Neural network-based microscale weather prediction using IoT sensor data in urban areas

Svetlana SEGARCEANU,
Beia Consult International, Bucharest, Romania
svetlana.segarceanu@beia.ro

Maria NICULAE,
Beia Consult International, Bucharest, Romania
maria.niculae@beia.ro

George SUCIU,
Beia Consult International, Bucharest, Romania
george@beia.ro

Mari-Anais SACHIAN,
Beia Consult International, Bucharest, Romania
anais.sachian@beia.ro

Abstract
Accurate weather prediction at the urban microscale is crucial for risk prevention, emergency response, and informed citizen decision-making. This study aims to improve short-term weather forecasting in Bucharest using Internet of Things (IoT) sensor data and neural network models, with a focus on predicting temperature, wind speed, and wind direction, parameters that are key to managing local environmental challenges such as air pollution dispersion and fire propagation. Previous research on urban weather forecasting has primarily relied on mesoscale models and limited sensor coverage. Recent studies have demonstrated that IoT-based approaches, coupled with machine learning, can enhance spatial and temporal resolution. Building on these findings, our work integrates local sensor networks and pollution monitoring platforms to deliver improved microscale predictions. We collected and analyzed time-series data from multiple IoT-based sensor platforms located in Bucharest and surrounding areas, including the Airly network, which provides both meteorological and air quality parameters (e.g., PM10 and PM2.5). Several neural network architectures were implemented and compared to evaluating their predictive accuracy. The proposed models demonstrated promising performance, particularly in short-term wind direction prediction, enabling more accurate assessment of pollutant dispersion and fire spread dynamics. The outcomes support data-driven decision-making for urban planners, environmental agencies, and emergency management organizations. This research contributes to a scalable approach to integrating IoT data and neural network-based modeling for fine-grained urban weather prediction, offering an adaptable framework applicable to other metropolitan areas.

Keywords: atmospheric monitoring, data-driven modeling, pollution dispersion, smart cities, environmental risk assessment.

1. Introduction
Weather events have a huge impact on several spheres of development of our planet. In particular, they have a strong effect in urban areas. Forecasting urban phenomena is contained in the set of tactics that municipalities and management organizations should use to prepare for and take appropriate measures in an opportune manner. In this respect weather forecast models are essential for accurate weather forecasting.

In the context of rapid urbanization, the need for accurate weather forecasts at microscale level has become increasingly critical. Urban environments generate distinct microclimates influenced by building density, surface materials, traffic emissions, and the distribution of green areas, which can lead to significant meteorological variations over very short distances. Phenomena such as the urban heat island effect alter local air circulation patterns and directly impact pollutant dispersion, energy consumption, and human comfort. As a result, traditional mesoscale forecasting models often struggle to capture the fine-grained dynamics that characterize modern metropolitan areas.

There are three main types of weather forecast models: global models that cover the entire planet, mesoscale models that focus on specific regions, and microscale models that are centered on individual cities or neighborhoods. Different forecasting models may produce varying outputs due to differences in their underlying algorithms, spatial and temporal resolutions, and initial conditions. From the perspective of prediction horizons, short-range forecasts typically cover the interval from 12 to 48 hours in advance. Medium-range forecasts (MRF) extend up to approximately 132 hours, or 3 to 7 days, while long-range models provide probabilistic estimates over extended time periods.

Microscale and mesoscale models are used for applications requiring high-resolution simulations, in urban planning, or pollutant dispersion analysis. Mesoscale models provide insight of the regional context, while microscale models handle detailed local features. Integrating these models produces a more accurate understanding of the large-scale atmospheric conditions and their effects. One important application resides in wind energy assessment [1]. Regional wind maps are generated and combined with local wind resource information at a specific wind farm location. Short-term wind power forecasting aids in operational planning for wind farms.

