Barbara Vantaggi • Giulianella Coletti • Thierry Denoeux • Anne Laurent • Davide Petturiti • Enrique Miranda • Jesús Medina • Bernadette Bouchon-Meunier • Ronald R. Yager (Editors)

# IPMU 2026

15$^{th}$ – 19$^{th}$ June 2026, Rome, Italy

21$^{st}$ International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems

## Short Paper Proceedings


ROME TECHNOPOLE
INNOVATION ECOSYSTEM

SAPIENZA
UNIVERSITÀ DI ROMA

Thumbs up to IPMU's
40th anniversary

© 2026 The authors

Published by Associazione per la Matematica Applicata alla Rappresentazione dell'Incertezza per le Decisioni Economiche e Finanziarie.

**ISBN:** 979-12-243-3084-4

***Editors:*** Barbara Vantaggi, Giulianella Coletti, Thierry Denœux, Anne Laurent, Davide Petturiti, Enrique Miranda, Jesús Medina, Bernadette Bouchon-Meunier, Ronald R. Yager

The papers published in this proceedings are —unless indicated otherwise— covered by the Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 International (CC-BY-ND 3.0). You may copy, distribute, and transmit the work, provided that you attribute it (authorship, proceedings, publisher) in the manner specified by the author(s) or licensor(s), and that you do not use it for commercial purposes. The full text of the licence may be found at https://creativecommons.org/licenses/by-nc-nd/3.0/deed.en.

Credits: Cover image realized by Associazione per la Matematica Applicata alla Rappresentazione dell'Incertezza per le Decisioni Economiche e Finanziarie.

# Contents

# Preface

We are very pleased to present the proceedings of the Short Papers of the 21st International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2026). The conference was held on June 15–19, 2026, at the Faculty of Economics of Sapienza University of Rome in Rome, Italy.

The IPMU conference is held every two years and aims at bringing together scientists working on methods for the management of uncertainty and aggregation of information in intelligent systems. Since 1986, the IPMU conference has been providing a forum for the exchange of ideas between theoreticians and practitioners working in these areas and related fields.

In addition to many contributed scientific papers, the conference has attracted prominent plenary speakers, including Nobel Prize winners Kenneth Arrow, Daniel Kahneman, and Ilya Prigogine. A very important feature of the conference is the presentation of the Kampé de Fériet Award for outstanding contributions to the management of uncertain and imprecise information. Past winners of this prestigious award include Lotfi A. Zadeh (1992), Ilya Prigogine (1994), Toshiro Terano (1996), Kenneth Arrow (1998), Richard Jeffrey (2000), Arthur Dempster (2002), Janos Aczel (2004), Daniel Kahneman (2006), Enric Trillas (2008), James Bezdek (2010), Michio Sugeno (2012), Vladimir N. Vapnik (2014), Joseph Halpern (2016), Glenn Shafer (2018), Barbara Tversky (2020), Tomaso Poggio (2022), and Judea Pearl (2024). In this 2026 edition, the recipient of the Kampé de Fériet Award is Robert Kowalski (Imperial College London, United Kingdom) for his seminal work on human-oriented models of computing and computational models of human thinking. We warmly congratulate him.

The IPMU 2026 conference offered a rich and comprehensive scientific program. It featured four plenary talks delivered, in addition to Robert Kowalski, by Luc De Raedt (KU Leuven, Belgium and Örebro University, Sweden), María Ángeles Gil (University of Oviedo, Spain), and Marco Zaffalon (IDSIA, Switzerland).

The 2026 edition of IPMU celebrated the 40th anniversary of the first edition of the conference that took place in Paris, France. Rome was a particular fitting choice for this special occasion, as Rome and Paris have been exclusively and reciprocally twinned since April 1956.

The IPMU 2026 program consisted of 183 presentations divided into 22 special sessions and a general track, authored by researchers from more than 25 countries.

For IPMU 2026, the authors had the option to submit either regular papers or short papers. The present volume contains 51 accepted short papers. All accepted papers underwent a thorough review process and were evaluated by at least two reviewers. Furthermore, all papers were examined by the program chairs. The reviewing process respected the usual conflict-of-interest standards, so that all papers received multiple independent evaluations.

