Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Forecasting Lab Climates: A Machine Learning Decision Support System for Humidity-Sensitive Experiments

Lukas Tilch

tilch@imp.uni-hannover.de

Leibniz University Hannover

Maximilian Heumann

Leibniz University Hannover

Jan Fabian Drexler

Leibniz University Hannover

Marc Mueller

Leibniz University Hannover

Article

Keywords:

Posted Date: July 15th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9940600/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Forecasting Lab Climates: A Machine Learning Decision Support System for Humidity-Sensitive Experiments

Lukas Tilch[1,*], Maximilian Heumann[2], Jan Fabian Drexler[1], and Marc Mueller[1]

[1]Institute for Multiphase Processes, Leibniz University Hannover, Garbsen, 30823, Germany
[2]Information Systems Institute, Leibniz University Hannover, Hannover, 30167, Germany
[*]tilch@imp.uni-hannover.de

ABSTRACT

The reproducibility of electrospinning is strongly influenced by the ambient environment, as temperature and relative humidity affect solvent evaporation, jet stability, and fiber formation. In practice, these variables are often documented only retrospectively, which limits the ability of researchers to anticipate weather-driven shifts in laboratory conditions and in turn can hinder reproducibility.

Here, we present a machine-learning-based decision-support framework for forecasting indoor laboratory relative humidity from local meteorological measurements and forecasts and internal climate measurements. To preserve room-specific influences such as orientation, ventilation context, and local thermal buffering, we trained and evaluated separate regression models for four laboratories within the same building rather than assuming a single shared indoor-outdoor transfer function. We then used site-specific weather forecasts to assess predictive performance and uncertainty across forecast horizons from 1 to 7 days.

To examine user-related indoor variability without relying on direct occupancy sensing, we additionally analyzed working- and non-working-day patterns using a calendar-based occupancy proxy and Wasserstein-distance heatmaps. The resulting framework provides a seven-day indoor humidity forecast with lead-day-specific prediction intervals intended to support experiment planning under weather-responsive indoor conditions.

Because the study was conducted in four laboratories within a single building, the trained models should be understood as site-specific rather than directly transferable to other laboratory environments. The broader contribution of this work therefore lies in the forecasting and decision-support workflow, which can be adapted to other settings using local data. At its current stage, the framework is intended to indicate likely indoor relative humidity shifts and their uncertainty; the experimental significance of those shifts remains to be interpreted by the researcher in the context of the specific polymer system and target fiber outcome.

Introduction

Electrospinning is a versatile route for producing micro- and nanofibrous materials, and the final fiber architecture is highly sensitive to the coupled effects of solution properties, electrical forcing, distance to the collector, feed rate, and ambient conditions during jet flight and solidification[1]. Among the environmental variables, relative humidity (RH) is especially influential because it affects charge dissipation, solvent evaporation, phase separation, fiber solidification and in turn morphology, mechanical behavior, and, in many polymer-solvent systems, crystalline structure[2,3]. As a result, RH is not merely a background laboratory parameter but a major determinant of process stability, repeatability, and product quality in electrospinning.

The effect of RH on electrospun fibers is strongly dependent on polymer-solution composition and desired result, which makes humidity control scientifically important, but deciding whether to increase or decrease humidity a complex decision. In hydrophilic or water-compatible polymer-solution systems, increasing RH can slow evaporation and delay solidification, enabling prolonged jet stretching, bead formation, fiber fusion, or film-like defects, whereas in hydrophobic systems the same humidity increase can accelerate water-induced phase separation and promote thicker, rougher, wrinkled, or porous fibers[2].

In principle, Heating, Ventilation, and Air Conditioning (HVAC) systems should buffer electrospinning environments from outdoor weather. In practice, however, indoor humidity often can retain a measurable dependence on external conditions over hourly to seasonal timescales[4]. Studies on indoor-climate prediction have shown that future indoor RH can be forecast from recent indoor and outdoor climate histories, although predictive accuracy depends on building behavior, data quality, and the degree of coupling to external disturbances[5,6]. Likewise, predictive building-control research shows that indoor states evolve over finite horizons under the combined influence of weather, infiltration, fresh-air exchange, internal gains, and building thermal mass, implying that future indoor conditions are dynamically structured rather than random.[7,8]

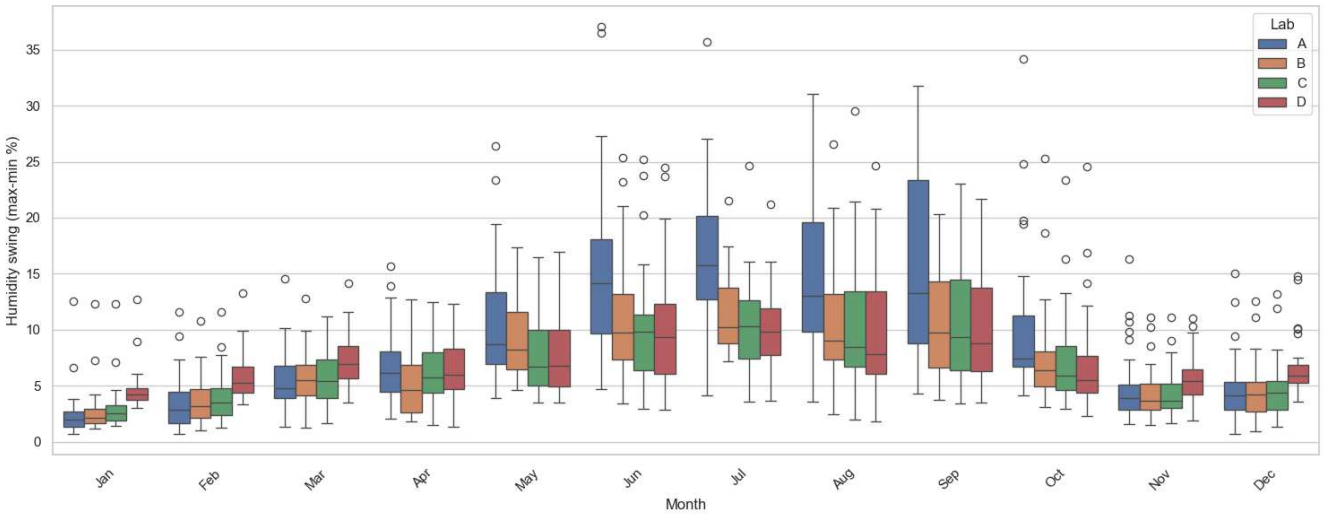


Figure 1. Boxplots of daily daytime relative humidity swing in laboratories A-D by month. Humidity swing was defined as the difference between the daily maximum and minimum relative humidity ($RH_{\mathrm{max,day}} - RH_{\mathrm{min,day}}$) measured between 08:00 and 18:00. The plots highlight both seasonal variation and inter-laboratory differences in the magnitude and dispersion of intra-day humidity fluctuations relevant to electrospinning process stability.

This point is especially relevant for electrospinning laboratories, where even modest weather-driven humidity shifts may move a process outside morphology-sensitive operating conditions before the operator recognizes the risk from current sensor readings alone[2,3]. However, although the electrospinning literature has described how humidity affects fiber outcomes and building-climate studies have demonstrated the predictability of indoor environmental dynamics, these strands have rarely been combined into a framework designed to support laboratory decision-making.

In this work, we propose a decision support system (DSS) for electrospinning laboratories that links outdoor weather evolution with indoor humidity response to provide actionable forecasts for researchers. Rather than prescribing autonomous control actions, the DSS is intended to warn operators ahead of likely weather-driven indoor RH shifts and thereby support improved scheduling, preparation, and interpretation of humidity-sensitive experiments. Because indoor-outdoor humidity transfer depends on room-specific boundary conditions such as orientation, ventilation context, and local thermal buffering, we treat the forecasting problem at the room level rather than assuming a single transferable response across laboratory spaces. In this way, the framework translates indoor-climate forecasting into practical laboratory guidance while keeping the trained models aligned with the local conditions under which they are deployed.

Results

Humidity drift over the year

Across all four validation laboratories, the daytime relative humidity swing varied over the year and differed between rooms (Fig. 1). Here, humidity swing was defined for each day as the difference between the maximum and minimum relative humidity recorded between 08:00 and 18:00.

The smallest daily swings occurred in winter. In January, the median swing ranged from 1.95% in Lab A to 4.23% in Lab D, and in February from 2.83% in Lab A to 5.24% in Lab D. Median swings increased during spring, with April values between 4.64% and 6.14%, and May values between 6.67% and 8.73%.