At the same time, the emergence of dense Internet of Things (IoT) sensor networks has created new opportunities for capturing high-resolution environmental data in real time. Compared to conventional meteorological stations, IoT platforms offer significantly improved spatial coverage and can detect short-term atmospheric fluctuations that are often missed by coarser observational systems. This rich stream of localized data provides a valuable foundation for data-driven forecasting models capable of learning complex, nonlinear relationships characteristic of urban microclimates. Consequently, the integration of IoT-based measurements with modern machine learning techniques represents a promising direction for enhancing short-term weather prediction in densely populated urban environments Similar IoT-based approaches for fine-grained weather monitoring at household level have been reported in recent smart city studies, which highlight discrepancies between city-wide forecasts and local microclimatic conditions [2].

Another application concerns avalanche forecasting where site-specific and time-specific forecasts of an avalanche can get support through mesoscale modelling of various weather parameters using a high-end workstation [3].

In recent years, deep learning techniques are more frequently taken into account as a means of improving or supporting solutions in a range of fields such as financial, environmental research and preservation, and in weather forecasting. In this paper we will present a deep learning framework to predict some weather parameters like temperature, wind speed or wind direction, based on the measurements furnished by two stations located not far from each other. The paper is organized as follows: The next section presents a state of the art in meteorology and weather forecasting. Chapter 3 presents the method and the data we have employed. The last part of the paper makes interpretations of the obtained results and draws the conclusions and directions for future research.

2. State of the art

Two well-known global weather models are the Global Forecast System (GFS) and the European Center for Medium-Range Weather Forecasts (ECMWF) model.

Global and regional numerical weather prediction (NWP) models often struggle to represent urban-scale meteorological phenomena with sufficient accuracy, even when using high-resolution grids. For example, a recent study showed that applying machine-learning bias correction and downscaling to a regional forecast model improved near-surface temperature predictions at 100 m horizontal resolution over an urban area, significantly reducing errors compared to the base forecast [4]. This evidence highlights the need for downscaling strategies when forecasting weather in dense urban environments, where fine-scale spatial heterogeneity strongly influences local microclimate and heat / wind distribution.

Microscale meteorology or micrometeorology is the study of temporary atmospheric phenomena smaller than mesoscale, on areas stretching about 1 kilometer (0.6 mi) or less [5]. This branch of meteorology is usually grouped with mesoscale as "mesoscale and microscale meteorology" (MMM) and together rather study phenomena detail features than synoptic scale. Microscale meteorology describes the most important mixing and dilution processes in the atmosphere [6]. Important topics in microscale meteorology include heat transfer and gas exchange between soil, vegetation, and/or surface water and the atmosphere caused by near-ground turbulence.

Numerical Weather Prediction (NWP) modeling is the most widely used and accurate method for weather forecasting. NWP involves solving a set of mathematical equations that represent the fundamental laws of physics governing the atmosphere. It became affordable and produced realistic outcome with the advent of computer simulation in the 1950s.

WRF (Weather Research and Forecasting model) [7] is a mesoscale numerical weather prediction system designed for both atmospheric research and operational forecasting applications, developed at the National Center for Atmospheric Research (NCAR), Boulder, CO, U.S.A. It relies on two dynamical components: the Advanced Research WRF (ARW) and the NCEP Non-hydrostatic Mesoscale Model (NMM). Both involve the flux form of the Boussinesq, non-hydrostatic Navier–Stokes equations that resolve advection, viscosity, pressure gradients, and Coriolis acceleration. The WRF-ARW is the most widely used dynamical core, based on an Arakawa C-grid for spatial discretization and a Runge-Kutta time integration scheme.

Although WRF includes several urban canopy parameterizations, recent studies have shown that the accuracy of urban simulations strongly depends on the quality of underlying land-use and urban morphology data. Modern approaches integrate Local Climate Zone (LCZ) classifications or WUDAPT datasets to better capture building density, vegetation fraction, and surface roughness. A 2022 study demonstrated that using LCZ-driven urban parameters significantly improved near-surface temperature and wind predictions in dense metropolitan areas compared to default WRF urban schemes. These advances highlight the importance of harmonized urban morphological data for accurate microscale prediction in cities [8].

Due to its open-source nature, broad community support, and continuous development, the WRF-ARW has become the foundation in both research and operational meteorology.

Mesoscale numerical weather prediction models focus on specific regions such as entire nations and provide forecasts at a higher resolution compared to global models. These models are particularly useful for predicting localized weather phenomena and severe weather events.