Organizing a conference would not be possible without the assistance, dedication, and support of many people and institutions, and IPMU 2026 is no exception. We are particularly grateful to the organizers of special sessions. Such sessions, dedicated to a variety of topics and organized by experts, have always been a characteristic feature of IPMU conferences.

We would like to acknowledge all members of the IPMU 2026 Program Committee, as well as the additional reviewers who played an essential role in the reviewing process, ensuring a high-quality conference. We thank them sincerely for their work and commitment. We gratefully acknowledge the technical co-sponsorship of the IEEE Computational Intelligence Society and the European Society for Fuzzy Logic and Technology (EUSFLAT).

We also acknowledge the support from Sapienza University of Rome, and in particular from the Faculty of Economics and the Department of Methods and Models for Economics, Territory and Finance; from the Rome Technopole Foundation; from the Springer team who managed the publication of these proceedings. We want to thank EUSFLAT for funding five student grants; the *International Journal of Approximate Reasoning* (– Elsevier) for funding a best paper award; Springer for funding a best paper award. Our very special and greatest gratitude goes to the authors who submitted the results of their work and presented them at the conference. Without them, this conference would not have been possible.

We hope that these proceedings provide readers with multiple ideas leading to numerous research activities, significant publications, and stimulating presentations at future IPMU conferences.

June, 2026
Rome, Italy

Barbara Vantaggi
Giulianella Coletti
Thierry Denœux
Anne Laurent
Davide Petturiti
Enrique Miranda
Jesús Medina
Bernadette Bouchon-Meunier
Ronald R. Yager

# Organization

## General Chair

| | |
|---|---|
| Barbara Vantaggi | Sapienza University of Rome, Italy |

## Program Committee Chairs

| | |
|---|---|
| Giulanella Coletti | University of Perugia, Italy |
| Thierry Denœux | Université de Tecnologie de Compiègne, France |
| Anne Laurent | LIRMM, Université de Montpellier, France |
| Davide Petturiti | Sapienza University of Rome, Italy |

## Publication Chair

| | |
|---|---|
| Enrique Miranda | University of Oviedo, Spain |

## Special Session Chair

| | |
|---|---|
| Jesús Medina | University of Cádiz, Spain |

## Social Media Chairs

| | |
|---|---|
| Andrea Cinfrignini | University of Siena, Italy |
| Silvia Lorenzini | University of Perugia, Italy |

## Advisory Board

| | |
|---|---|
| João Paulo Carvalho | Istituto Superior Tecnico, Portugal |
| Davide Ciucci | University of Milano-Bicocca, Italy |
| Giulianella Coletti | University of Perugia, Italy |
| Miguel Delgado | University of Granada, Spain |
| Mario Fedrizzi | University of Trento, Italy |
| Laurent Foulloy | Université de Savoie, France |
| Salvatore Greco | University of Catania, Italy |
| Julio Gutierrez-Rios | Polytechnic University of Madrid, Spain |
| Eyke Hüllermeier | Ludwig Maximilians University, Germany |
| Uzay Kaymak | Eindhoven University of Technology, The Netherlands |
| Anne Laurent | LIRMM, Université de Montpellier, France |
| Marie-Jeanne Lesot | Sorbonne University, France |
| Luis Magdalena | Polytechnic University of Madrid, Spain |
| Christophe Marsala | Sorbonne University, France |
| Benedetto Matarazzo | University of Catania, Italy |
| Jesús Medina Moreno | University of Cádiz, Spain |
| Manuel Ojeda-Aciego | University of Granada, Spain |
| Maria Rifqi | Université Panthéon-Assas, France |
| Lorenza Saitta | University of Piamonte, Italy |
| Olivier Strauss | LIRMM, Université de Montpellier, France |
| Enric Trillas | University of Spain |
| Llorenç Valverde | University of the Balearic Islands, Spain |
| José Luis Verdegay | University of Granada, Spain |
| Maria-Amparo Vila | University of Granada, Spain |

## Executive Directors

| | |
|---|---|
| Bernadette Meunier | Bouchon-Sorbonne Université-CNRS, France |
| Ronald Yager | Iona University, USA |