The highest median daytime swings were recorded from June to September. Lab A showed the largest median values in this period, reaching 14.12% in June, 15.78% in July, 12.99% in August, and 13.28% in September. Over the same months, median swings in Labs B-D were lower and remained between 7.81% and 10.31%.

Large day-to-day excursions were also evident in summer and early autumn. In Lab A, some individual days showed daytime humidity swings above 30%, with the highest values occurring in September. In Labs B, C, and D, individual daily swings also exceeded 15% during this period.

Differences between laboratories were present throughout the year. Lab A generally showed the largest daytime swings from May to September. By contrast, Lab D showed the highest winter medians among the four rooms, with median swings of 5.40% in November, 5.89% in December, 4.23% in January, and 5.24% in February.

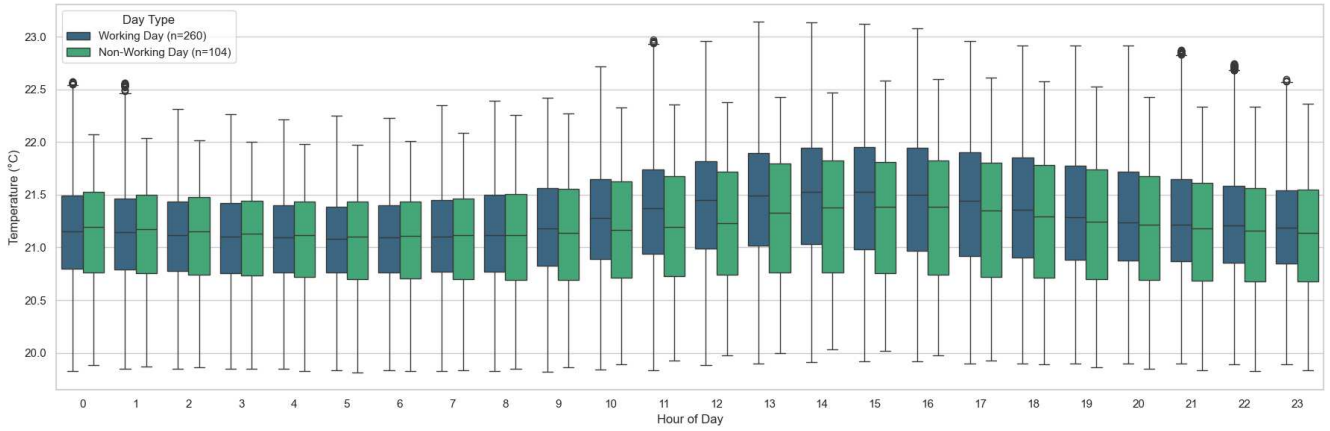


Figure 2. Workday impact on variance in given timeframes for Lab A: Days are categorized into working and non-working days with working day categorization (15). The data is plotted over hour of day to compare variance in form of whisker spread (Tukey's Fences)

Working-day occupancy proxy analysis

To assess whether human activity introduced a systematic and learnable signal into the indoor climate, indoor temperature was analyzed using a calendar-based working-day proxy. Across the four laboratories, the comparison between working and non-working days revealed clear time-of-day differences, indicating that room usage contributed measurably to short-term indoor climate variability. Temperature was particularly suitable for this comparison because occupancy-related effects appear more directly and rapidly in the thermal signal: human presence, door opening, and experiment-related activity can alter indoor temperature over short time scales, whereas relative humidity is typically more buffered by moisture storage in the room and pressure disturbances tend to equalize quickly[9,10].

The hourly boxplots in figure 2 showed that temperature distributions during working days were consistently broader than during non-working days, particularly within the core daytime period. This effect was expressed less by a strong shift in median temperature than by an increase in spread, with the whisker extent generally exceeding those of non-working days.

The day-by-hour Wasserstein distance analysis in figure 3 confirmed these observations and provided a more explicit view of weekly laboratory use patterns. When non-working days were used as the low-occupancy baseline, the resulting heatmaps showed elevated distances primarily during normal working hours, demonstrating that weekday temperature distributions deviated systematically from the baseline state. The strongest deviations were concentrated in the daytime, whereas nighttime periods remained comparatively similar to the non-working reference, consistent with reduced room activity outside regular operating hours.

A recurring weekly structure was visible across the laboratories. Monday and Friday generally showed smaller deviations from the non-working baseline than the mid-week period, whereas the central weekdays showed larger and more persistent deviations.

Room-specific differences were also apparent. Laboratories with standard daytime use showed broad weekday daytime deviation windows, the laboratory characterized by earlier and longer experiments showed an earlier onset of deviation from the non-working baseline, and the intermittently used laboratory showed stronger deviations concentrated on mid-week days.

Model performance

The blocked temporal validation showed that indoor relative humidity in Lab C could be predicted from engineered outdoor, temporal, solar-geometry, and calendar-based features. Figure 4 shows that across the model achieved mean absolute errors between 0.532 and 2.154 percentage points RH and R^2 values between 0.9179 and 0.9927.

Model performance was strongest in January, February, November, and December, where the MAE remained below 0.8 percentage points RH and R^2 ranged from 0.9826 to 0.9928. October showed an MAE of 0.857 percentage points RH and $R^2 = 0.9817$.

Performance was lower from March through September. Within this period, the highest MAE was observed in September at 2.154 percentage points RH, while the lowest R^2 was observed in March at 0.918.

When the final model was trained on the full available development dataset and evaluated on a fully unseen period from 8 December 2025 to 9 March 2026, it achieved an MAE of 0.747 percentage points RH and an R^2 of 0.982, as shown in Fig. 5.

Feature-importance analysis in Fig. 6 showed a similar ranking across the tree-based models used for Lab C. In both

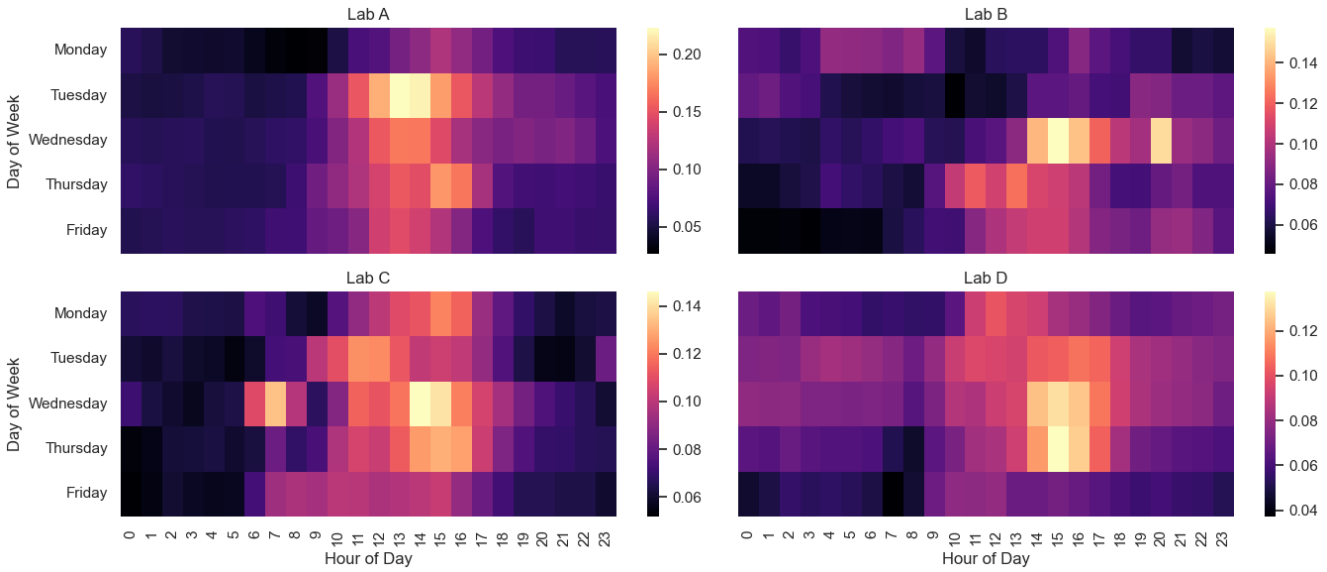


Figure 3. Probability distribution shifts for indoor temperature (Wasserstein): Computed by comparing working and non-working days and plotted on a Day-Of-Week Hour-Of-Day heatmap to visualize temporal patterns in lab usage. Brighter colors indicate a higher variance difference between working and non-working days for the given hour of day and day of week.