Machine learning and AI techniques are integrated to an increasing extent into weather forecasting models. These technologies enable the extraction of patterns from huge amounts of weather data, improving forecast accuracy and helping identification of minute atmospheric conditions.

Climavision [9] uses cutting edge AI technology to create AI models like Horizon AI Subseasonal to Seasonal model, also improving the accuracy of NWP models Horizon AI Global, Point and HIRES. It brings together the power of a high resolution weather radar and satellite network, with advanced weather prediction modelling to close significant weather observation gaps and drastically improve forecast speed and accuracy.

At present the Internet of Things (IoT) is the fundamental idea behind smart cities' ability to predict the weather. Massive volumes of data about atmospheric conditions, temperature, humidity, and wind patterns are gathered by Internet of Things devices, such as weather stations, sensors on buildings, and even smartphone applications. The collected data is sent to a central hub for real-time processing and analysis. The gathered data are converted into precise weather forecasts with the aid of machine learning algorithms, offering vital information that helps people, organizations, and governmental bodies make educated decisions [10]. In parallel, recent work emphasizes the role of edge computing in reducing latency and improving the reliability of IoT-based services in smart cities, by bringing computation closer to the data sources [11].

Beyond weather prediction, IoT-based sensing architectures are increasingly deployed for fine-grained environmental monitoring in smart city applications, for example in museum environments where microclimatic conditions must be tightly controlled for cultural heritage preservation [12].

One strategy used to predict extreme urban weather events is statistical forecasting, such as time series analysis. Historical meteorological data analysis can reveal trends and patterns that can then be utilized to predict urban drought factors [10].

A time series is a collection of observations or data points on a single variable over time. Time series forecasting employs a variety of methods. These data-driven models have been shown to be extremely accurate in meteorological forecasting applications [13]. Model performance is usually evaluated using statistical metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), or Mean Absolute Percentage Error (MAPE).

When forecasting using time series, it is important to evaluate the data properties, such as trends, seasonality, and other aspects.

ARIMA (Autoregressive Integrated Moving Average) forecasting is a widely used method for predicting the future values in a time series based on its historical values. The model incorporates autoregressive, moving average, and differencing components to accurately depict the dependencies and patterns of the series [10].

The ARIMA model is adapted to the SARIMA model which delivers seasonal forecasts data by including three seasonal components.

In recent years, deep learning has made considerable contributions in manifold fields such as agriculture [14], healthcare [15], energy usage [16], and finance [17]. This development is also considered to devise new solutions for the time series prediction problem, as neural network models have the ability of processing large amounts of non-linear data and capture distant data dependencies. Nevertheless, most existing prediction models based on neural networks sometimes fail to notice critical issues.

One strategy used to predict urban extreme weather events is statistical forecasting, such as time series analysis. Historical meteorological data analysis can reveal trends and patterns that can then be utilized to predict urban drought factors.

Among the traditional methods used to solve time-series forecasting problems, Linear Regression, SVM, Decision Trees, can be mentioned [13]. More recently deep neural networks, especially RNNs and LSTMs, or neural networks implementing the attention mechanism have been steadfastly experimented and operated. Although RNN and LSTMs have been popular for time-series forecasting, they also come with their own sets of challenges. RNNs [18], are prone to issues like vanishing and exploding gradients for the case long sequences, making them less effective for capturing long-term. LSTMs, designed to mitigate some of these issues, are computationally expensive and still might require substantial parameter tuning for optimal performance [19]. The authors in [13] design a novel Time-Series Neural Network framework incorporating an original Time-Attention mechanism and a Kernel Filter.

Recent breakthroughs in deep learning have introduced Transformer-based architectures and neural operators capable of learning complex spatiotemporal atmospheric dynamics more efficiently than recurrent networks. Models such as Google DeepMind's GraphCast [20] and NVIDIA's FourCastNet [21] can generate accurate global forecasts at significantly lower computational cost than traditional NWP. A 2024 evaluation showed that Transformer-based models outperform LSTMs and CNNs in short-range prediction of temperature and wind, while neural operators such as FNO achieve competitive accuracy for high-resolution atmospheric fields. These architectures represent a promising direction for future urban-scale prediction frameworks.