## Program Committee

| | |
|---|---|
| Rui Jorge Almeida | Maastricht University, The Netherlands |
| Nahla Ben Amor | Institut Superieur de Gestion de Tunis, Tunisia |
| Isabelle Bloch | Sorbonne Université, France |
| Ulrich Bodenhofer | University of Applied Sciences Upper Austria, Austria |
| Gloria Bordogna | CNR, Italy |
| Andrea Capotorti | University of Perugia, Italy |
| Christer Carlsson | Abo Akademi University, Finland |
| Joao Paulo Carvalho | Istituto Superior Tecnico, Portugal |
| Davide Ciucci | Universitá de Milano-Bicocca, Italy |
| Oscar Cordón | University of Granada, Spain |
| Maria Eugenia Cornejo Piñero | University of Cádiz, Spain |
| Keeley Crockett | Manchester Metropolitan University, United Kigndom |
| Bernard De Baets | Ghent University, Belgium |
| Marzieh Derakhshannia | University of Toulouse, France |
| Sébastien Destercke | Université de Technologie de Compiegne, France |
| Antonio Di Nola | University of Salerno, Italy |
| Scott Dick | University of Alberta, Canada |
| Ivana Dragovic | University of Belgrade, Serbia |
| Krzysztof Dyczkowski | Adam Mickiewicz University, Poland |
| Zied Elouedi | Institut Superieur de Gestion de Tunis, Tunisia |
| Juan Carlos Figueroa-García | Universidad Distrial Francisco José de Caldas, Colombia |
| Lluis Godo | Artificial Intelligence Research Institute, Spain |
| Fernando Gomide | University of Campinas, Brasil |
| Uzay Kaymak | Eindhoven University of Technology, The Netherlands |
| Laszlo T. Koczy | Budapest University of Technology and Economics, Hungary |
| Rudolf Kruse | University of Magdeburg, Germany |
| Marie-Jeanne Lesot | Sorbonne University, France |
| Christophe Marsala | Sorbonne University, France |
| Trevor Martin | University of Bristol, United Kingdom |
| Jerry Mendel | University of Southern California, USA |
| Radko Mesiar | University of Bratislava, Slovak Republic |
| Javier Montero | Complutense University, Spain |
| Ignacio Montes | University of Oviedo, Spain |
| Jacky Montmain | IMT Mines Ales, France |

| | |
|---|---|
| Serafín Moral | University of Granada, Spain |
| Amira Mouakher | University of Budapest, Hungary |
| Vilém Novák | University of Ostrava, Czech Republic |
| Hannu Nurmi | University of Turku, Finland |
| Dickson Owuor | Strathmore University, USA |
| David Pelta | University of Granada, Spain |
| Irina Perfilieva | University of Ostrava, Czech Republic |
| Frédéric Pichon | Université d'Artois, France |
| Olivier Pivert | IRISA-ENSSAT, France |
| Henri Prade | IRIT - CNRS |
| Raúl Pérez-Fernández | University of Oviedo, Spain |
| Barbara Pekala | University of Rzeszów, Poland |
| Benjamin Quost | University of Technology of Compiègne, France |
| Eloisa Ramírez Poussa | University of Cádiz, Spain |
| Marek Reformat | University of Alberta, Canada |
| Maria Rifqi | Université Panthéon-Assas, France |
| M. Dolores Ruiz | University of Cádiz, Spain |
| Thomas A. Runkler | Siemens AG, Germany |
| Daniel Sanchez | University of Granada, Spain |
| Giuseppe Sanfilippo | University of Palermo, Italy |
| Grégory Smits | IMT Atlantique, France |
| Martin Stepnicka | University of Ostrava, Czech Republic |
| Eulalia Szmidt | Polish Academy of Sciences, Poland |
| Andreja Tepavcevic | University of Novi Sad, Serbia |
| Matthias Troffaes | Durham University, United Kigndom |
| José Luis Verdegay | University of Granada, Spain |
| Paolo Vicig | University of Trieste, Italy |
| Susana Vieira | Universidade de Lisboa, Portugal |