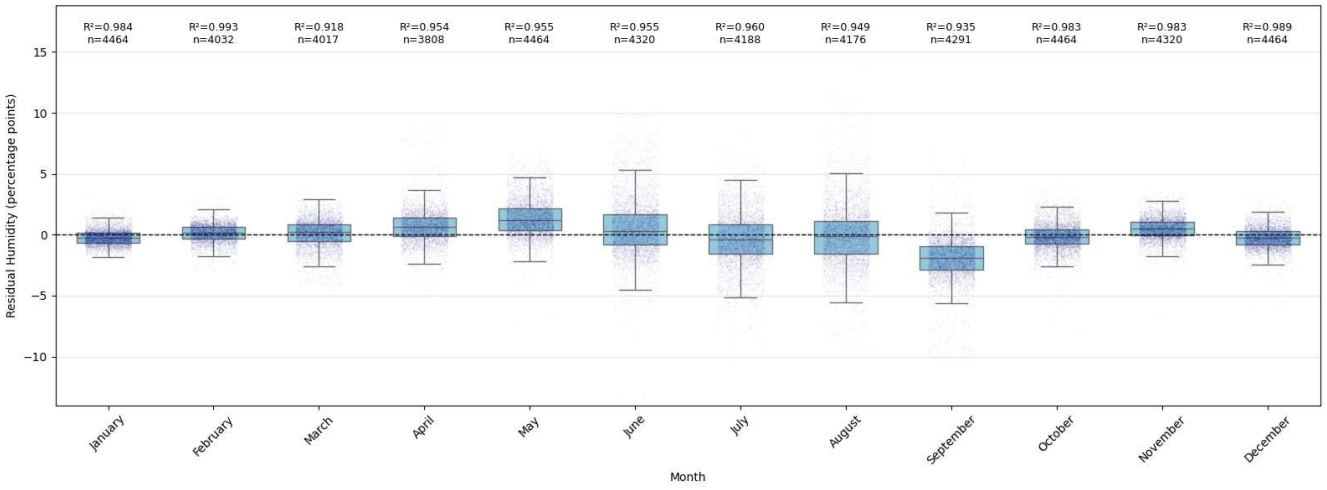


Figure 4. Monthly distribution of residual errors from blocked-month validation of the XGBoost model for indoor relative humidity prediction. Residuals are defined as $RH_{predicted} - RH_{observed}$ and are reported in percentage points for all 10-min observations within each held-out calendar month. The figure therefore summarizes month-specific model performance and spread across the annual cycle, while the labels above each month give the corresponding coefficient of determination (R^2) and the number of evaluated samples (n).

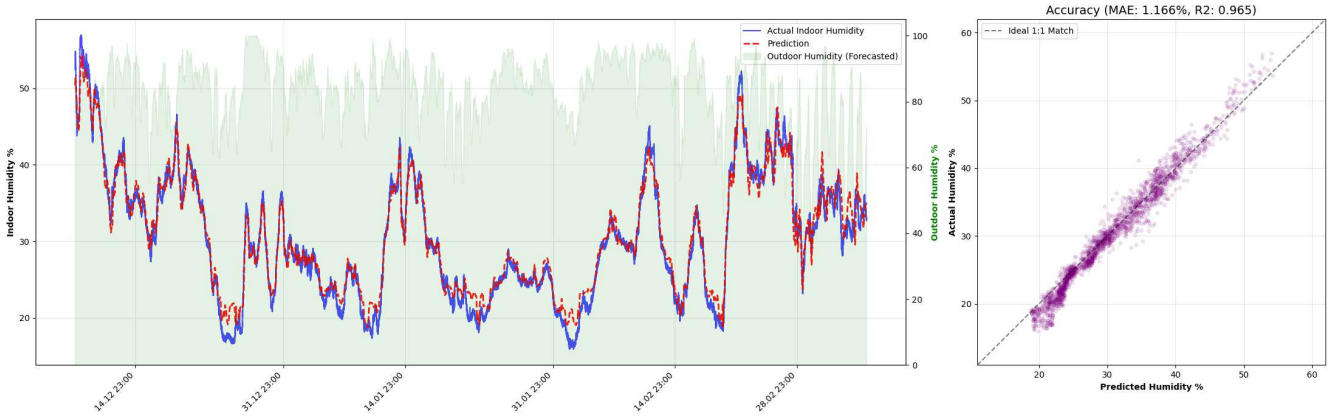


Figure 5. Prediction of indoor relative humidity in Lab C using the most recently updated forecast[11]. The left panel shows observed indoor relative humidity and the corresponding forecasted indoor relative humidity over time, together with forecasted outdoor relative humidity on the secondary axis. The right panel shows predicted versus observed indoor relative humidity, with the close clustering around the 1:1 reference line indicating high predictive accuracy.

Random Forest and XGBoost, outdoor dew-point variables received the highest importance scores, with the 1 h, 2 h, and 3 h moving averages ranked highest in both models, whereas outdoor temperature, solar-geometry, and calendar-based variables showed lower importance values.

No lagged indoor humidity terms were included in this baseline configuration.

Model forecast performance across prediction horizons

Model performance decreased systematically with increasing forecast horizon for all evaluated approaches. Across days 0-7, mean absolute error increased, R^2 declined, and the 95% predictive interval width broadened, indicating the expected loss of forecast skill with lead time, as shown in figure 7.

The multilinear regression (LR) baseline performed worst throughout. Its MAE increased from 2.74 on day 0 to 6.70 on day 7, while R^2 decreased from 0.814 to −0.10. In parallel, the 95% predictive interval width increased from 15.26 to 34.48, showing that the linear model was not able to represent the longer-horizon humidity dynamics reliably.

At short forecast horizons, the random forest (RF) model achieved the strongest overall performance. On day 0, RF yielded the lowest MAE (0.17), the highest R^2 (0.965), and the narrowest 95% predictive interval width (5.37). RF remained the best-performing model through day 2, with consistently lower errors, higher explained variance, and narrower predictive intervals than XGB over the near-term horizon.

At intermediate and longer forecast horizons, the performance gap between the two non-linear models became small, and XGB generally took the lead. Beginning on day 3, XGB achieved slightly lower MAE, higher R^2, and narrower predictive intervals than RF in most cases. For example, on day 4, XGB reached an MAE of 2.86, an R^2 of 0.778, and with an interval width of 16.38, marginally outperforming RF in this regard.

By the longest horizon, all models showed substantial degradation, but RF provided the strongest overall performance among the non-linear approaches. On day 7, XGB achieved the lowest MAE (4.93), the highest R^2 (0.409), and the narrowest 95% predictive interval width (26.19), while RF followed closely with similar but slightly weaker results. This suggests that RF was preferable at short lead times, whereas XGB was more robust at longer forecast horizons.

Overall, the comparison revealed a clear separation between the simple linear baseline and the more flexible non-linear models. RF was most effective for short-term prediction, whereas XGB was more accurate at longer horizons, particularly when overall robustness across error, explained variance, and predictive uncertainty was taken into account. Such horizon-dependent behavior is consistent with previous indoor-climate forecasting studies, in which predictive performance typically declines with lead time and depends on the extent to which the model captures non-linear building response and external disturbances.[5–7].

Because all methods were supplied with the same forecast input, the similar increase in error at longer lead times indicates that horizon-dependent input uncertainty dominates the total prediction error; the methods therefore do not converge, but deteriorate similarly under the same forecast constraints.

