Preprints.org

Article

Deep Learning for Atmospheric Modeling: A Proof of Concept Using Fourier Neural Operator on WRF Data to Accelerate Transient Wind Forecasting at Multiple Altitudes

Paulo Alexandre Costa Rocha [*] , Jesse Van Griensven Thé , Victor Oliveira Santos , Bahram Gharabaghi

Article

Deep Learning for Atmospheric Modeling: A Proof of Concept Using Fourier Neural Operator on WRF Data to Accelerate Transient Wind Forecasting at Multiple Altitudes

Paulo Alexandre Costa Rocha [1,*], Jesse Van Griensven Thé [2], Victor Oliveira Santos [3] and Bahram Gharabaghi [3]

[1] Mechanical Engineering Department, Technology Center, Federal University of Ceará, Fortaleza, Ceará 60020-181, Brazil

[2] Lakes Environmental, 170 Columbia St. W, Waterloo, ON N2L 3L3, Canada

[3] School of Engineering, University of Guelph, 50 Stone Rd E, Guelph, ON N1G 2W1, Canada

[*] Correspondence: paulo.rocha@ufc.br

Abstract: This study addresses the problem of computational cost of transient CFD simulations, which rely on iterative time-step calculations, by employing deep learning to generate optimized initial conditions for accelerating the Weather Research and Forecasting (WRF) model. To this end, we forecasted wind speed for short time frames over Houston region using the WRF model data from 2018 to 2022, training the models to predict the wind speed X-component (U). The so-called global FNO model, trained across all atmospheric heights, was first tested, achieving competitive results. A more refined approach was tested to improve it further, training separate models for each altitude level, enhancing accuracy significantly. These ad hoc models outperformed surface and middle atmosphere persistence, achieving 27.64% and 20.46% nRMSE, respectively, while remaining competitive at higher altitudes. Variable selection played a key role, revealing that different physical processes dominate at various altitudes, necessitating distinct input features. The results highlight the potential of deep learning, particularly FNO, in atmospheric modeling, suggesting that tailored models for specific altitudes may enhance forecast accuracy. Thus, this study demonstrates that a deep learning model can be designed to start the iterations of a transient simulation, reducing convergence time and enabling faster, lower-cost predictions.

Keywords: deep learning; WRF; weather forecasting; computational fluid dynamics; artificial intelligence

1. Introduction

Nowadays, state-of-the-art meteorological computational models are important in decision-making, risk assessment, and investment planning, among several other benefits for modern society. They enable accurate weather forecasting, which aids daily activities, agriculture, and energy management [1,2]. These models can also predict severe weather events, saving lives and reducing economic losses [3,4]. In a broader sense, they are key to understanding climate change, enabling the analysis of trends, and permitting the development of strategies for adaptation and mitigation [5,6].

Meteorological models are also applied to support the monitoring of air quality [7,8], issue alerts for at-risk populations [9,10], and predict weather-influenced disease outbreaks [11,12]. They may enhance transportation safety, guide resilient infrastructure design, and improve emergency response and resource allocation. Furthermore, from an economic perspective, they support industries like insurance and logistics by assessing risks and optimizing operations [13,14].

Additionally, they drive advancements in atmospheric science, improving our understanding of weather and climate impacts [15–17].

The Weather Research and Forecasting (WRF) model is a extensively used meteorological computational tool that fits seamlessly into this context as a powerful framework for atmospheric research and operational forecasting [18]. Developed collaboratively by institutions such as the National Center for Atmospheric Research (NCAR) [19], WRF is designed to provide high-resolution and accurate weather predictions, making it invaluable across numerous applications, including the support of daily planning, agricultural optimization, and renewable energy management [20–23]. Its high spatial and temporal resolution makes it suitable for predicting localized severe weather events like hurricanes, tornadoes, and floods, thus playing a critical role in disaster preparedness and response [3,22].

In this sense, WRF serves as an essential tool in public health, where it also aids in air quality modeling by forecasting pollutant dispersion and supporting early warnings for vulnerable populations [24,25].