## Additional Reviewers

Stefano Aguzzoli
Roberto G. Aragón
Sally Giuseppe Arcidiacono
Michal Baczynski
Paolo Baldi
David Barilla
Juan Baz
Akram Bensalem
Matteo Bianchi
Maria del Mar Bibiloni-Femenias
Henri Bollaert
Jana Borzová
Agustina Bouchet
Matteo Brunelli
Vittoria Bruni
Humberto Bustince
Marco Calautti
Antonio Calcagni
Andrés Cano
Andrea Campagner
Catrin Campbell-Moore
Giuseppe Caristi
Davide Cazzorla
Martine Ceberio
Fernando Chacón
William D. Chicaiza
Tiziana Ciano
Salvatore Corrente
Esther Anna Corsi
Anderson Cruz
Martina Dankova
Rocío de Andrés Calle
Wannes de Maeyer
Pasquale de Meo
Laura de Miguel
José Angel Díaz García
Francesco DiFatta
Anton Dolhopolov
Paweł Drygaś
Phuc Do
Jorge Elorza
Alexander Erreygers
Juan Manuel Escaño
Javier Fernández
Carlos Fernandez Basso
Raquel Fernández Peralta
Massimiliano Ferrara
Giuseppe Filippone
Tomasso Flaminio
Alberto Gaetano Valerio
Diego García Gil
Ramón García
Salvatore Greco
Przemyslaw Grzegorzewski
Zuzana Haniková
Piotr Helbin
Ondrej Hutník
Simon James
Gabriel Jaume Martin
Balasubramaniam Jayaram
Enol Junquera
Katarzyna Kaczmarek-Majer
Martin Kalina
Vladik Kreinovich
Konrad Kułakowksi
Gianmarco La Rosa
Dominique Laurent
David Lobo
Victor Lopes de Sousa
Salvador López Joya
Miguel López Pérez
Manuel Mancini
Brice Mayag
Jiri Mazurek
Arnau Mir Fuentes
Yacine Mokhtari
Samuel Molina Ruiz
Serafín Moral García
Andrea Morales Garzón
Kavit Nanavati
Serena Nicolazzo
David Nieto-Barba
Francisco José Ocaña
Manuel Ojeda-Hernández
Jonathan Okz
Bartolomé Ortiz
Bruno Antonio Pansera
José M. Pavía

Banu Pazar Varol
Graçaliz Pereira Dimuro
Boris Pérez Cañedo
Francisco Pérez Gámez
Sandro Preto
Nageswara Rao Vemuri
José María Rivas
Umberto Rivieccio
Jesús Rodríguez López
Maciej Romaniuk
Aleksandra Rutkowska
Mikel A. Serra-Moll
Luca Spada
Nik Stopar
Umberto Straccia
Marco Elio Tabacchi
Zdenko Takáč
Osamu Takahashi
Menelaos Tasiou
José Antonio Torné Zambrano
Leonardo Valera
Arthur Van Camp
Amanda Vidal
Binbin Xu
Nicolás Zumelzu

## Web Chair

Domenico Vitulano — Sapienza University of Rome, Italy

## Webmaster

Massimiliano Coppola — Sapienza University of Rome, Italy

# Plenary talks

# Medium-Range Temperature Forecasting: Neural Networks or a Linear Model?

Luca Vincenzo Ballestra

Department of Statistical Sciences, Alma Mater Studiorum – University of Bologna, Bologna, Italy
luca.ballestra@unibo.it

**Abstract.** This paper compares a seasonal vector autoregression, VAR(1), with four neural-network architectures for medium-range forecasting of daily mean temperature and mean sea-level pressure over horizons from 1 to 90 days. Forecast accuracy is evaluated with path-averaged multi-step MAE and RMSE. The neural alternatives are a direct multilayer perceptron, a recursive multilayer perceptron, a CNN–LSTM encoder–decoder, and an LSTM sequence-to-sequence model. All nonlinear specifications incorporate harmonic seasonality, and the direct multi-step architectures exploit the known future seasonal sequence at the forecast origin. The empirical evidence from daily Milan data is mixed: LSTM–Seq2Seq is strongest at the shortest horizon, at the 30-day temperature horizon, and for 5-day pressure forecasts; MLP-Rec attains the lowest 5-day temperature MAE and the lowest 90-day temperature errors; and VAR(1) is strongest for temperature at 10 days and for pressure from 10 days onward. No neural model dominates uniformly across variables and horizons, while VAR(1) is substantially faster. An additional LSTM–Seq2Seq experiment with shortwave radiation and wind speed yields selective improvements, mainly at short horizons for temperature and more modestly across several horizons for pressure.