Operator-facing decision support output

To translate the forecasting framework into a form usable during laboratory planning, the final model output was rendered as an operator-facing decision-support visualization for Lab C. The DSS presents indoor relative humidity as a continuous 7-day

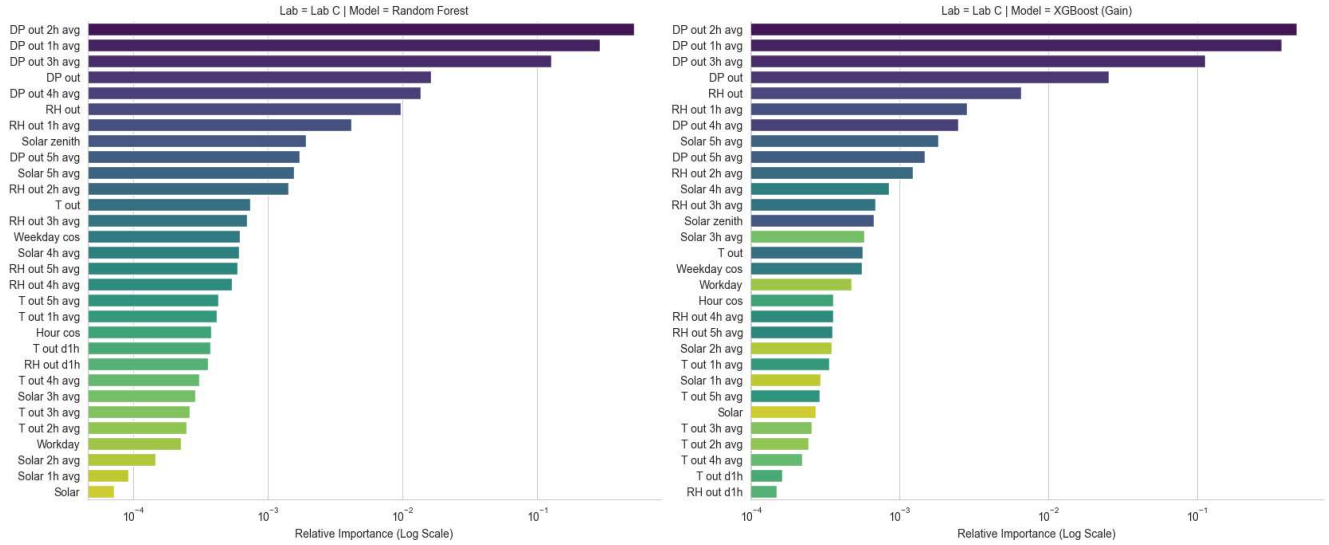


Figure 6. Feature importance for indoor humidity prediction in the electrospinning lab (Lab C) using Random Forest and XGBoost models. Relative importance is shown on a logarithmic scale. In both models, outdoor dew-point variables, particularly the 1 h, 2 h, and 3 h moving averages, rank highest.

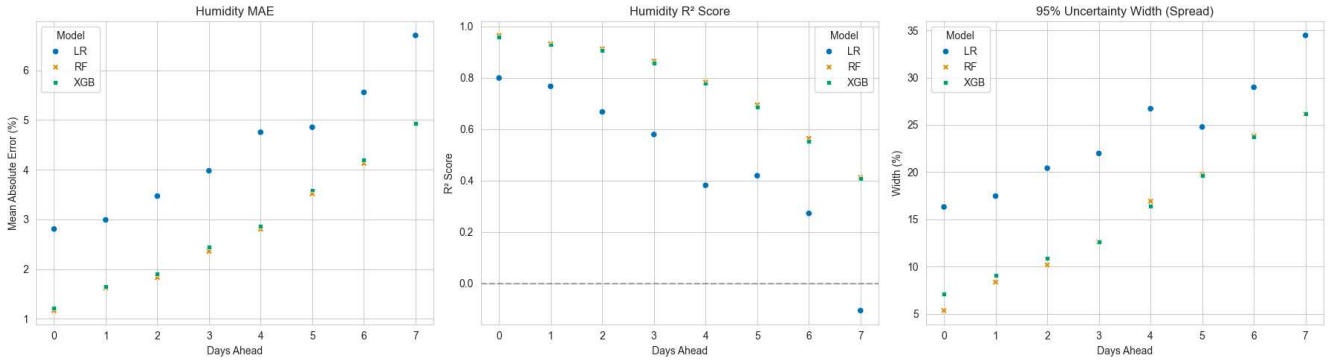


Figure 7. Forecast performance of multilinear regression (LR), Random Forrest (RF), and Extreme Gradient Boosting (XGB) models for indoor humidity prediction at increasing lead times. Model performance is shown for prediction horizons from 0 to 7 days ahead using three metrics. The left panel presents humidity MAE, the middle panel presents humidity R^2, and the right panel presents the 95 % uncertainty width, allowing direct comparison of accuracy, explained variance, and predictive spread across models and forecast horizons.

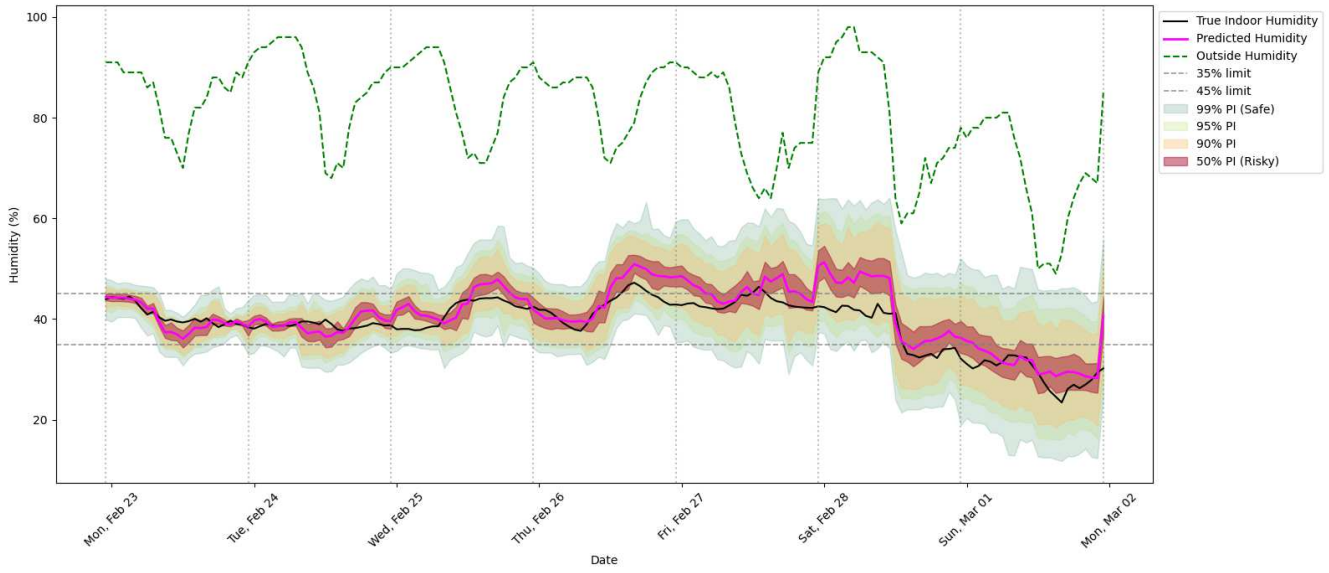


Figure 8. Operator-facing 7-day decision-support output for Lab C. The magenta line shows the forecast indoor relative humidity. The black line indicates measured indoor relative humidity and is shown in this figure for evaluation and visual comparison only; it is not part of the operational DSS, which relies exclusively on forecast information. The dashed green line shows forecast outdoor relative humidity as contextual weather forcing. Shaded bands indicate calibrated prediction intervals (50 %, 90 %, 95 %, and 99 %) applied separately for each forecast lead day, such that uncertainty increases with prediction horizon. Horizontal dashed lines are user defined working upper and lower limits for a specific polymer solvent combination. Vertical dotted lines mark day boundaries and support identification of the timing and confidence of potential RH excursions relevant to electrospinning operation.

forecast together with the corresponding outdoor humidity signal and nested calibrated prediction intervals, allowing expected indoor RH evolution and forecast uncertainty to be assessed simultaneously.

Figure 8 shows that the forecast output retained short-term structure in the indoor humidity trajectory while also communicating the progressive loss of certainty at longer lead times. Because prediction intervals were calibrated separately for each forecast day, uncertainty expanded in a horizon-aware manner rather than being represented by a single fixed band across the entire 7-day window.

The resulting display functioned as a warning-oriented summary of the forecast state. It allowed periods of probable RH excursion to be identified in advance, together with the timing and confidence of those excursions, thereby converting the probabilistic forecast into information directly relevant for electrospinning planning under weather-responsive indoor conditions.

Discussion

The present study shows that relative humidity in laboratory spaces connected to air handling units cannot be treated as a static background condition, even when the rooms are embedded within a centrally conditioned building. Instead, the indoor environment behaved as a damped but dynamically responsive system in which outdoor weather signals remained visible indoors over short time spans. For electrospinning, this is practically important because RH directly affects charge dissipation, solvent evaporation, jet solidification, phase separation, surface morphology, pore formation, and fiber diameter[2].

A central implication of the results is that building-level climate control alone did not fully suppress weather-driven variability in the investigated laboratories. The observed humidity drift suggests that indoor RH emerged from the interaction of outdoor moisture forcing, envelope buffering, air exchange, solar exposure, and room-specific operating conditions rather than from strict humidity regulation at room level. This interpretation is consistent with indoor-climate forecasting studies showing that future indoor conditions depend on both external disturbances and the dynamic response of the building, including inertia, ventilation, and control regime[5,12].