The Weather Research and Forecasting (WRF) model employs a Computational Fluid Dynamics (CFD) paradigm specifically developed for atmospheric processes. WRF solves the Navier-Stokes equations and additional equations for thermodynamics, moisture, and other atmospheric properties. These equations are solved numerically using advanced discretization techniques to simulate the atmosphere's behaviour.

General-purpose CFD models, such as OpenFOAM and ANSYS Fluent, are designed to simulate fluid flows in a variety of settings (e.g., pipes, turbines) [26–28], while WRF specializes in simulating fluid dynamics in the Earth's atmosphere. It incorporates physical processes like radiation, cloud microphysics, surface interactions, and boundary layer turbulence, which are critical for weather and climate modeling [29].

In a similar way as a transient CFD model works, WRF uses iterative processes to search for future values. Most of the processing time in WRF is iterating to achieve convergence of the solution. WRF starts the iterations using the latest value, which entails high computational costs to achieve convergence. To be more specific, when a CFD model starts the iterations of a new time step, it uses the variable fields of the last result to start the iterative calculation of the new values until convergence is achieved. In this context, employing a strategy that initiates iterations with values closer to the future value results in a significant reduction of computational resources in terms of time and economic costs.

Thus, this study aims to demonstrate, via a proof of concept, that artificial intelligence tools can be used to generate solution fields that approximate future hourly values more closely than simply using the current value. Machine learning and deep learning communities denotes the last variable field as the "persistence model". Thus, it can be seen that, if an improved field closer to the final result is applied to start the CFD iterations, the converged solution will be reached earlier. Using such fields to initiate temporal iterations makes it possible to achieve convergence in a significantly lower number of WRF convergence iterations, producing results in less time and with potentially significant computational savings. This initial guess commonly applied in CFD, which consists of using the latest fields to start the temporal iterations, is referred to throughout the whole manuscript as the "persistence" model.

2. Literature Review

Physics-Informed Neural Networks (PINNs) are a class of Deep Learning approaches that aim to merge physical laws into neural network architectures for solving partial differential equations (PDEs). Unlike traditional deep learning models, PINNs use mechanisms to induce predictions to adhere to governing equations, targeting accuracy and generalizability, particularly with limited data. PINNs provide an efficient alternative to computationally expensive numerical methods, making them suitable for inverse problems and high-dimensional PDEs [30]. Recent research reports PINNs' success in fluid mechanics, heat transfer, and materials science, working on improving their

capability to reconstruct velocity fields and model stress-strain relationships. PINN frameworks could even outperform experimental correlations and other data-driven methods, providing more reliable predictions in complex physical systems [31].

However, challenges such as spectral bias and optimization difficulties persist., Despite their good results, PINNs are commonly developed in *ad-hoc* situations, thus requiring a specifically developed architecture and weights for each problem. Strategies like adaptive loss weighting and hybrid approaches are being explored to enhance performance [32]. Under this ecosystem, the Neural Operators are developed to promptly solve classes of ODEs/PDEs. Pre-trained Neural Operators can also serve as surrogates to represent the physical relationships in addressing hybrid problems. This capability is regarded as one of the major functions of Neural Operators in the literature, since they mimic the physical representation of each one of the parcels of a PDE.

Nevertheless, there is always an inherent set of different noises in PINN solution development, which is probably caused by the nature of neural networks [33]. As statistical models, they tend to "blur" the solution as the optimization process progresses, potentially impacting accuracy [34]. The baseline PINN has been noted to struggle with convergence and accuracy when solving "stiff" PDEs. Addressing this limitation, the Self-Adaptive PINNs (SA-PINNs), a fundamentally new approach to train PINNs adaptively, have been introduced. This methodology dynamically adjust the training process to enhance both stability and accuracy.

The L^2 Physics-Informed Loss is the prevailing standard for training PINNs, commonly employed alongside additional loss terms that enforce compliance with fundamental physical laws, such as mass and energy conservation. However, this approach poses challenges due to the complex interplay between the overall loss function and the fidelity of the learned solution. PINNs often exhibit a smoothing effect, as the imposed physics-based constraints mitigate but do not entirely eliminate deviations from physical consistency. In conventional Computational Fluid Dynamics (CFD) simulations, users explicitly define error thresholds to establish convergence criteria, ensuring controlled numerical accuracy. In contrast, neural network-based approaches lack an equivalent mechanism for precisely defining stopping conditions, which complicates their integration into traditional numerical frameworks. Although PINNs have not yet achieved the same level of accuracy and reliability as CFD solvers for general problem-solving, their incorporation into numerical simulations presents promising opportunities. This study investigates the potential of utilizing PINNs to enhance the initialization of transient CFD simulations, aiming to improve computational efficiency and accelerate convergence.