**Keywords:** Temperature forecasting · vector autoregression · neural networks · LSTM · medium-range forecasting

## 1 Introduction

Accurate daily forecasts of near-surface atmospheric conditions matter for energy systems, agriculture, infrastructure, and weather-related risk management. Although numerical weather prediction dominates the very short run, forecast skill deteriorates with lead time because atmospheric dynamics are chaotic [1, 4]. This makes medium-range forecasting a natural setting in which to compare parsimonious statistical baselines with more flexible machine-learning models.

The paper asks whether nonlinear neural architectures yield gains large and robust enough to justify their added complexity relative to a strong seasonal linear benchmark on a single-site daily forecasting problem. This is informative because seasonality and persistence already explain a substantial part of the predictable variation, so a careful benchmark remains useful even when the evidence is mixed.

Accordingly, we compare a seasonal VAR(1) with four neural-network architectures for forecasting daily mean temperature and mean sea-level pressure in Milan.

*B. Vantaggi, G. Coletti, T. Denœux, A. Laurent, D. Petturiti, E. Miranda, J. Medina, B. Bouchon-Meunier, R. R. Yager (eds.)*
*Proceedings of the IPMU2026 Short Papers. June 2026, Rome, Italy*, p. 337–342

The contribution is a controlled benchmark that emphasizes forecast accuracy, computational cost, and whether any nonlinear gains remain meaningful once both are considered. A brief additional experiment tests whether supplementing the strongest recurrent architecture with shortwave radiation and wind speed improves the forecasts.

The main message is cautious. Neural gains are selective rather than uniform: LSTM–Seq2Seq is strongest at the shortest horizon, at the 30-day temperature horizon, and for 5-day pressure forecasts; MLP-Rec is best for temperature MAE at 5 days and for both temperature metrics at 90 days; and VAR(1) becomes the strongest model for pressure from 10 days onward.

## 2 Methods

Let $T_t$ and $P_t$ denote daily mean temperature and mean sea-level pressure on day $t$. The bivariate target vector is $\mathbf{y}_t = (T_t, P_t)'$. Seasonality is represented by the harmonic vector

$$\mathbf{s}_t = \big(\sin(2\pi d_t/365),\, \cos(2\pi d_t/365)\big)',$$

where $d_t$ is the day of the year. All models use these seasonal features, but they differ in how past observations and known future seasonal covariates are mapped into multi-step forecasts. For the direct multi-step neural architectures, the future harmonic sequence $(\mathbf{s}_{t+1}, \ldots, \mathbf{s}_{t+H})$ is treated as deterministic information available at the forecast origin and supplied together with the historical input window. Background references on vector autoregressions include [2], [3].

*VAR(1).* The linear benchmark is the bivariate seasonal vector autoregression

$$\mathbf{y}_t = \boldsymbol{\mu} + A\mathbf{y}_{t-1} + C\mathbf{s}_t + \boldsymbol{\varepsilon}_t, \qquad \mathbf{y}_t = (T_t, P_t)'.$$

The model is estimated by ordinary least squares and then iterated recursively to obtain multi-step forecasts.

*Neural models.* The MLP-Dir uses the previous 30 days of temperature and pressure together with the horizon-specific future harmonic sequence $(\mathbf{s}_{t+1}, \ldots, \mathbf{s}_{t+H})$ to predict the entire forecast path $\widehat{\mathbf{y}}_{t+1:t+H|t}$ directly. The MLP-Rec uses only the latest lag ($L = 1$), is trained for one-step prediction, and is then iterated recursively; along the recursive path, the seasonal harmonics are updated day by day using the corresponding forecast date. The CNN–LSTM-ED and LSTM–Seq2Seq take a length-30 input sequence with four channels $(T, P, \sin, \cos)$ and produce the full $H$-step path in direct multi-step fashion, while their decoders receive the future harmonic sequence $(\mathbf{s}_{t+1}, \ldots, \mathbf{s}_{t+H})$ as known ex ante seasonal information. The CNN–LSTM-ED uses two causal 1D convolutional layers before decoding, whereas the LSTM–Seq2Seq is purely recurrent.

*Common training choices.* All neural networks were implemented in R through the Keras/TensorFlow interface and trained with Adam (learning rate $10^{-3}$), batch size 32, maximum 100 epochs, 10% validation split, and early stopping with patience 10. Input scaling used training-sample moments only, and all neural results are based on a fixed random seed equal to 123. The direct models were re-estimated separately for each horizon $H \in \{1, 5, 10, 30, 90\}$; the recursive VAR(1) and MLP-Rec were trained once and then iterated forward.