The magnitude of the observed variability is also relevant for electrospinning practice. In the present data, daytime RH swings of around 15 percentage points occurred within a single working day, and in Lab A some summer and early autumn days exceeded 30 percentage points. Changes of this order are large enough to move the process across RH ranges that are known to

alter electrospun fiber morphology and size[3]. Across the electrospinning literature, RH shifts have been reported to induce or suppress bead-on-string formation, change pore density and surface roughness, and produce large diameter changes, including an approximately 300% increase for PS between 10 and 70% RH and a fourfold decrease for PVA between 4 and 60% RH[2].

For multi-hour runs, this means that fibers deposited at the beginning and end of a single production sequence may not be formed under equivalent ambient conditions even when all nominal machine settings are kept constant. In such cases, variation may develop within the product itself rather than only between separate runs. Because RH affects not only mean fiber diameter but also porosity, surface structure, and fiber-fiber interactions, this type of intra-run drift can complicate interpretation of product quality, especially when the resulting heterogeneity is distributed through the mat rather than confined to a clearly separable top or bottom layer.

The occupancy proxy analysis further indicates that not all unexplained indoor variability should be attributed to weather. The clear weekday and hour-of-day structures found in the temperature-based working-day Wasserstein heatmaps (Fig 3) suggest that user presence, door opening, and experimental activity introduced additional short-term disturbances that were systematic rather than random. This matters for predictive modelling because it implies that part of the residual climate signal might in principle learnable from temporal and calendar features, even in the absence of direct occupancy sensing[6,7].

The month-wise validation results indicate that forecast difficulty was not constant throughout the year. Lower errors in winter and late autumn, together with weaker performance in late spring, summer, and early autumn, suggest that the indoor-outdoor coupling was not stationary across the monitoring period. A plausible explanation is that the relationship between outdoor forcing and indoor humidity changed with the thermal operating state of the building, for example between heating-dominated, cooling-dominated, and transitional periods. This interpretation remains tentative because the present dataset does not yet cover multiple years.

Regarding feature-importance analysis, across both Random Forest and XGBoost, the most influential predictors were outdoor dew-point variables, particularly the 1 h, 2 h, and 3 h moving averages. The steep gradient in feature importance suggests that, for this building, outdoor dew point is the dominant driver of indoor RH prediction, likely because the centralized ventilation system continuously exhausts indoor air and supplies fresh outdoor air, thereby creating strong coupling between indoor and outdoor moisture conditions. This is in line with findings from Mendes et. al. with strong coupling of absolute outside and inside humidity in a similar vented clean room[4] and Psomas et. al. also state that ventilation is a critical factor regarding relative indoor humidity[13]. The lower-ranked contribution of solar, temperature, and calendar-based variables does not imply that these factors were irrelevant. They still contributed to overall model performance, although their predictive effect was weaker than that of the moisture-related descriptors within the present feature set and study period.

The model results support that interpretation. Even under the deliberately restrictive baseline formulation, in which no lagged indoor humidity terms were provided, indoor RH in Lab C could be reconstructed with high accuracy from engineered outdoor, temporal, solar-geometry, and calendar-based predictors alone. This indicates that the selected exogenous feature set captured a substantial part of the physically meaningful forcing structure governing indoor humidity. At the same time, the month-wise blocked validation showed that forecast difficulty varied across the year, with lower errors in winter and late autumn and weaker performance in late spring, summer, and early autumn. Although the current dataset does not yet justify a strong annual-seasonality claim, the pattern suggests that the indoor-outdoor transfer function was not stationary and may have changed with the thermal operating state of the building, particularly between heating-dominated, cooling-dominated, and transitional periods.

A likely methodological extension would be to include lagged indoor humidity or temperature terms and to propagate predictions recursively across the forecast horizon. Such an approach could plausibly improve predictive performance, especially at short lead times, because many indoor-climate forecasting studies use historical indoor-state information to represent inertia and delayed system response[5,6]. However, it would also make the framework more operationally demanding by increasing model complexity, tightening the dependence on uninterrupted recent indoor measurements, and introducing additional error propagation as each forecast step becomes an input to the next. In the present context, the decision to rely on exogenous predictors and current indoor sensing therefore served not only as a conservative modeling test, but also as a deliberate design choice in favor of accessibility, portability, and low-barrier deployment across laboratories with limited monitoring infrastructure.

An important methodological implication is that these indoor-outdoor transfer characteristics should not necessarily be expected to collapse into a single shared response across rooms. The four laboratories differed in window orientation, room size, ventilation context, and typical usage pattern, and the occupancy-proxy analysis showed that these usage differences were reflected in distinct temporal disturbance structures. For that reason, the use of separate room-level models should be understood not as a weakness of the framework, but as a deliberate design choice that preserves laboratory-specific climate dynamics relevant to operational forecasting. In the present application, the aim was therefore not to maximize room-to-room generalization, but to obtain the most decision-relevant forecast for each target laboratory under its own local boundary conditions.

The horizon-dependent model comparison adds a second practical insight. Forecast skill decreased systematically with lead time for all approaches, which is expected because longer-range predictions must propagate uncertainty from both the indoor response and the weather forecast itself. The strongest near-term performance of the non-linear models, together with the weaker performance of the linear baseline, suggests that the humidity dynamics of the laboratories are not well described by simple linear relationships alone. This agrees with prior building-forecasting studies, which report that predictive performance depends on how well the model captures non-linear interactions between recent boundary conditions, internal loads, and the dynamic state of the building[5–7].

From a laboratory operations perspective, these findings justify the use of a decision support system rather than relying only on current sensor readings. In electrospinning, the critical issue is often not that humidity eventually changes, but that it crosses morphology-sensitive ranges before the operator has time to react[2,3]. A forecast-centered DSS can therefore create practical value even if it does not eliminate uncertainty: it gives the operator time to postpone a run, bring forward a run, adapt polymer concentration or solvent composition, prepare the room earlier, or interpret later fiber deviations against forecast humidity stress. This operator-facing concept also fits the actual needs of electrospinning laboratories better than autonomous HVAC control, because the relevant intervention is dependent on the desired outcome and often needs a manual change in process setup rather than a direct actuation of the building system.

The present empirical demonstration was carried out in four laboratories within a single building, so the specific transfer functions learned here will likely require re-estimation or adaptation for laboratories with different constructions or operating regimes. The proposed model was developed and validated in a laboratory environment where each room is served by a centralized air-handling unit. However, the broader contribution of this work is not a single globally transferable fitted model, but a forecasting and decision-support framework that can be adapted to other laboratories through the same pipeline of indoor and outdoor data acquisition, preprocessing, feature engineering, uncertainty calibration, and user-facing warning output. In that sense, the framework is intended to be portable, while the trained models remain site-specific and should be re-estimated using local laboratory data.

Second, the available monitoring period did only represent one year, so any interpretation of seasonal dependence must remain provisional. Third, the occupancy analysis was based on a calendar proxy rather than direct measurements of presence, door state, or experiment schedules. This may leave some user-driven disturbances only indirectly represented, but it is also a deliberate strength of the framework, because it keeps the required instrumentation limited to indoor climate sensing, outdoor weather data, and forecast inputs, thereby improving accessibility and lowering the barrier to adoption in laboratories that do not have access to more specialized building-monitoring infrastructure. Fourth, although Lab C was the electrospinning laboratory of primary interest and therefore the main focus of detailed exogenous-feature validation, future studies should evaluate the full framework across a broader range of laboratories and climates to test how readily the feature set, model classes, and uncertainty-calibration strategy transfer to other operational settings. Finally, the current forecasting framework predicts indoor humidity but does not formally encode polymer-specific response thresholds, the translation from forecasted RH to expected fiber consequences is intentionally left to operator expertise, with formal polymer-specific thresholds identified as a key target for future enhancement.

These limitations point directly to the next development steps. Longer monitoring periods covering multiple full seasonal cycles would allow the stability of the indoor-outdoor coupling to be tested more rigorously and would improve blocked temporal validation under truly repeated annual regimes. Incorporating direct occupancy indicators, door events, or local ventilation-state information could still be explored as optional extensions for settings where such data are available, but they are not a prerequisite for using the present framework. Equally important, the framework is well suited to continuous refinement during operation, because newly collected local sensor data can be incorporated into periodic retraining and recalibration cycles, allowing the forecasting model to improve over time as it observes additional seasons, weather regimes, and laboratory usage patterns. Most importantly for electrospinning practice, future work should link the humidity forecast to formulation-specific process windows so that the DSS can move from a general warning system to a polymer-aware advisory tool that translates predicted RH excursions into concrete, human-readable recommendations for laboratory operation.