3. Methodology

3.1. The Study Area

The study area is geographically located between the latitude coordinates 28.180214°N and 31.322086°N, and the longitude coordinates 97.1828°W and 93.5202°W, encompassing the city of Houston, Texas, and its surrounding regions. This area includes both urban and rural landscapes, offering a diverse range of environmental and climatic conditions for investigation [35,36]. The region under investigation is depicted in Figure 1.
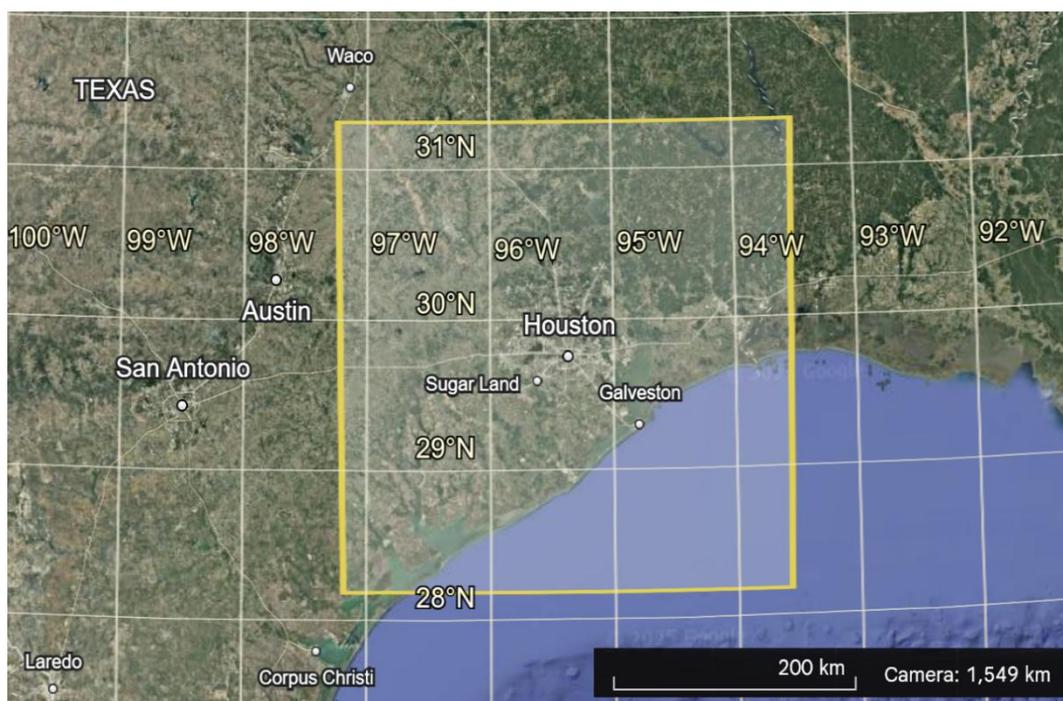
Figure 1. Image of the study location, highlighted by the yellow square [37].

The proximity of Houston to the Gulf of Mexico makes the region particularly influenced by maritime weather systems, such as hurricanes, intense convective activity, and high humidity levels [38,39]. The city's dense urban infrastructure also creates localized phenomena, such as urban heat islands, which interact with broader atmospheric dynamics [40,41]. The region is characterized by complex wind regimes shaped by coastal and inland influences. These include consistent southeasterly winds originating from the Gulf of Mexico, which dominate much of the year, as well as periodic shifts driven by frontal systems and cyclonic activity. These variable wind patterns significantly affect local weather, pollutant dispersion, and energy resource management, making them a critical factor in this study [42–45].