3. Method and experimental results

The experiments relied on data furnished by two stations situated not far from each other, an Airly station and an ADCON station. The Airly station provides hourly measurements of various air quality and weather parameters. Air quality parameters include particulate matter (PM1, PM2.5, PM10), and gas concentrations like nitrogen dioxide (NO2), ozone (O3), sulfur dioxide (SO2), carbon monoxide (CO), nitrogen oxide (NO), and hydrogen sulfide (H2S). Weather monitoring concerns parameters like temperature, pressure, humidity, wind speed(m/s) and wind direction. The ADCON station provides more frequent (usually about each 15 minutes) measurement values for weather parameters temperature, pressure, precipitation, humidity, wind speed(km/h) and wind direction. Although the two stations are close to each other, the parameter values are slightly different, as can be seen in Fig 1.

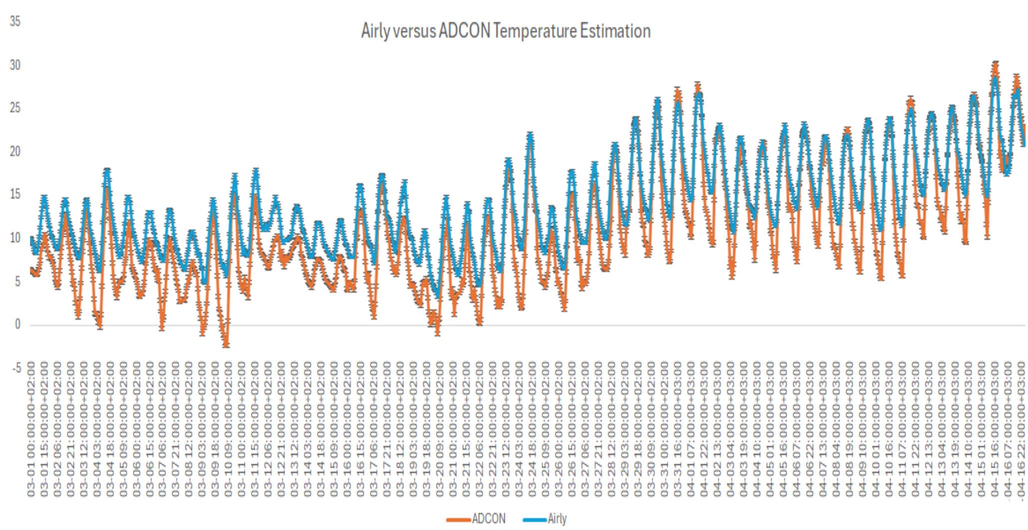

Fig. 1. Temperature values provided by Airly and ADCON stations.

In the first setting we employed the set of weather data furnished by the Airly station, with hourly estimations of temperature, pressure, humidity, wind speed(m/s) and wind direction, starting with the 2022-08-01, 00:00, till 2025-06-30 23:00.

The goal of the experiment was to predict the next day's temperature and wind speed, based on the values of the above mentioned parameters in the previous 11 days. We split the data into a training and a testing set, 772 days for training and 293 for testing.

The input vector, for a particular day in the above interval, contains 11 concatenated vectors corresponding to the 11 days on which the prediction relies. Each day is, at its turn, represented by hourly measurements of the 5 parameters mentioned above, for the 24 hours of the day, which means 120 components. Thus the input vector contains 1320 components.

We have evaluated the prediction of the next day for each hour of the day, using sheer Feedforward Neural Networks, with four layers, a number of neurons between 10 to 40 on the first level, up to 15 on the second and 5, respectively 4 on the remaining two layers.

We have assessed the quality of prediction using the estimations of MAE and RMSE. Table 1 presents the best estimations of Mean Absolute Error, for different neural network architecture, in the case of next day temperature prediction.

We have chosen the first configuration for the further tests. Table 2 presents estimations of the Mean Absolute Error and Root Mean Squared Error for tests related to next day temperature forecast at various hours of the day. One first conclusion is that the forecast for early or very late hours is more accurate than for the rest of the day.