Methods

Validation laboratories

To validate the proposed forecasting and decision-support approach, four laboratories within the same centrally conditioned building were selected as test environments. The buildings air exchange system consists of a central large air handling unit with filters, a heat recovery system and heating and cooling coils to precondition the fresh air and releases the exhaust air. The exhaust air from the fume hoods is vented separately. The rooms were chosen to span relevant boundary conditions for indoor humidity dynamics while maintaining broadly comparable building services. Laboratory A is a small north-facing room with heating, ventilation, and centralized air conditioning, used mainly during standard daytime operation. Laboratory B is a large north-facing room with the same building services but a more intermittent usage profile, with activity concentrated

primarily on Wednesdays and Thursdays. Laboratory C is a small east-facing room with comparable building services and is characterized by an early daily start and longer-duration experiments. Laboratory D is a large east-facing room with the same service configuration and standard daytime use. The Laboratories all have an air exchange rate of 50 m³/h are located in the Leibniz University Hannover campus for mechanical engineering in Garbsen Germany and were built in 2019.

Separate room-level models were used rather than assuming a single shared indoor-outdoor transfer function across all spaces, because local boundary conditions were expected to modulate how outdoor weather signals propagate indoors. In particular, differences in room size, orientation, usage schedule, and local cooling context were expected to influence solar gains, buffering behavior, and the effective coupling between outdoor and indoor humidity conditions. The observed weekday structure was broadly consistent with operational impressions and with German hybrid-work statistics showing that workplace presence is typically higher on Tuesdays through Thursdays than on Mondays and Fridays[14].

Data acquisition

Indoor climate data were collected at 1 min resolution using Bosch BME280 sensor modules interfaced with Raspberry Pi units. The sensors measured air temperature, relative humidity, and barometric pressure, and all observations were transmitted via the local area network to a backend server running a database for storage and retrieval. Continuous indoor monitoring data were available from 14 April 2025 onward.

Meteorological boundary conditions were compiled from two external sources. Deutscher Wetterdienst (DWD) data were used for model training owing to their finer temporal resolution (10 min) and included air temperature, relative humidity, pressure, diffuse sky radiation, global radiation, and sunshine duration. Open-Meteo data were used both to reconstruct past forecasts for the calibration of prediction intervals and to provide operational forecast inputs at hourly resolution. All time series were cleaned and synchronized prior to analysis[11,15–17].

DWD station data were used as meteorological inputs; however, solar radiation variables for the Langenhagen station were no longer available from 16 December 2025 onward. Consequently, station-based solar predictors were only used up to this date.

Data preprocessing

Raw outdoor data from DWD were parsed from compressed station files and converted to timestamped tabular records. Air-temperature and solar datasets were merged on their common time index, non-informative metadata columns were removed, and the remaining variables were renamed to physically interpretable labels. For the DWD station record, only observations after 14 April 2025 were retained for further processing.

Outdoor data quality control was performed before interpolation. Records were retained only when air temperature and relative humidity were both available and within physically plausible bounds (-100 to $100\,^{\circ}$C for temperature and 0 to 100% for relative humidity). The cleaned outdoor series were then resampled to 1 min resolution and linearly interpolated onto a regular time grid. To avoid bridging major outages, gaps of 20 min or longer were identified prior to interpolation and flagged; timestamps falling within such gaps were excluded from the final merged dataset.

Indoor sensor exports were concatenated across files, timestamps were parsed and rounded to 1 min resolution, and the measurement columns were converted to numeric format. For each room, temperature, pressure, and relative humidity were aggregated to 1 min means. Indoor records were retained only when temperature and relative humidity were present and within the same plausibility bounds applied to the outdoor data.

The cleaned indoor and outdoor datasets were merged by timestamp for each room. Rows associated with previously flagged large interpolated outdoor gaps were removed, and rows lacking valid outdoor temperature or humidity after merging were discarded. The resulting room-specific datasets were exported as analysis-ready time series for subsequent feature engineering and model development.

Feature engineering

From the cleaned indoor-outdoor time series, additional predictors were derived to represent moisture state, temporal periodicity, solar exposure, recent weather history, and expected occupancy-related effects. This follows common indoor-climate and building-prediction practice, in which derived psychrometric variables and calendar-related descriptors are added to the predictor set to better represent non-linear environmental dynamics[6,7].

Moisture-state variables. Dew-point temperatures were calculated from air temperature and relative humidity using the Magnus-Tetens approximation[18]. An intermediate term, $\gamma(T,\mathrm{RH})$, was first defined, after which the dew-point temperature, T_{dp}, was obtained as follows:

$$\gamma(T,\mathrm{RH}) = \ln\left(\frac{\mathrm{RH}}{100}\right) + \frac{T}{c+T}, \tag{1}$$

$$T_{dp} = \frac{c\,\gamma(T,\mathrm{RH})}{b-\gamma(T,\mathrm{RH})}. \tag{2}$$

where $b = 17.625$ and $c = 243.04\,^{\circ}\mathrm{C}$ corresponding to a commonly used Magnus coefficient set for dew-point calculations.[19]. The inclusion of derived moisture variables is consistent with earlier indoor-climate forecasting studies, in which psychrometric transformations such as absolute humidity were introduced to represent moisture dynamics more effectively than raw temperature and relative humidity alone[5,6,20,21].

Temporal descriptors. To preserve the cyclical structure of time, temporal variables were encoded using circular transformations. For hour of day h and day of week d_{week} cosine encodings were defined as

$$h_{\cos} = \cos\left(\frac{2\pi h}{24}\right), \tag{3}$$

$$d_{\mathrm{week,cos}} = \cos\left(\frac{2\pi d_{\mathrm{week}}}{7}\right), \tag{4}$$

Such time-based descriptors are well aligned with prior building-prediction studies, in which calendar information is used to represent recurring patterns in occupancy, equipment use, and indoor environmental response[6,7,20,22–24].

Solar-exposure variables. To represent room-specific solar loading, solar-radiation variables from DWD and Open-Meteo were first harmonized to a common irradiance basis consisting of global horizontal irradiance, I_{GHI}, diffuse horizontal irradiance, I_{DHI}, and direct normal irradiance, I_{DNI}. The solar zenith angle, θ_s, and solar azimuth, γ_s, were evaluated at the midpoint of each interval. For DWD observations, the provided zenith angle was used directly; for forecast records, solar position was derived from timestamp and site location.

For a surface with tilt β and azimuth γ_w, the angle of incidence, θ_i, was calculated as

$$\cos(\theta_i) = \cos(\theta_s)\cos(\beta) + \sin(\theta_s)\sin(\beta)\cos(\gamma_s - \gamma_w). \tag{5}$$

For vertical windows ($\beta = 90^{\circ}$), this reduces to

$$\cos(\theta_i) = \sin(\theta_s)\cos(\gamma_s - \gamma_w). \tag{6}$$

The beam-exposure factor, F_{beam}, was defined as

$$F_{\mathrm{beam}} = \begin{cases} \cos(\theta_i), & \text{if } \theta_s < 90^{\circ} \text{ and } \theta_i < 90^{\circ}, \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

When direct normal irradiance was not available directly, it was estimated from the horizontal beam component as

$$I_{\mathrm{DNI}} = \begin{cases} \max\left(0, \dfrac{I_{\mathrm{GHI}} - I_{\mathrm{DHI}}}{\cos(\theta_s)}\right), & \text{if } \theta_s < 90^{\circ}, \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

The effective solar-impact feature was then defined as

$$I_{\mathrm{eff}} = I_{\mathrm{DNI}} F_{\mathrm{beam}} + I_{\mathrm{DHI}} \frac{1 + \cos(\beta)}{2}. \tag{9}$$

For vertical windows ($\beta = 90^{\circ}$), this simplifies to

$$I_{\mathrm{eff}} = I_{\mathrm{DNI}} F_{\mathrm{beam}} + \frac{1}{2} I_{\mathrm{DHI}}. \tag{10}$$

Outdoor weather-memory and change variables. To provide the model with short-term meteorological context, rolling averages were computed for outdoor variables over a 6 h window. For a generic outdoor variable X_t, the rolling mean was defined as

$$\bar{X}_t = \frac{1}{N} \sum_{i=0}^{N-1} X_{t-i}, \tag{11}$$

where N denotes the number of samples within the 6 h window.

