstract

Indoor temperature prediction is often utilized to maintain the right environment in which wine can be fermented or stored. However, forecasting these indoor microclimates is challenging due to inaccuracies in outdoor temperature predictions combined with the difficulty of modelling a building's thermal inertia, which changes due to unpredictable environmental factors. This paper proposes a two-stage machine learning and state-estimation pipeline to predict temperatures inside a winery up to an 18-hour horizon. First, feature extraction is performed to train an Autoregressive with exogenous inputs (ARX) model. Then, to mitigate the drift caused by factors such as noise in the API predictions or errors caused by the model, a Two-State Kinematic Kalman Filter (KKF) is introduced. The proposed KKF lies in its dual-state formulation which tracks both the magnitude of the prediction error and its latent velocity. In this way, the filter effectively models the building's thermal inertia and identifies the steady, low-frequency velocity of the thermodynamic drift, resulting in a stable and physically realistic forecast.

**Keywords:** Microclimate forecasting, Kinematic Kalman Filter, ARX modelling, smart agriculture, building thermal inertia, time-series error correction, winery climate control.

## 11.1 Introduction and Background

Fermentation processes in viticulture and winemaking are highly sensitive to thermal fluctuations and extreme temperatures; therefore, indoor temperatures must often be maintained within specific limits [1]. To prevent detrimental thermal instability and maintain ideal conditions, facility managers or IoT infrastructure need reliable, multi-hour forecasts of indoor microclimates to proactively manage indoor conditions [2]. Meteorological APIs powered by global atmospheric models, such as Open-Meteo, can supply baseline predictions of future outdoor conditions that will eventually affect indoor temperature. Despite these advantages, predicting indoor microclimates using external macro-weather data with standard ARX models presents challenges. These challenges arise because forecasts made by API often contain inaccuracies for the specific area of interest. As these global models operate on wide spatial grids, they often fail to capture localized temperature and solar radiation anomalies that impact the building's exterior and eventually the indoor temperature [3]. Furthermore, the agricultural buildings have thermal inertia often influenced by unpredictable circumstances such as open doors or windows [4]. Consequently, when predicting iteratively several hour horizons, this macro-weather noise cascades into the indoor thermal model. The small, step-by-step prediction errors can compound causing the base model to drift significantly from the actual indoor temperature trajectory [5]. If left uncorrected this cumulative drift results in significant control latency and can trigger HVAC activations, undermining the energy efficiency and stability goals of smart agricultural Edge AI systems [6].

### 11.1.1 Contribution

This paper proposes a hybrid, two-stage forecasting pipeline to alleviate the problems arising from noise in the data.

- An ARX model was developed using Ridge Regression. To train this model correctly, we created and tested several features related to the inertia of the winery and identified the most valuable ones using L1 regularization (LassoCV). This feature selection was performed seasonally to capture shifting thermal dependencies, such as the increased impact

of solar radiation during summer or the dominance of indoor thermal mass in winter.

- To correct the multi-step drift inherent in recursive ARX predictions, a Two-State Kinematic Kalman Filter (KKF) is introduced. Unlike traditional filters that treat error as random noise, the proposed KKF models the prediction bias as a dynamic state with both position and latent velocity components. By focusing on the velocity of the temperature drift, the filter effectively respects the building's thermal inertia and ignores transient observation noise. A Differential Evolution algorithm is utilized to optimize the baseline noise matrices, the decaying bias parameters, and a thermal anchoring weight  $\beta$ . This ensures that the correction filter identifies genuine thermodynamic trends, providing a stable and physically realistic forecast up to 18-hour horizons.

## 11.2 Data Preprocessing and Feature Engineering

To implement the base prediction model, we must collect data for the winery that we are interested in, along with the corresponding meteorological predictions for the particular area from the external API. After the acquisition of the raw IoT and API datasets, it is important to construct and select features designed to capture the building's thermal inertia without overparameterizing the ARX model. This section outlines the data acquisition, feature engineering, and dimensionality reduction steps required to build the forecasting ARX baseline.