Table 1. Mean Absolute Error estimations for different neural network architecture

Feed Forward Neural Network Architecture (Neurons on 4 layers)				MAE
27,	12,	5,	4	3.1834
29,	10,	5,	4	3.3524
15	11,	5,	4	3.5922

Table 3 presents the real values and the predicted values of temperature at 0:00 and 13:00 in several days of the set test. Examining these values one can infer that although the forecast is reasonable in most cases, sharp leaps of temperature, e.g. from 27, to 19 degrees, cannot be predicted only taking into account the values of the 5 parameters

Table 2. Mean Absolute Error and Root Mean Squared Error estimations for tests related to next day temperature forcast at various hours of the day

hour	MAE	RMSE	hour	MAE	RMSE
0	1.7467	2.272	14	3.8195	4.762
1	1.8212	2.335	15	3.7596	4.663
2	1.8652	2.395	16	3.9272	5.003
3	1.9367	2.481	17	4.0796	5.235
4	1.9539	2.570	18	3.5029	4.512
5	2.2050	2.864	19	3.4310	4.353

Table 3. Real values versus predicted values of temperature at 0:00 and 13:00 in several days

Predicted versus real temperature values at 0:00		Predicted versus real temperature values at 13:00	
Predicted value	Real value	Predicted value	Real value
17.6372	19.1400	20.4660	20.6300
17.0806	18.0300	25.964	24.6100
18.8574	18.0200	28.7792	27.8300
21.4221	18.7600	24.959	19.2600
17.3500	12.3500	21.3820	20.1100
15.2542	12.8600	21.6306	19.2300

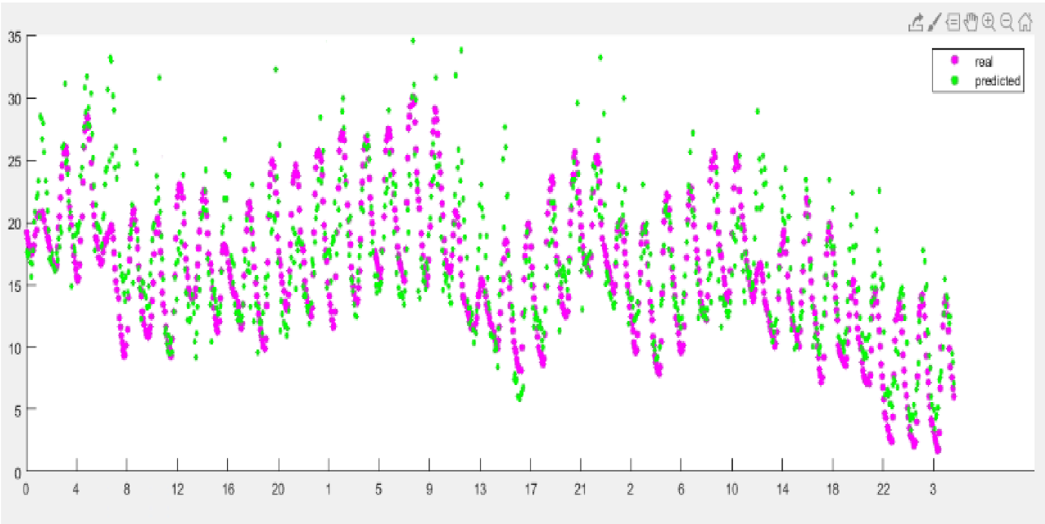

Fig. 2. The daily real temperature values against the predicted values for an interval starting with October 2024; values on the x axis are hours of the days, to emphasize the behaviour at different time slots

Figure 2 presents the daily temperature real values against the predicted values for an interval starting with October 2024. The values on the x axis represent the hours of the day of the temperature estimation.

We have also run some tests to forecast temperature the day after next day, at all the hours of the day, using the same Feedforward Network configuration. The Mean Absolute Error and Root Mean Squared Error for these tests were 4.020563 and 5.15. Table 4 presents estimations of the Mean Absolute Error for tests concerning temperature forecast the day after next day at various hours of the day. A similar conclusion can be drawn, the forecast for early or very late hours is more accurate than for the rest of the day. The errors are more substantial, because of more incorrect predictions.