In addition, 1 h first-difference features were computed for outdoor temperature and relative humidity to capture the direction and short-term rate of change in meteorological forcing:

$$\Delta T_t^{\mathrm{out}} = T_t^{\mathrm{out}} - T_{t-1\mathrm{h}}^{\mathrm{out}}, \tag{12}$$

$$\Delta RH_t^{\mathrm{out}} = RH_t^{\mathrm{out}} - RH_{t-1\mathrm{h}}^{\mathrm{out}}. \tag{13}$$

These features were introduced to represent delayed indoor response to recent outdoor forcing. Whereas the rolling averages describe the recent exposure history, the first-difference terms capture whether outdoor conditions are increasing or decreasing over the last hour. This formulation is consistent with prior building time-series studies that combine weather variables with historical information to account for inertia and lagged system behavior[5–7,12,23].

Working-day indicators To account for occupancy-related differences between regular working days and reduced-activity periods, calendar indicators were derived from German public holidays and predefined university closure periods. Let D denote the set of public holidays and institutional closure dates. The holiday indicator, H_t, and the working-day indicator, W_t, were defined as follows:

$$H_t = \begin{cases} 1, & \text{if date}(t) \in D, \\ 0, & \text{otherwise}, \end{cases} \tag{14}$$

$$W_t = \mathbb{1}(\text{dayofweek}(t) < 5) \wedge (H_t = 0). \tag{15}$$

This feature design is consistent with earlier building-prediction work showing that time and calendar variables can serve as proxies for occupancy schedules and related internal loads[6,7,20,25,26].

Working-day occupancy proxy analysis

To estimate usage-related effects on the indoor climate without relying on direct occupancy sensing, a calendar-based occupancy proxy analysis was performed for each room. Following the common use of calendar information as a surrogate for building-use patterns, days were classified as working days or non-working days based on weekday, German public holidays, and a predefined winter closure period from 22 December 2025 to 4 January 2026[6,7,20,25] All timestamps were converted to the local time zone (Europe/Berlin), and the hour of day and day of week were extracted from the time index.

Two complementary visual analyses were used. First, hourly distributions of the selected indoor variable were compared between working and non-working days using boxplots. These plots were used to assess whether the central tendency and spread of the variable changed systematically between regular operation and reduced-activity periods.

Second, a distribution-based distance analysis was carried out using the first Wasserstein distance. To reduce the influence of extreme values, outliers were removed using the 1tile range rule. For each hour of the day, the non-working-day observations were taken as the baseline distribution, representing an approximate low-activity room state. The first Wasserstein distance W_1 was then calculated between this hourly baseline distribution and the corresponding distribution for each day-of-week and hour combination, as defined in Equation 16. The resulting distances were arranged as a day-by-hour heatmap, where larger values indicate a stronger shift away from the non-working baseline and thus a larger usage-related perturbation of the indoor climate.

$$W_1 := W_1(F_{\text{day,hour}}, F_{\text{non-working,hour}}) = \int_0^1 \left| F^{-1}_{\text{day,hour}}(u) - F^{-1}_{\text{non-working,hour}}(u) \right| du. \tag{16}$$

Forecasting model development and blocked temporal validation

To test whether the engineered exogenous feature set was physically informative, a deterministic outdoor-only baseline model was first developed using observed outdoor weather data from DWD together with measured indoor relative humidity. This baseline was intended as a feature-validation step under known boundary conditions rather than as the final deployment model. The target variable was indoor relative humidity in Lab C, as it is the laboratory used for electrospinning, while the predictor set consisted exclusively of engineered outdoor, temporal, solar-geometry, and calendar-related features. No lagged indoor measurements or other indoor history terms were included in this baseline configuration. After feature generation, rows with missing values were removed.

Model development was performed using XGBoost regressor. Temporal generalization performance was assessed using a blocked temporal validation strategy. For each available calendar month, all samples belonging to that month were withheld as an external test block, while the remaining months formed the development pool. This design avoided evaluating the model on randomly mixed neighboring samples from the same period and instead tested whether a model trained on other parts of the year could reproduce indoor humidity dynamics in an unseen monthly regime. This type of time-aware evaluation is consistent with prior indoor-climate forecasting practice, where the newest portion of the time series is commonly reserved for testing so that performance reflects prediction on unseen future data rather than on randomly shuffled observations[5,8,20,24].

Within the development pool, a random 90:10 split was used to separate training and internal evaluation data for early stopping only. Thus, randomness was used only inside the training pool, whereas the final assessment remained strictly month-blocked.

Model performance for each withheld month was quantified using the mean absolute error (MAE) and the coefficient of determination (R^2). In this way, the baseline quantified how well indoor humidity could be inferred from observed outdoor forcing and calendar-dependent structure alone, without access to lagged indoor-state variables. This formulation was retained deliberately as a conservative and low-barrier modeling choice, because it isolates the informativeness of exogenous weather and calendar features. Similar outdoor and calendar predictors have been shown to be informative for indoor temperature and humidity forecasting in prior work, typically in combination with indoor history inputs rather than as stand-alone exogenous models[12,21–23].

Probabilistic forecasting and interval calibration

After establishing the feature-validity baseline under observed outdoor forcing, a forecast-driven prediction model was developed to represent the intended operational decision-support setting. Unlike the baseline experiment, this stage used forecast weather inputs from Open-Meteo so that the full forecast-to-warning pipeline could be evaluated under realistic deployment conditions. The model was designed to estimate both expected indoor relative humidity and the uncertainty of the resulting forecasts.

For point prediction, indoor relative humidity was modeled with XGBoost at each forecast time step. The same cleaned laboratory dataset used in the deterministic analysis was retained, restricted to physically plausible values, aggregated to 10 min resolution, and enriched with the engineered forecasting predictors described above. Indoor relative humidity served as the target variable, and prediction was based on a predefined feature set.

Hyperparameters were optimized with Optuna. The search space comprised the learning rate, tree depth, L1 and L2 regularization, subsampling ratio, column subsampling ratio, and minimum child weight. To preserve temporal ordering, model selection used a 3-fold time series split on the training pool and minimized the mean validation score across folds. This follows the general principle that building time-series models should be evaluated on unseen future segments rather than on randomly shuffled samples.

For each trial, the XGBoost model was trained with early stopping and evaluated on validation-fold predictive accuracy. After completion of the Optuna study, the model was retrained with the best-performing hyperparameter set and saved for downstream forecasting.

Uncertainty quantification was not based on in-sample residuals or model-internal interval estimates. Instead, prediction intervals were constructed empirically from archived past forecasts and their realized indoor humidity outcomes, following the out-of-sample forecast-error approach of Lee and Scholtes[27]. For each forecast horizon $h \in \{0, \ldots, 7\}$, out-of-sample forecast errors were defined as

$$e_{t,h} = y_{t+h} - \hat{y}_{t+h|t}. \tag{17}$$

For a nominal central coverage level $1-\alpha$, the empirical prediction interval was then given by

$$I_{t,h}^{1-\alpha} = \left[\hat{y}_{t+h|t} + Q_{\alpha/2}(e_{t,h}),\ \hat{y}_{t+h|t} + Q_{1-\alpha/2}(e_{t,h})\right], \tag{18}$$

with quantiles estimated separately for each lead day.

Multi-day forecasts were generated for lead days 0 to 7, together with empirical prediction intervals at central coverage levels of 50%, 90%, 95%, and 99%. Because the forecast-error distribution was estimated separately for each lead day, interval width increased naturally with forecast horizon and reflected the observed deterioration in predictive skill at longer lead times.[27]

Model performance was assessed in two complementary ways. Point forecast accuracy was evaluated using MAE and R^2, while interval quality was assessed using empirical coverage and average interval width for each nominal confidence level. This separation between the observed-weather baseline and the forecast-driven uncertainty model made it possible to distinguish the contribution of feature design from the additional uncertainty introduced by operational weather forecasts.

Benchmark models

To evaluate the proposed forecasting framework independently of any single learner, three model classes were implemented within the same methodological pipeline: multiple linear regression, random forest regression and XGBoost. All models used the same predictor set, forecast targets, lead-day definitions, and blocked temporal train, calibration, and evaluation periods, such that performance differences could be attributed to model structure rather than differences in preprocessing, tuning effort, or data handling.