### 11.2.1 Data Sources

To test our model, we retrieved historical 24-hour-ahead predictions for solar radiation and outdoor temperature from the Open-Meteo API [8]. These forecasts were generated by the AROME France model [7] for the specific location of the winery.

The dataset used consists of indoor temperature measurements collected from an IoT sensor deployed within a winery facility.

### 11.2.2 Feature Construction and Thermal Lag Analysis

To find how exogenous drivers such as outdoor temperature and solar radiation affect the indoor temperature, we must quantify the temporal behaviour of the winery and explore how the indoor temperature responds, and with

what delay, to these external influences. Therefore, we implemented a statistical analysis of the dataset for both indoor and outdoor data to find minimum, maximum, and average measurements, along with cross-correlation analysis to quantify the exact physical delay or thermal lag between external forcing and internal response in three representative intervals of the year:

- **Summer (June 3 – July 21):** During this interval outdoor API temperatures averaged  $20.63^{\circ}\text{C}$  reaching peaks of  $39.40^{\circ}\text{C}$  and exhibited an average daily temperature swing of  $11.25^{\circ}\text{C}$  and a maximum of  $19.90^{\circ}\text{C}$ . The solar heat peaked at  $918.00\text{ W/m}^2$  driving the average indoor daily swing to  $7.47^{\circ}\text{C}$  and pushing the absolute maximum indoor temperature to  $31.50^{\circ}\text{C}$ . Using Pearson cross-correlation, we found a delay of about 6 hours for the outdoor temperature to completely affect the indoor temperature with  $r = 0.947$ , and an 8-hour delay for the penetration of solar radiation with  $r = 0.805$ .
- **Autumn (Sep 10 – Nov 17):** In the autumn period, the mean API outdoor temperature dropped to  $13.23^{\circ}\text{C}$ , reaching an absolute peak of  $29.70^{\circ}\text{C}$ , while the indoor mean was maintained at  $16.87^{\circ}\text{C}$  with an absolute maximum of  $26.68^{\circ}\text{C}$ . The average outdoor daily swing of  $7.27^{\circ}\text{C}$ , a maximum daily swing of  $17.30^{\circ}\text{C}$ , and solar radiation peaking at  $720.00\text{ W/m}^2$  the indoor microclimate experienced only moderate average diurnal swings of  $3.55^{\circ}\text{C}$ . As the external thermal load decayed, the thermal resistance of the facility shifted accordingly, reducing the optimal cross-correlation lags to 5 hours for outdoor temperature  $r = 0.869$  and 7 hours for solar radiation  $r = 0.737$ .
- **Winter (Dec 9– Jan 31):** Conversely, the winter dataset demonstrated internal stability as the mean API outdoor temperature dropped to  $3.49^{\circ}\text{C}$ , reaching an absolute peak of  $11.60^{\circ}\text{C}$ , with an average daily swing of  $4.65^{\circ}\text{C}$  and a maximum daily swing of  $11.10^{\circ}\text{C}$ . External solar radiation was minimal, averaging just  $32.48\text{ W/m}^2$  and peaking at only  $319.80\text{ W/m}^2$ . Because of the building's massive thermal persistence and the lack of external solar pressure, the indoor temperature was stable, maintaining a mean of  $11.24^{\circ}\text{C}$  with an absolute maximum of only  $15.30^{\circ}\text{C}$  and a mere  $1.51^{\circ}\text{C}$  average daily swing. The temperature lag reduced to 4 hours and the correlation for solar radiation collapsed entirely  $r = 0.246$  at a 5-hour lag, proving that direct solar heat transfer is negligible during the winter months.