Table 4. Estimations of Mean Absolute Error for tests related to temperature forecast the second next day at various hours of the day

hour	MAE	hour	MAE
0	3.1884	14	5.2475
1	2.7572	15	4.9237
2	2.7123	16	4.7244
3	3.233	17	4.4064
4	3.2031	18	4.359
5	3.4702	19	4.1125

Other range of experiments concerned the prediction of wind speed and wind direction. We used again the neural network architecture with 4 layers, 27 neurons on the first layer, 15 on the second, 5, 4 respectively on the last two layers.

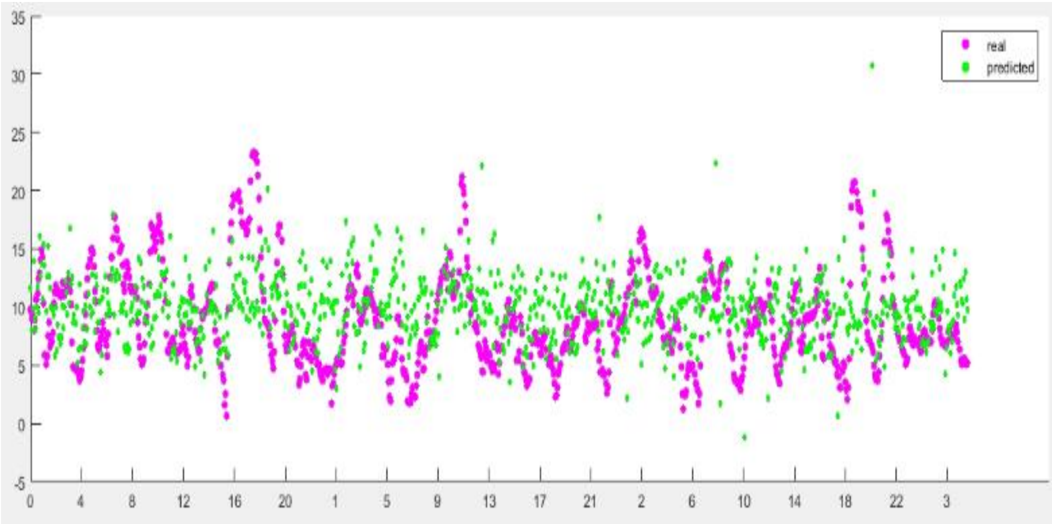

Fig. 3.The daily real wind speed values against the predicted values for an interval starting with October 2024

The results were poorer. The achieved Mean Absolute Error was 4.61, and the Root Mean Squared Error 5.82. Figure 3 shows the daily wind speed real values against the predicted values for an interval starting with October 2024.

The results for wind speed direction were also irrelevant with a Mean Absolute Error of 85.45, meaning an alteration with 85 degrees of the real wind direction by prediction.

4. Conclusions

The present paper presents our research concerning short -term weather parameters forecasting. The work is based on the weather data provided by two stations situated not far from each other, an Airly station and an ADCON station, containing 1065 days, out of which 770 was used for training. We have tested a Feed Forward Neural Network framework, using layers with a variable number of neurons on the first two layers and 5, respectively 4 on the last two layers. The prediction concerned the temperature, wind speed, and wind direction at certain hours (in fact all the hours) of the next day based on the values of five parameters (humidity, wind speed, wind direction, temperature, air pressure) in the previous 11 days. We have assessed the quality of the forecast using the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) measures and have chosen the network architecture with the best MAE value in the preliminary tests concerning temperature prediction.

There are some facts that came out. Temperature forecast at small hours (early morning or night) is more accurate than at noon or afternoon (MAE < 2). Although the accuracy seems not bad in many cases, the neural network approach could not forecast steep changes of temperature, which probably needs a more complex approach than time series of the values of the 5 parameters. The results for wind parameters are poorer.

There was no conclusion to be inferred concerning the architecture of the neural network, e.g. number of layers and neurons.

As future directions we intend to extend the volume of data to be included in the training and testing sets, use the data from both stations, or maybe one situated not in the immediate proximity of the two. We also intend to link the research to the air quality assessment and quality. We intend to use other popular deep neural network approaches. like LSTM or approaches that implement the attention mechanism.