## 3 Data, evaluation, and baseline results

The data come from Open-Meteo (https://open-meteo.com) and cover Milan (Europe/Rome local time, CET/CEST) from 1 January 2000 to 1 January 2024. The sample was split chronologically into 80% training observations and 20% test observations. Supervised windows were built after this split. Training origins were therefore restricted so that both input lags and forecast targets lie within the training block. Test origins start on the first day of the test block; for the earliest test forecasts, lagged inputs may use observations from the end of the training block, which are known at the forecast origin and are therefore admissible. For each horizon $H$, only forecast origins with all future realizations observed were retained, so late incomplete test windows were dropped. Over the full sample, temperature has mean 13.2°C and standard deviation 8.08°C; pressure has mean 1015.85 hPa and standard deviation 7.56 hPa.

For each horizon H, test forecast paths were formed so that the first predicted day lies in the test block and all $H$ future realizations are observed. Hence the number of complete test paths is $M_H = N_{\text{test}} - H + 1$. Reported errors are path-averaged: for each $H$, they average forecast errors over all retained windows and all leads $h = 1, \ldots, H$. Consequently, the statistic reported for $H = 30$ is not the error at lead 30 alone, but the average error over the full path from lead 1 through lead 30. For a generic variable $Y \in \{T, P\}$,

$$\mathrm{MAE}_H(Y) = \frac{1}{M_H H} \sum_{t=1}^{M_H} \sum_{h=1}^{H} \left| \widehat{Y}_{t+h|t} - Y_{t+h} \right|,$$

$$\mathrm{RMSE}_H(Y) = \sqrt{\frac{1}{M_H H} \sum_{t=1}^{M_H} \sum_{h=1}^{H} \left( \widehat{Y}_{t+h|t} - Y_{t+h} \right)^2}.$$

These metrics summarize average performance over the entire forecast path up to horizon $H$ rather than isolated terminal-lead accuracy.

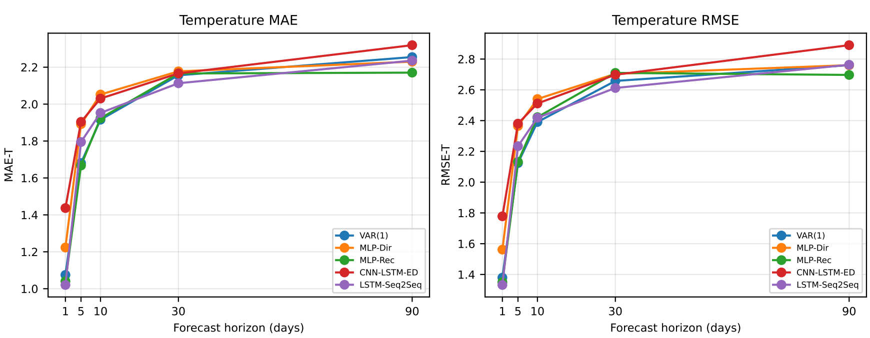


**Fig. 1.** Temperature MAE and RMSE over forecasting horizons for the five baseline models.

**Table 1.** Path-averaged forecast errors for temperature ($T$) and pressure ($P$). Panel A reports MAE and Panel B reports RMSE. Best value in each column is in bold.

*Panel A: MAE*

| | Temperature [°C] | | | | | Pressure [hPa] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Model | 1 | 5 | 10 | 30 | 90 | 1 | 5 | 10 | 30 | 90 |
| VAR(1) | 1.075 | 1.681 | **1.916** | 2.156 | 2.255 | 3.210 | 4.898 | **5.329** | **5.602** | **5.646** |
| MLP-Dir | 1.223 | 1.892 | 2.052 | 2.178 | 2.230 | 3.077 | 4.996 | 5.467 | 5.731 | 5.767 |
| MLP-Rec | 1.041 | **1.668** | 1.923 | 2.166 | **2.171** | 3.185 | 4.894 | 5.460 | 6.024 | 6.285 |
| CNN–LSTM-ED | 1.437 | 1.904 | 2.030 | 2.167 | 2.319 | 3.077 | 5.030 | 5.628 | 5.761 | 5.856 |
| LSTM–Seq2Seq | **1.020** | 1.795 | 1.952 | **2.113** | 2.236 | **2.855** | **4.853** | 5.446 | 5.782 | 5.808 |