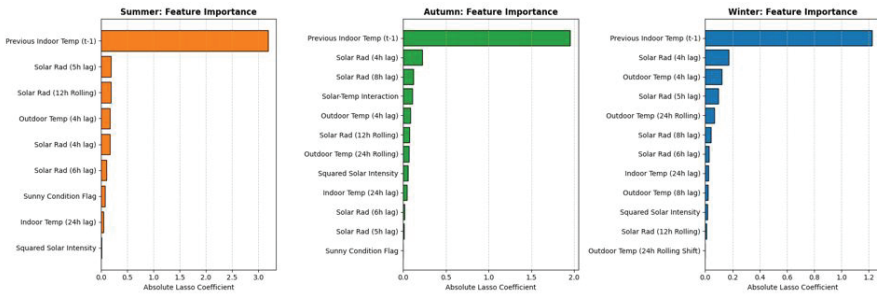
### 11.2.3 Feature Construction

Driven by these empirical lag variables, three distinct sets of temporal features were engineered from the API data to enable the linear ARX model to capture the winery's specific thermal profile:

- **Lagged Variables:** To model the thermal inertia, discrete lagged versions of the raw API data outdoor temperature and solar radiation were generated. Specifically, we explored a discrete interval of 4, 5, 6, 7, and 8 hours, as this represents the established time required for exogenous variables to noticeably penetrate the building. Furthermore, an immediate previous-step indoor temperature  $t - 1$  and a 24-hour lagged indoor temperature  $t - 24$  were included to capture the facility's daily autoregressive thermal persistence.
- **Rolling Windows:** To capture long-term environmental heat accumulation, moving average windows of 8 hours were calculated for both solar radiation and outdoor temperature. This effectively aligns the immediate API predictions with the building's broader thermal envelope.
- **Interaction and Non-Linear Terms:** High ambient heat combined with direct sunlight impacts the heating of the building non-linearly. To ensure the standard linear ARX model could capture these dynamic non-linear interaction features were constructed. First, an interaction term was created by multiplying the predicted outdoor temperature with shortwave solar radiation to capture their combined thermal load. Second, squared solar intensity was calculated. Finally, Boolean flags were engineered to isolate specific solar states: a baseline daylight flag (taking the value 1 when solar radiation exceeds  $20 \text{ W/m}^2$  and a flag extreme (taking the value 1 when radiation exceeds  $600 \text{ W/m}^2$ . These flags isolate extreme solar events, which have a particularly disruptive impact on the building's thermal response.

### 11.2.4 Feature Selection

To identify the most important features affecting the winery's microclimate and avoid redundant variables, we applied Lasso L1 regularization [9] independently for each season, implemented via the scikit-learn library [10]. The winery facility exhibits shifting physical thermodynamics across the year. For instance, during the summer period, the optimizer selected a stronger penalty of  $\alpha = 0.00336$ , and the dataset was reduced to just 9 out of 19 features. In contrast, during the stable autumn and winter periods, the model retained



**Figure 11.1** Magnitude of retained feature coefficients following Lasso L1 regularization.

more features (12 active variables in winter) using weaker environmental signals, as seen in Figure 11.1. In all seasons, however, the building’s thermal mass was the dominant predictive factor, with the immediate previous indoor temperature (t-1) exerting the highest influence by a wide margin. Crucially, this influence scaled with the seasonal thermal load: Lasso coefficients peaked at 3.18 in summer, dropped to 1.95 in autumn, and reached their lowest point at 1.22 in winter. Beyond this internal baseline, the seasonal coefficient shifts directly quantify the building’s structural thermal delay. The 4-hour to 5-hour lagged external variables specifically solar radiation and outdoor temperature emerged as the strongest exogenous predictors across all three datasets. This proves that peak external heat loads require a minimum of 4 to 5 hours to have a noticeable effect on the indoor microclimate. Furthermore, the autumn and winter models prioritized complex interaction terms and long-term rolling averages (such as the 24-hour rolling outdoor temperature). In winter specifically, the model compensated for the lack of intense, direct solar heat by leaning on these slower diurnal heat transfers to anchor the autoregressive forecast.