Acknowledgements

This research was supported by the project "Romanian Hub for Artificial Intelligence - HRIA", Smart Growth, Digitization and Financial Instruments Program, MySMIS no. 351416 and by CHRISS project that has received funding from the HORIZON EUSPA-2021-SPACE research and innovation programme under grant agreement No. 101082440.

References

[1] S. E. Haupt, "Mesoscale to microscale coupling for wind energy applications: Addressing the challenges," in *Journal of Physics: Conference Series*, 2020.

[2] I. F. VOICU, "Precision long-term monitoring of local weather at household leve," in *Smart Cities International Conference (SCIC) Proceedings*, 2025.

[3] K. Srinivasan, A. Ganju and S. S. Sharma, "Usefulness of mesoscale weather forecast for avalanche forecasting," *Current Science*, pp. 921-926, 2005.

[4] L. P. Blunn, F. Ames, H. L. Croad, A. Gainford, I. Higgs, M. Lipson and C. H. B. Lo, "Machine learning bias correction and downscaling of urban heatwave temperature predictions from kilometre to hectometre scale," *Meteorological Applications*, 2024.

[5] T. Foken, "Micrometeorology," in *Applied Meteorology*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, pp. 315-326.

[6] Wang, K. Lawrence and P. C. Norman, Air and Noise Pollution Control: Volume 1, Springer Science & Business Media, 2012.

[7] J. G. Powers, J. B. Klemp, W. C. Skamarock, C. A. Davis, J. Dudhia and D. O. Gill, "The weather research and forecasting model: Overview, system efforts, and future directions," *Bulletin of the American Meteorological Society*, p. 1717–1737, 2017.

[8] M. Demuzere, J. Kittner, A. Martilli, G. Mills, C. Moede and I. D. Stewart, "A global map of Local Climate Zones to support earth system modelling and urban scale environmental science," *Earth System Science Data Discussions*, pp. 1-57, 2022.

[9] "Climavision," [Online]. Available: https://climavision.com/ .

[10] M. Todorov, N. Tihi, S. Popov and B. Stamatovic, "Time series models for weather forecasting in smart cities," in *International Scientific Conference ”ALFATECH–Smart Cities and modern technologies*, Belgrade, 2024.

[11] A. CRĂCEA, "The influence of edge computing on the development of smart cities," in *SCIC*, 2025.

[12] A. DANILA, R. A. STRECHE, F. E. OSIAC, A. M. STERE, C. DOBRE and G. SUCIU, "IoT systems for efficient preservation of cultural heritage," in *SCIC*, 2025.

[13] L. Zhang, R. Wang, Z. Li , Y. Ge, S. Wa, S. Huang and C. Lv, "Time-series neural network: a high-accuracy time-series forecasting method based on Kernel filter and time attention," *Information*, 2023.

[14] Y. Zhang, H. Wang, R. Xu, X. Yang, Y. Wang and Y. Liu, "High-precision seedling detection model based on multi-activation layer and depth-separable convolution using images acquired by drones," *Drones,* p. 152, 2022.

[15] Y. Zhang, "GenU-Net++: An Automatic Intracranial Brain Tumors Segmentation Algorithm on 3D Image Series with High Performance," *Symmetry,* 2021.

[16] M. R. Maarif, "Energy usage forecasting model based on long short-term memory (LSTM) and eXplainable artificial intelligence (XAI)," *Information,* 2023.

[17] H. Huo, "An accelerated method for protecting data privacy in financial scenarios based on linear operation," *Applied Sciences,* p. 1764, 2023.

[18] W. Zaremba, I. Sutskever and O. Vinyals, "Recurrent neural network regularization," arXiv preprint, 2014.

[19] A. Graves, "Long short-term memory," in *Supervised sequence labelling with recurrent neural networks*, 2012, pp. 37-45.

[20] R. Lam, "Learning skillful medium-range global weather forecasting," *Science,* pp. 1416-1421, 2023.

[21] J. Pathak, "Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators," arXiv preprint, 2022.