*Panel B: RMSE*

| | Temperature [°C] | | | | | Pressure [hPa] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Model | 1 | 5 | 10 | 30 | 90 | 1 | 5 | 10 | 30 | 90 |
| VAR(1) | 1.379 | **2.124** | **2.391** | 2.658 | 2.760 | 4.351 | 6.670 | **7.151** | **7.461** | **7.521** |
| MLP-Dir | 1.562 | 2.365 | 2.540 | 2.704 | 2.760 | 4.170 | 6.748 | 7.297 | 7.610 | 7.666 |
| MLP-Rec | 1.350 | 2.132 | 2.422 | 2.710 | **2.697** | 4.326 | 6.691 | 7.343 | 7.967 | 8.265 |
| CNN–LSTM-ED | 1.778 | 2.380 | 2.511 | 2.698 | 2.890 | 4.182 | 6.787 | 7.517 | 7.657 | 7.776 |
| LSTM–Seq2Seq | **1.332** | 2.234 | 2.418 | **2.612** | 2.763 | **3.886** | **6.591** | 7.283 | 7.677 | 7.752 |

Table 1 and Fig. 1 for temperature show a stable qualitative picture across MAE and RMSE. At $H = 1$, LSTM–Seq2Seq attains the lowest error for both variables under both loss functions, although MLP-Rec remains close for temperature. At $H = 5$, the ranking is split by variable and metric: MLP-Rec yields the lowest temperature MAE, VAR(1) has the lowest temperature RMSE, and LSTM–Seq2Seq is best for pressure under both metrics. At $H = 10$, VAR(1) has the lowest temperature errors and is also best for pressure under both MAE and RMSE.

Over longer forecast paths, the gains of the neural models remain selective rather than broad. At $H = 30$, LSTM–Seq2Seq yields the lowest temperature MAE and RMSE, but its advantage over VAR(1) is modest in absolute magnitude (2.113 versus 2.156 in MAE, and 2.612 versus 2.658 in RMSE). At $H = 90$, MLP-Rec attains the lowest temperature errors (2.171 in MAE and 2.697 in RMSE), improving on both VAR(1) and the direct MLP for this variable. For pressure, however, the seasonal VAR(1) is the best model from $H = 10$ onward under both MAE and RMSE. This means that no neural architecture dominates uniformly across variables and horizons.

CNN–LSTM-ED is often the weakest specification, although its updated 10- and 30-day temperature errors are closer to the strongest alternatives than in the earlier run. Overall, nonlinear flexibility yields only selective gains on this dataset, whereas a parsimonious seasonal linear benchmark remains difficult to outperform robustly.

### 3.1 Computational cost

Table 2 reports training, test, and total runtimes for the five baseline models, aggregated over the five forecast horizons. All computations were performed on a laptop equipped with a 12th Gen Intel Core i5-1245U processor and 16 GB RAM.

VAR(1) is more than an order of magnitude faster than every neural alternative. Among the neural models, LSTM–Seq2Seq is the cheapest in total runtime, requiring about 33 seconds across the five horizons, versus less than half a second

**Table 2.** Training, test, and total runtimes (seconds) for the five baseline models, summed over $H \in \{1, 5, 10, 30, 90\}$.

| Model | TrainTime | TestTime | TotalTime |
|---|---|---|---|
| VAR(1) | 0.01 | 0.45 | 0.46 |
| MLP-Dir | 54.37 | 1.75 | 56.12 |
| MLP-Rec | 18.74 | 19.94 | 38.68 |
| CNN–LSTM-ED | 85.23 | 8.09 | 93.32 |
| LSTM–Seq2Seq | 16.01 | 16.84 | 32.85 |

**Table 3.** LSTM–Seq2Seq with additional encoder inputs (shortwave radiation and wind speed): path-averaged errors and runtimes.

| $H$ | $\text{MAE}_T$ | $\text{RMSE}_T$ | $\text{MAE}_P$ | $\text{RMSE}_P$ | Train | Test | Total |
|---|---|---|---|---|---|---|---|
| 1 | 1.012 | 1.328 | 2.837 | 3.787 | 186.42 | 4.58 | 191.00 |
| 5 | 1.752 | 2.193 | 4.806 | 6.534 | 232.73 | 4.33 | 237.06 |
| 10 | 2.098 | 2.592 | 5.473 | 7.267 | 203.22 | 5.34 | 208.56 |
| 30 | 2.159 | 2.681 | 5.710 | 7.635 | 159.73 | 5.28 | 165.01 |
| 90 | 2.348 | 2.901 | 5.788 | 7.736 | 243.39 | 4.09 | 247.48 |

for VAR(1). MLP-Rec is also relatively inexpensive because it is trained once and then iterated forward, but its recursive test-time cost is larger. Even where recurrent networks improve accuracy, those gains carry a substantial computational premium.