### 11.3 Methodology

A machine learning and state-estimation pipeline combination is proposed that first utilizes an autoregressive linear forecast to predict the baseline of the temperature trajectory and then introduces a dynamic error-correction Rolling-Window Adaptive Kalman Filter (AKF).

#### 11.3.1 The Base ARX Model

The initial predictions are generated by an AutoRegressive with eXogenous inputs (ARX) architecture, formulated using the specific subset of features

isolated by the L1 algorithm discussed in Section 3.4. The intrinsic properties of the ARX model are highly suited for modelling building thermodynamics: the autoregressive (AR) component natively captures the winery's massive thermal inertia by leveraging lagged historical indoor temperatures, while the exogenous (X) inputs integrate external meteorological driving forces, such as API-derived outdoor temperatures and solar radiation. To stabilize this architecture against the residual multicollinearity of overlapping temporal features, the model is constructed via Ridge Regression with an L2 penalty. While these linear properties provide a mathematically interpretable and physically grounded baseline for the winery's thermal trajectory, the static nature of the ARX structure cannot dynamically adapt to real-time microclimate disturbances or unmeasured API inaccuracies. Consequently, recursive multi-step predictions remain susceptible to compounding predictive drift.

### 11.3.2 Two-State Kinematic Kalman Filter

To tackle the cumulative prediction drift discussed in this chapter, a Two-State Kinematic Kalman Filter (KKF) was implemented. Traditional adaptive filters often treat ARX prediction errors as purely stochastic noise. However, in agricultural buildings, temperature errors compound through slow, continuous thermodynamic drift driven by the facility's thermal inertia. To physically model this behaviour, the proposed pipeline employs a kinematic state-space architecture that tracks both the magnitude of the prediction bias and its rate of change. First, to account for continuous meteorological volatility, the model estimates the real-time ARX prediction error as a two-dimensional state vector  $X_t$ , containing both the bias position  $X_{\text{pos}}$  and the latent bias velocity  $X_{\text{vel}}$ . For a given time, step  $t$ , the raw ARX prediction error is calculated as:

$$Y_{\text{err}, t} = Y_{\text{actual}, t} - Y_{\text{pred}, t} \quad (11.1)$$

The KKF defines the state transition matrix  $F$  and the observation matrix  $H$  to model a constant-velocity kinematic system. Because the ARX model outputs at regular hourly intervals, the transition matrix assumes the error position updates linearly based on its velocity, while only the position error is directly observable:

$$F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad H = [1 \ 1] \quad (11.2)$$

A critical, physics-informed enhancement in this architecture is the structural formulation of the process noise covariance matrix,  $\mathbf{Q}$ . Rather than dynamically reacting to all outliers equally, the algorithm isolates the noise into a diagonal matrix where the position variance  $Q_{\text{pos}}$  and velocity variance  $Q_{\text{vel}}$  are optimized independently:

$$\mathbf{Q} = \begin{bmatrix} Q_{\text{pos}} & 0 \\ 0 & Q_{\text{vel}} \end{bmatrix} \quad (11.3)$$

By isolating these parameters, the Differential Evolution optimizer is allowed to actively constrain  $Q_{\text{pos}}$  toward zero in high-inertia environments. When this occurs, the filter acts as a mathematical lock against non-physical, instantaneous temperature shocks. All thermodynamic volatility is strictly routed through the velocity variance, ensuring that the filter ignores high-frequency sensor noise and actively tracks the smooth, slow-moving drift inherent to passive building heat loss. During the iterative execution of the filter, the a priori state estimate  $X_{\text{pred},t}$  and error covariance  $P_{\text{pred},t}$  are projected forward in the predict step:

$$\mathbf{X}_{\text{pred},t} = \mathbf{F}\mathbf{X}_{t-1} \quad (11.4)$$

$$\mathbf{P}_{\text{pred},t} = \mathbf{F}\mathbf{P}_{t-1}\mathbf{F}^T + \mathbf{Q} \quad (11.5)$$