Selecting this region offers an opportunity to test and validate the use of artificial intelligence tools for improving initial conditions in transient CFD-based meteorological models like WRF. By focusing on a geographically and climatically dynamic area with intricate wind regimes, the study aims to enhance weather simulations' computational efficiency and accuracy, particularly in scenarios involving evolving weather systems rapidly.

3.2. Data Specifications and Variable Definitions

The study utilizes high-resolution meteorological data generated by the WRF model. The dataset is divided into two periods to facilitate the training and testing of the proposed methodologies. The dataset has a time resolution of 1 hour.

The training dataset comprises WRF simulation outputs spanning three years, from 2019 to 2021. This dataset captures a wide range of atmospheric conditions, including seasonal variations, extreme weather events, and changes in local wind regimes, aiming for a robust training process. Including several predictive parameters is often beneficial, deeming the model capable of generalizing across diverse scenarios [42,46].

For testing purposes, the study employed WRF simulation data from 2022. This separated dataset allows for an independent evaluation of the proposed methods, ensuring that the results accurately reflect the model's ability to predict future conditions without bias from the training data [47,48].

WRF output files contain numerous meteorological and environmental variables, including 2D surface fields and 3D atmospheric profiles. The variables were limited to a shorter list of key 2D and 3D attributes (Tables 1 and 2) to streamline the analysis and maintain consistency with the simulations. This focused selection ensures that the inputs used in the AI-based initialization are directly aligned with those employed in the actual WRF simulations. This improves the computational efficiency and preserves the critical information needed for accurate predictions.

Table 1. List of 2D variables present in the WRF simulations and used in the variable selection process. 'Times' is a 1D variable, but it is inserted here for the sake of space.

WRF name	Units	Description
'Times'	hh:mm:ss	Time
'XLAT'	degree	Latitude of the collocated grid
'XLONG'	degree	Latitude of the collocated grid
'XLAT_U'	degree	Latitude of the staggered grid in the X direction
'XLONG_U'	degree	Longitude of the staggered grid in the X direction
'XLAT_V'	degree	Latitude of the staggered grid in the Y direction
'XLONG_V'	degree	Longitude of the staggered grid in the Y direction
'COSZEN'	degree	Cosine of the Zenith Angle of the Sun
'PSFC'	Pa	Surface Pressure
'T2'	K	2-meter Temperature
'Q2'	kg water vapor/kg air	2-meter Specific Humidity
'U10'	m/s	10-meter U-Component of Wind
'V10'	m/s	10-meter V-Component of Wind
'SST'	K	Sea Surface Temperature
'TSK'	K	Skin Temperature
'TSLB'	K	Soil Temperature at Specified Layers
'HGT'	m	Geopotential Height
'ALBEDO'	Fraction (0-1)	Surface Albedo

Table 2. List of 3D variables present in the WRF simulations and used in the variable selection process.

WRF name	Units	Description
'U'	m/s	U-Component of Wind
'V'	m/s	V-Component of Wind
'W'	m/s	W-Component of Wind
'T'	K	Air Temperature
'THM'	K	Moist Potential Temperature
'QVAPOR'	kg water/kg dry air	Water Vapor Mixing Ration
'QCLOUD'	kg water/kg air	Cloud Water Mixing Ration
'QRAIN'	kg rainwater/kg air	Rainwater Mixing Ratio
'P'	Pa	Perturbation Pressure
'P_HYD'	Pa	Hydrostatic Pressure
'PB'	Pa	Perturbation Base-State Pressure

The latitude-longitude data and wind speed components (U, V, and W) were always included as inputs for all developed models. Additional auxiliary variables were tested separately to evaluate their impact on the performance of the models. The target variable for prediction was maintained consistently throughout this work as the X-direction wind speed component, 'U', reflecting the proof-of-concept nature of the study. The aim is to focus on a single output variable, allowing for a more concise and objective presentation and analysis of the results, ensuring clarity and allowing a direct analysis.

This study limited the number of time steps used to predict the next step to two. This choice aligns with the standards applied by Microsoft Aurora and was adopted for computational efficiency

and memory usage considerations [49]. The approach balances predictive accuracy with resource constraints by leveraging a minimal yet effective input size, making it feasible for operational meteorological applications.