The evidence should nevertheless be interpreted cautiously. The benchmark uses a single location, a single chronological split, and path-averaged rather than lead-specific losses. These limitations narrow the scope of the conclusions.

## 4 LSTM–Seq2Seq with additional meteorological inputs

The additional-input experiment asks whether local meteorological variables improve forecasts beyond persistence and harmonic seasonality. We extend only LSTM–Seq2Seq because it remains the most attractive recurrent baseline: it is best at $H = 1$ for both variables, best for temperature at $H = 30$, best for pressure at $H = 5$, and the cheapest neural model in total runtime.

The extension augments the encoder input sequence with daily shortwave radiation and wind speed while leaving the forecast targets unchanged. Table 3 reports the resulting errors and runtimes. The gains are selective. For temperature, both MAE and RMSE improve at $H = 1$ and $H = 5$, but worsen from $H = 10$ onward. For pressure, the additional inputs improve RMSE at all horizons and MAE at most horizons, with modest gains.

The computational cost of the additional-input model is substantial. Summed over the five horizons, it requires 1049.11 seconds, compared with 32.85 seconds for the baseline LSTM–Seq2Seq in the updated baseline run. The extra-input specification is therefore best interpreted as a targeted robustness check rather than as a broadly superior model.

## 5 Conclusions

For daily data from Milan, the evidence does not support a uniformly superior nonlinear model for medium-range forecasting. Neural architectures can improve upon a seasonal linear benchmark, but the gains are selective, horizon-dependent, and variable-specific. LSTM–Seq2Seq is strongest at the shortest horizon and at the 30-day temperature horizon, as well as for 5-day pressure forecasts; MLP-Rec performs best for temperature MAE at 5 days and for both temperature metrics at 90 days; and VAR(1) remains the strongest model for pressure from 10 days onward.

The main implication is comparative. On this benchmark, a seasonal VAR(1) remains hard to beat once computational cost is taken into account, especially for pressure. The updated results give MLP-Rec a clearer advantage for long-horizon temperature forecasts, but this does not translate into broad dominance across variables and horizons. The additional-input extension of LSTM–Seq2Seq confirms the same message: supplementary meteorological information can help at very short horizons, but not enough to change the overall balance between accuracy and complexity.

Within the present setting, parsimonious seasonal linear dynamics remain a strong baseline for medium-range daily forecasting, and more complex nonlinear models must clear a high bar to justify their additional cost.

**Acknowledgements** We acknowledge the support of the project PRIN 2022 "Modeling and valuation of financial instruments for climate and energy risk mitigation", proposal code 2022FPLY97 - CUP J53D23004530006, funded by the European Union - NextGenerationEU and implemented by the Italian Ministry of University and Research (MUR) under the National Recovery and Resilience Plan (PNRR) - Mission 4 Education and research - Component 2 From research to business - Investment 1.1 Notice Prin 2022.

We also acknowledge the support of the project "Numerical methods for addressing traditional and emerging challenges in finance", funded by the Italian INDAM-GNCS - CUP E53C24001950001.

## References

1. Bauer, P., Thorpe, A., Brunet, G.: The quiet revolution of numerical weather prediction. Nature **525**, 47–55 (2015)
2. Hamilton, J.D.: Time Series Analysis. Princeton University Press, Princeton, NJ (1994)
3. Lütkepohl, H.: New Introduction to Multiple Time Series Analysis. Springer, Berlin, Heidelberg (2005)
4. Zhang, F., Sun, Y.Q., Magnusson, L., Buizza, R., Lin, S.J., Chen, J.H., Emanuel, K.: What is the predictability limit of midlatitude weather? Journal of the Atmospheric Sciences **76**(4), 1077–1091 (2019)