Once the new ARX prediction error is observed, the innovation and Kalman Gain  $K_t$  are calculated to optimally update the state vector, governed by the optimized measurement noise  $\mathbf{R}$ :

$$\text{innovation}_t = y_{\text{err},t} - \mathbf{H}\mathbf{X}_{\text{pred},t} \quad (11.6)$$

$$\mathbf{K}_t = \mathbf{P}_{\text{pred},t}\mathbf{H}^T (\mathbf{H}\mathbf{P}_{\text{pred},t}\mathbf{H}^T + \mathbf{R})^{-1} \quad (11.7)$$

Finally, the state vector and covariance matrix are updated (a posteriori) to be utilized in the subsequent time step:

$$\mathbf{X}_t = \mathbf{X}_{\text{pred},t} + \mathbf{K}_t \cdot \text{innovation}_t \quad (11.8)$$

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t\mathbf{H}) \mathbf{P}_{\text{pred},t} \quad (11.9)$$

### 11.3.3 Multi-Step Prediction Correction with Bias Decay

Because facility managers require reliable forecasts up to 18 hours in advance, iterative multi-step forecasting is performed. As the model predicts further into the future ( $j$  steps ahead), projecting the error velocity

in a continuous straight line from the Two-State Kinematic Kalman Filter (KKF) can lead to severe overcorrection, as environmental thermodynamics naturally fluctuate rather than drift infinitely. To safely project the state error, the KKF applies an optimized exponential decay factor  $\gamma$  to the kinematically projected bias. The corrected bias at future step  $j$  is calculated using the estimated error position  $X_{\text{pos}}$  and error velocity  $X_{\text{vel}}$  components of the state vector

$$\mathbf{Bias}_{\text{corr},j} = (\mathbf{X}_{\text{pos}} + \mathbf{X}_{\text{vel}} \cdot j) \cdot \gamma^j \quad (11.10)$$

Finally, recognizing that agricultural buildings possess massive thermal mass, the prediction output must be physically anchored. The final forecast is formulated as a weighted blend controlled by the optimized thermal persistence parameter  $\beta$  of the simulated indoor temperature from the previous time step  $y_{\text{sim},t+j-1}$  and the newly bias-corrected ARX forecast:

$$y_{\text{final},t+j} = \beta \cdot y_{\text{sim},t+j-1} + (-\beta) \cdot (y_{\text{ARX},t+j} + \mathbf{Bias}_{\text{corr},j}) \quad (11.11)$$

## 11.4 Experimental Setup and Results

This section details the training of the ARX model and the finetuning of the KKF via Differential Evolution (DE). In addition, we compare the hybrid ARX-KKF architecture against the standard ARX baseline for each season to display the predictive accuracy improvement of the system and its dynamic physical adaptation to the winery's shifting thermodynamics.

### 11.4.1 Hyperparameter Optimization

To validate the proposed pipeline and to test its adaptability, we trained the model for each specific season interval independently. For each season, we took the first 70% of the data to train the ARX model, used the next 15% of the data to calibrate the Kalman filter parameters, and used the rest as unseen samples to test the performance of the model. Specifically, the input features were normalized using standard scaling before being fitted to a Ridge Regression algorithm with a fixed regularization parameter of  $\alpha = 1.0$ . After that, we calibrated the Kalman filter using the validation dataset by employing DE to implement a global search in the multidimensional parameter space to find the optimal combination of hyperparameters that minimizes the multi-step prediction error. As the parameter space specifically the tuning of the noise covariance matrices is non-convex and prone to local minima, DE was selected for its proven ability to locate global optima in continuous

spaces without relying on gradient information [11]. The DE algorithm was deployed to simultaneously optimize five KKF parameters: the physical thermal anchoring weight  $\beta$ , the process noise variances for both the error position  $Q_{\text{pos}}$  and error velocity  $Q_{\text{vel}}$ , the baseline measurement noise,  $R_{\text{base}}$  and a kinematic projection decay factor  $\gamma$ .