Including hourly resolution data, using two timesteps, consistent use of wind and spatial variables, and reduced/optimized variable set ensure that both short-term and long-term atmospheric phenomena are accurately represented. These decisions collectively aim to enhance the reliability of the proposed techniques while maintaining the computational efficiency necessary for transient CFD models like WRF.

3.3. Applied Deep Learning Models

To evaluate the feasibility of using artificial intelligence to improve the initialization process of transient CFD models, two distinct deep learning architectures were tested: a Multilayer Perceptron (MLP) Network and a Fourier Neural Operator (FNO) Network. Each model was chosen for its capabilities to address the specific comparisons carried out in this work.

The Multilayer Perceptron (MLP) is a feedforward neural network where every neuron in one layer is connected to every neuron in the next layer. It is well-regarded for its simplicity and versatility in learning structured datasets, and it is theorized to serve as a universal function approximator [50]. An incremental approach of the hyperparameters was adopted to determine an optimal configuration. Initially, a zero hidden layer architecture was tested, where the network consisted only of an input layer and an output layer, providing a baseline for direct input-output mapping. Subsequently, the number of neurons in the input and output layers was varied. The best performance was achieved with 1024 neurons.

Once this configuration was established, the number of hidden layers was tested while freezing the number of neurons at 1024 per layer. The results revealed that the best outcomes on the testing dataset occurred when the network had 0 hidden layers, confirming that a shallow network was sufficient to capture the link between the inputs and the target variable (U, the X-axis component of wind velocity). For this setup, the Root Mean Squared Error (RMSE) for the testing set was 4.7551 m/s. While this result served as a useful baseline, it will be shown that the FNO significantly outperformed the MLP, demonstrating its suitability for the complex spatiotemporal patterns inherent to the WRF dataset.

The FNO represents a more advanced methodology tailored for learning operators in PDEs. Unlike traditional neural networks, the FNO models the solution operator, allowing it to predict not only a single instance but multiple instances simultaneously [28]. This property makes the FNO highly generalizable and capable of capturing a broader range of solution cases, as it is not constrained to a specific instance or condition. In our study, this feature was leveraged by training and testing the FNO across a variety of conditions.

The FNO's hyperparameters were optimized over two rounds (no more than that was needed), where the following parameters were systematically tested in sequence:

1. The size of the final hidden layer;
2. The number of Fourier modes, in powers of 2 (from 1 to 16);
3. The width, defined as the number of transforms applied to the Fourier modes;
4. The number of stacked Fourier layers, also in powers of 2 (from 1 to 16).

Via this iterative optimization process, the best architecture was found to have 1024 neurons in the final hidden layer, 16 Fourier modes, a width of 48, and only one Fourier layer. This configuration achieved a final testing RMSE of 1.2544 m/s, showcasing a significant improvement over the MLP model. The schematics of the used FNO architecture is presented in Figure 3.
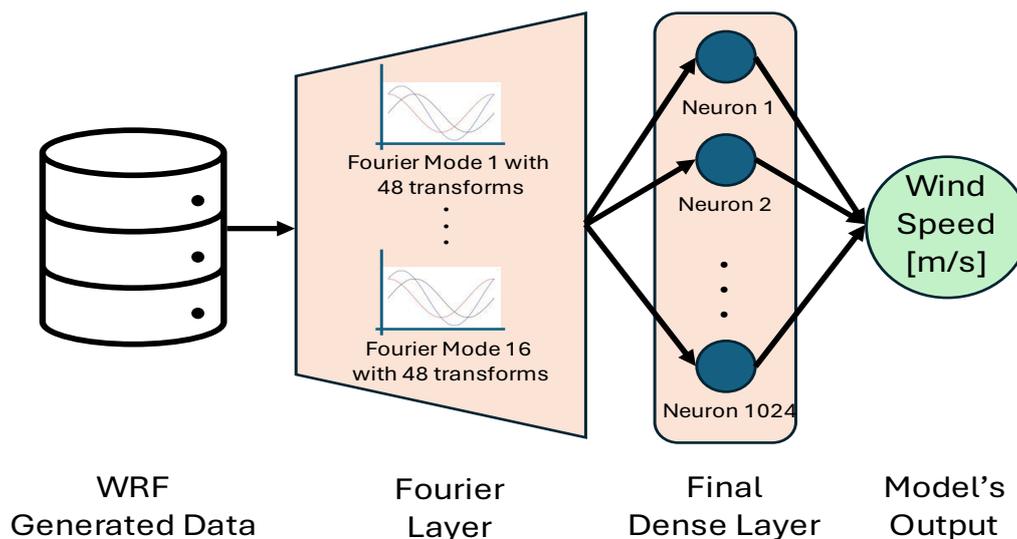
Figure 3. The selected FNO architecture.