Random forest and XGBoost were treated as alternative non-linear formulations of the same forecasting task. Hyperparameters for the non-linear models were optimized with Optuna under the same tuning framework, while multiple linear regression was included as an interpretable linear baseline. XGBoost served as the initial implementation of the forecasting framework and was subsequently compared against random forest, and multiple linear regression under the same evaluation protocol.

User-facing DSS output

To support day-ahead planning for humidity sensitive experiments, the user-facing DSS was implemented as a model-agnostic visualization layer that converts the hourly indoor humidity forecasts into actionable graph containing lead-time-specific prediction intervals as information for risk assessment. Its main purpose is not automated control, but advance decision support based on the expected timing, magnitude, and uncertainty of upcoming indoor RH changes.

For each issuance time, the DSS assembles a continuous forecast from day 0 through day 7 and applies prediction intervals separately for each lead day. This per-day calibration step accounts for the fact that forecast error is not constant across the horizon and typically increases with temporal distance from the forecast origin. The displayed uncertainty bands therefore expand in a structured, horizon-aware manner rather than using a single fixed interval width across the full 7-day period.

The interface visualizes the predicted indoor RH trajectory as the central forecast line and overlays nested calibrated intervals, such as 50%, 90%, 95%, and 99%, to communicate increasing levels of uncertainty. In addition, the corresponding outdoor humidity signal can be displayed as contextual information, allowing the operator to interpret predicted indoor RH changes in relation to approaching weather conditions. User defined boundaries for relative humidity can be displayed as dashed horizontal lines to help gauge, when predicted humidity leaves a safe zone.

The practical output of the DSS is therefore a forecast-centered warning view in which operators can identify when an RH excursion is most likely, how uncertain that forecast is on each day ahead, and how much confidence should be placed in preventive planning decisions.

References

1. Teo, W. E. & Ramakrishna, S. A review on electrospinning design and nanofibre assemblies. *Nanotechnology* **17**, R89, DOI: 10.1088/0957-4484/17/14/R01 (2006).
2. Mailley, D., Hébraud, A. & Schlatter, G. A review on the impact of humidity during electrospinning: From the nanofiber structure engineering to the applications. *Macromol. Mater. Eng.* **306**, 2100115, DOI: https://doi.org/10.1002/mame.202100115 (2021). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mame.202100115.
3. Szewczyk, P. K. & Stachewicz, U. The impact of relative humidity on electrospun polymer fibers: From structural changes to fiber morphology. *Adv. Colloid Interface Sci.* **286**, 102315, DOI: https://doi.org/10.1016/j.cis.2020.102315 (2020).
4. Paulino Mendes, A. S., Rutgersson, A. & Paulsson, M. Outdoor environmental effects on cleanrooms – A study from a Swedish hospital pharmacy compounding unit. *Eur. journal pharmaceutics biopharmaceutics* **177**, 100–106 (2022).
5. Boesgaard, C. *et al.* Prediction of the indoor climate in cultural heritage buildings through machine learning: first results from two field tests. *Herit. Sci.* **10**, 176, DOI: https://doi.org/10.1186/s40494-022-00805-3 (2022).
6. Zhao, Y. *et al.* Prediction of daily home indoor temperature and relative humidity using a deep ensemble machine learning approach. *Build. Environ.* **295**, 114392, DOI: https://doi.org/10.1016/j.buildenv.2026.114392 (2026).
7. Wang, H., Chen, Y., Kang, J., Ding, Z. & Zhu, H. An XGBoost-based predictive control strategy for HVAC systems in providing day-ahead demand response. *Build. Environ.* **238**, 110350, DOI: https://doi.org/10.1016/j.buildenv.2023.110350 (2023).
8. Afroz, Z., Urmee, T., Shafiullah, G. M. & Higgins, G. Real-time prediction model for indoor temperature in a commercial building. *Appl. Energy* **231**, 29–53, DOI: https://doi.org/10.1016/j.apenergy.2018.09.052 (2018).
9. Tang, J. *et al.* Door-opening motion can potentially lead to a transient breakdown in negative-pressure isolation conditions: The importance of vorticity and buoyancy airflows. *The J. Hosp. Infect.* **61**, 283–286, DOI: 10.1016/j.jhin.2005.05.017 (2005).
10. Bhattacharya, A., Ghahramani, A. & Mousavi, E. The effect of door opening on air-mixing in a positively pressurized room: Implications for operating room air management during the COVID outbreak. *J. Build. Eng.* **44**, 102900, DOI: 10.1016/j.jobe.2021.102900 (2021). 40477595.
11. Open-Meteo. Previous model runs api (2024). Accessed 2026-04-01.
12. Wang, S. *et al.* A real-time indoor temperature and relative humidity prediction method for exhibition hall based on long short-term memory network. *J. Build. Eng.* **98**, 111492, DOI: 10.1016/j.jobe.2024.111492 (2024).
13. Psomas, T., Teli, D., Langer, S., Wahlgren, P. & Wargocki, P. Indoor humidity of dwellings and association with building characteristics, behaviors and health in a northern climate. *Build. Environ.* **198**, 107885, DOI: 10.1016/j.buildenv.2021.107885 (2021).
14. ifo Institute. Friday and monday most common days for working from home in germany (2023).
15. Zippenfenig, P. Open-meteo.com weather api [computer software], DOI: 10.5281/zenodo.7970649 (2023).
16. DWD Climate Data Center. 10-minute station observations of air temperature for germany, station 02014, recent data. https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/10_minutes/air_temperature/recent/10minutenwerte_TU_02014_akt.zip (2026). Accessed 2026-04-01.
17. DWD Climate Data Center. 10-minute station observations of solar radiation for germany, station 02014, recent data. https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/10_minutes/solar/recent/10minutenwerte_SOLAR_02014_akt.zip (2026). Accessed 2026-04-01.
18. Tetens, O. Über einige meteorologische begriffe. *Zeitschrift für Geophys.* **6**, 297–309 (1930).

19. Lawrence, M. G. The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications. *Bull. Am. Meteorol. Soc.* **86**, 225–234, DOI: 10.1175/BAMS-86-2-225 (2005).

20. Hampo, C. C., Schinasi, L. H. & Hoque, S. Indoor temperature and humidity prediction in residential buildings: A review of the white box and black box modeling techniques. *Build. Environ.* **290**, 114125, DOI: 10.1016/j.buildenv.2025.114125 (2026).

21. Qadeer, K., Ahmad, A., Qyyum, M. A., Nizami, A.-S. & Lee, M. Developing machine learning models for relative humidity prediction in air-based energy systems and environmental management applications. *J. Environ. Manag.* **292**, 112736, DOI: 10.1016/j.jenvman.2021.112736 (2021).

22. Özbalta, T. G., Sezer, A. & Yıldız, Y. Models for prediction of daily mean indoor temperature and relative humidity: Education building in izmir, turkey. *Indoor Built Environ.* **21**, 772–781, DOI: 10.1177/1420326X11422163 (2012).

23. Mba, L., Meukam, P. & Kemajou, A. Application of artificial neural network for predicting hourly indoor air temperature and relative humidity in modern building in humid region. *Energy Build.* **121**, 32–42, DOI: 10.1016/j.enbuild.2016.03.046 (2016).

24. Alawadi, S. *et al.* A comparison of machine learning algorithms for forecasting indoor temperature in smart buildings. *Energy Syst.* **11**, 937–970, DOI: 10.1007/s12667-020-00376-x (2020).

25. Jiang, X., Dawson-Haggerty, S., Ortiz, J. & Culler, D. Sensing by proxy: Occupancy detection based on indoor CO2 concentration (2011).

26. Tien, P. W., Wei, S., Darkwa, J., Wood, C. & Calautit, J. K. Machine learning and deep learning methods for enhancing building energy efficiency and indoor environmental quality – a review. *Energy AI* **10**, 100198, DOI: 10.1016/j.egyai.2022.100198 (2022).

27. Lee, Y. S. & Scholtes, S. Empirical prediction intervals revisited. *Int. J. Forecast.* **30**, 217–234, DOI: 10.1016/j.ijforecast.2013.07.018 (2014).

Author contributions statement

L.T. conceived the study, developed the methodology, conducted the experiments, analyzed the results, and drafted the manuscript. J.F.D. contributed to discussion of the electrospinning experiments and revised the manuscript. M.H. revised the methodology section and reviewed the manuscript. M.M. supervised the work, critically revised the manuscript, and approved the final version. All authors reviewed and approved the manuscript.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Additional information

Accession codes: Not applicable. Competing interests: The authors declare no competing interests.