### 11.4.2 Multi-Horizon Performance Analysis

The results of running the experiments described in Section 4.1 are outlined in Table 4.1. They confirm the efficacy of the proposed framework, but also the hyperparameters that the DE chose reveal how the filter shifts its behaviours for each specific season. By analysing the optimal thermal anchoring weights  $\beta$ , kinematic process noise variances  $Q_{\text{pos}}$  and  $Q_{\text{vel}}$ , and projection decay factors  $\gamma$  selected across the three seasons, a physical narrative emerges regarding how the algorithm adapts to shifting thermodynamics through the seasons.

- **Summer:** In the summer interval, the Hybrid KKF provided a 15.02% improvement at the 6-hour mark, though performance degradation accelerated at the 12- and 18-hour horizons. The aggressive diurnal solar loads caused the baseline trajectory to drift so significantly that traditional thermal persistence became a liability. To compensate for this extreme volatility, the optimizer practically eliminated the thermal blending weight  $\beta < 0.08$ , maximized the velocity process noise  $Q_{\text{vel}}, > 0.08$ , and raised the decay factor  $\gamma > 0.56$  for extended horizons. The DE acknowledged that large, sudden temperature swings in summer are true physical realities not transient observation noise and therefore shifted the forecasting weight entirely onto the active kinematic tracking of the thermodynamic drift.
- **Autumn:** In the autumn interval, the architecture achieved the highest relative improvements of the entire study, peaking at 22.89% for the 12-hour horizon. Because the external weather was not as violently volatile as mid-summer, yet not as static as winter, the KKF was able to successfully track and update the state bias without becoming saturated. The optimizer manages to balance moderate thermal anchor  $\beta \approx 0.20$  with active velocity tracking  $Q_{\text{vel}}, \approx 0.01$  to 0.04 and an elevated decay factor  $\gamma \approx 0.55$ . This configuration allowed the kinematic projection to confidently ride out the steady, predictable seasonal drift while maintaining a physically realistic anchor, yielding the most robust and stable multi-step forecasts of the year.

**Table 11.1** Predictive Performance Comparison (MAE) and Optimized KKF Hyperparameters Across Seasonal Forecasting Horizons

Season	Forecast	ARX	KKF	Impr.%	$\beta$	$Q_{\text{pos}}$	$Q_{\text{vel}}$	$R_{\text{base}}$	$\gamma$
	horizon	MAE °C	MAE °C						
Summer	6h	0.2623	0.2029	<b>22.65</b>	0.029	0.0704	0.0001	0.000	0.50
	12h	0.3607	0.2782	<b>22.89</b>	0.063	0.0284	0.0826	0.002	0.51
	18h	0.3912	0.3154	<b>19.39</b>	0.074	0.0661	0.0813	0.000	0.56
autumn	6h	0.2623	0.2029	<b>22.65</b>	0.190	0.0538	0.0159	0.223	0.50
	12h	0.3607	0.2782	<b>22.89</b>	0.222	0.0873	0.0408	0.857	0.55
	18h	0.3912	0.3154	<b>19.39</b>	0.200	0.0000	0.0172	0.161	0.51
Winter	6h	0.2314	0.2069	<b>10.61</b>	0.459	0.0000	0.0000	0.593	0.50
	12h	0.3191	0.2917	<b>8.59</b>	0.416	0.0892	0.0000	0.686	0.50
	18h	0.3754	0.3420	<b>8.90</b>	0.361	0.0000	0.0000	0.969	0.50

- Winter: During the winter months, external thermal loads and overall absolute temperatures are lower. To adapt to this low-energy, highly persistent environment, the optimizer selected a remarkably high thermal blending weight  $\beta \approx 0.46$  (at the 6-hour mark) while mathematically neutralizing the velocity tracking  $Q_{\text{vel}}, \approx 0.0$ . Furthermore, the kinematic decay factor hit the absolute lower boundary  $\gamma = 0.500$  across all time horizons. Because the winter indoor microclimate strongly resists rapid change, the filter essentially learned to distrust noisy, high-frequency errors from the ARX baseline. By setting the decay to its maximum dampening effect relying instead on the building's massive physical inertia to anchor the forecast.