For this step, the choice of input variables is aimed at maintaining the physical consistency of the system. We started with variables considered relevant to the underlying physics, and the initial dataset was limited to the following:

- 3D variables: U, V, W (wind speed components in X, Y, and Z directions) and T (temperature).
- 2D variables: latitude-longitude, while U10 and V10 (wind speed components at 10 m above the surface) were used to provide information on surface conditions.

The presented selected auxiliary variables were used as inputs, while the target variable was the X-direction wind speed component (U). The decision to predict a single output variable was made results more concisely since the results obtained are intended to be proof of concept, which can be expanded later to a complete weather forecasting application.

The FNO's reliance on Fourier transforms enables it to efficiently model spatial and temporal interactions, making it particularly well-suited for the distributed and transient nature of WRF data. The results highlight the FNO's potential to reduce computational costs and enhance numerical weather prediction systems by accelerating the convergence of transient CFD models.

3.4. Variable Selection

In this study, the initial approach to variable selection involved applying a single model across all Z-levels, encompassing and predicting all vertical layers in the dataset, what we call here and throughout the remainder of this manuscript "the global model". However, as it will be demonstrated later, this approach can be modified and enhanced by using one model per reference height, totaling 34 individual height levels. This adjustment aligns with the distinct atmospheric conditions that exist at different altitudes and allows for a more tailored and precise modeling approach.

The initial variable set used in the modeling process included U, V, and W (wind speed components in the x-easting, y-northing, and z-elevation directions), T (temperature), and U10 and V10 (wind speed components at 10 meters above ground level). This selection is consistent with the variables chosen for the FNO model discussed in the previous subsection.

A systematic, iterative process was implemented to refine the input variables and enhance prediction accuracy. In each round of testing, all available variables from the dataset (but U, V, W, and latitude-longitude, which were always kept) were considered as potential substitutes for the original set. The process alternated between testing only 2D variables (e.g., surface variables) and testing only 3D variables (e.g., variables across vertical layers).

The two-dimensional (2D) variables' surface-specific characteristics often result in minimal coupling with higher-altitude atmospheric dynamics. Consequently, if the inclusion of a given 2D variable in an iteration failed to yield a contraction in the RMSE, it was systematically excluded from subsequent training rounds to optimize model performance. In contrast, three-dimensional (3D) variables were retained across all iterations, as their incorporation consistently led to RMSE improvements. This persistent enhancement underscores the critical role of vertical atmospheric profiling in accurately resolving the physical mechanisms governing wind flow dynamics within the computational fluid dynamics (CFD) framework.

Finally, the variable set comprising HGT (height, 2D), QVAPOR (water vapor mixing ratio, 3D), and THM (perturbed moist potential temperature, 3D) emerged as the best-performing set for the case of using a single model across all Z-levels. This selection highlighted the relevance of combining surface and vertical atmospheric variables to accurately represent the physical conditions affecting wind flow.