## 11.5 Conclusion

To address the drift that often affects indoor temperature autoregressive predictions, we proposed a two-stage hybrid pipeline. We did this by integrating an ARX model with a Two-State Kinematic Kalman Filter (KKF) and a thermal smoothing anchor to correct the cascading errors caused by noise in the API data or other unpredictable events. We evaluated the model across three distinct seasonal intervals. For each interval, we trained the model from scratch and optimized the kinematic and thermal persistence parameters via Differential Evolution. The proposed pipeline was able to reduce multi-step forecasting errors by up to 22.89% across forecast horizons of up to 18 hours. Future work will explore replacing the linear ARX baseline with advanced recurrent neural network (RNN) architectures. Transitioning to deep learning

could allow the base model to better capture the extreme, non-linear thermal dynamics observed during peak summer solar loads, thereby enhancing the framework's overall accuracy.

## References

- [1] I. Kovačević, I. Aleksi, T. Keser, and T. Matic, "Winnie: A sensor-based system for real-time monitoring and quality tracking in wine fermentation," *Appl. Sci.*, vol. 15, no. 21, p. 11317, Oct. 2025, doi: 10.3390/app152111317.
- [2] S. Kontogiannis, M. Tsoumani, G. Kokkonis, C. Pikridas, and Y. Kotseridis, "Proposed SmartBarrel system for monitoring and assessment of wine fermentation processes using IoT nose and tongue devices," *Sensors*, vol. 25, no. 13, p. 3877, 2025.
- [3] A. Worthy, M. Ashayeri, J. Marshall, and N. Abbasabadi, "Bridging the simulation-to-reality gap: A comprehensive review of microclimate integration in urban building energy modeling (UBEM)," *Energy and Buildings*, vol. 331, p. 115392, 2025, doi: 10.1016/j.enbuild.2025.115392.
- [4] R. Felez and J. Felez, "Advanced energy management for residential buildings optimizing costs and efficiency through thermal energy storage and predictive control," *Appl. Sci.*, vol. 15, no. 2, p. 880, 2025, doi: 10.3390/app15020880.
- [5] A. Benechehab, G. Paolo, A. Thomas, M. Filippone, and B. Kégl, "Multi-timestep models for model-based reinforcement learning," arXiv preprint arXiv:2310.05672, 2023, doi: 10.48550/arXiv.2310.05672.
- [6] K. Almazam, O. Humaidan, N. M. Shannan, F. M. Bashir, T. Gammoudi, and Y. A. Dodo, "Innovative energy efficiency in HVAC systems with an integrated machine learning and model predictive control technique: A prospective toward sustainable buildings," *Sustainability*, vol. 17, no. 7, p. 2916, 2025, doi: 10.3390/su17072916.
- [7] Demortier A, Mandement M, Pourret V, Caumont O. Assimilation of temperature and relative humidity observations from personal weather stations in AROME-France. *Nat Hazards Earth Syst Sci.* 2025;25:429–449. <https://doi.org/10.5194/nhess-25-429-2025>
- [8] Zippenfenig P. Open-Meteo.com Weather API [Internet]. Zenodo; 2023 [cited 2026 Mar 13]. Available from: <https://doi.org/10.5281/ZENODO.7970649>

- [9] R. Tibshirani, Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society: Series B (Methodological)* 58 (1) (1996) 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- [10] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, et al., Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* 12 (2011) 2825–2830.
- [11] R. Storn, K. Price, Differential evolution – A simple and efficient heuristic for global optimization over continuous spaces, *Journal of Global Optimization* 11 (4) (1997) 341–359. <https://doi.org/10.1023/A:1008202821328>