The selection rounds continued until no further improvement was observed in the testing RMSE. This iterative optimization allowed us to determine the most relevant variables for predicting the target output, ensuring that unnecessary inputs were excluded and that computational efficiency was maintained (Figure 2).
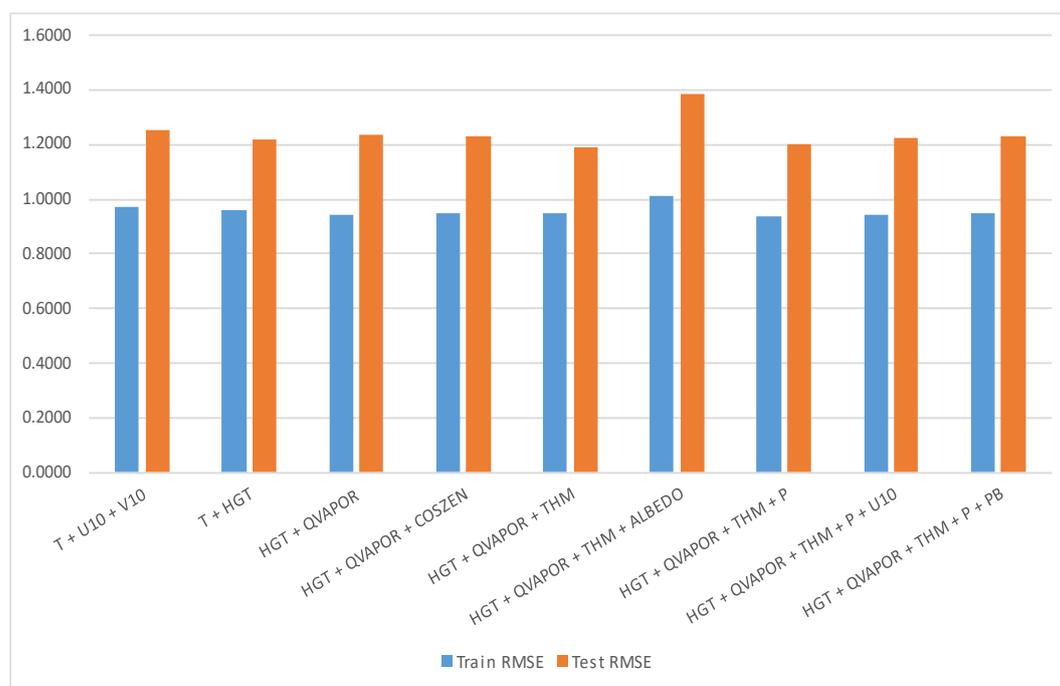
Figure 2. Results for the variable selection process using the training and test sets. This step regards the "global model", as explained, the one that predicts U for all the Z-levels simultaneously. Even though the best result was achieved for 'HGT'+'QVAPOR'+'THM', all the presented variable sets were used later to assess the localized models, developed specifically for each Z-level.

This methodology aimed to underscore the importance of variable selection in reducing model complexity and improving prediction accuracy using the lowest computer resources. The transition from a universal model across all Z-levels to individualized models per height, combined with an optimized variable set, presented in the next section, forms the foundation for the enhanced results presented in this study.

4. Results and Discussion

4.1. Global Model Performance

To assess the overall accuracy of the proposed approach, the performance of the global model, which simultaneously utilizes and predicts variables across all z-levels, was evaluated using a normalized metric. The RMSE values were normalized by dividing them by the average absolute value of the U velocity component across the entire domain [51]. This normalization facilitates a clearer interpretation of the model's performance by providing a relative measure of error magnitude in relation to the predicted variable.

$$nRMSE = \frac{RMSE}{\frac{\sum_i^n |U_i|}{n}} \tag{1}$$

For the global model, the resulting normalized RMSE (nRMSE) was 17.15%, indicating the percentage deviation from the expected values. In contrast, the persistence model, which represents the baseline approach commonly used in all CFD transient algorithms (e.g., [51,52]), where the velocity field of the previous time step is directly applied to the next step, achieved a lower nRMSE of 13.69% for the entire 3D domain.

The performance of the global model was further analyzed at three specific Z-levels: 0 (surface), 16 (middle atmosphere), and 33 (top of atmosphere). The results indicated that the global model did not outperform the persistence approach at any of these levels. At Z-level 0 (surface), the normalized root mean squared error (nRMSE) of the global model was 41.93%, whereas the persistence model achieved a lower error of 33.37%. The increased error at the surface is likely attributed to the presence of strong gradients caused by irregular terrain, boundary layer effects, and additional constraints imposed by empirical models used to approximate surface interactions, all of which introduce further uncertainties. [54].

At Z-level=16 (middle atmosphere), the nRMSE was 29.84%, while the persistence approach achieved 26.55%. The results at this altitude are closer between the two models, probably because the atmosphere becomes more predictable and stable, reducing variability and making persistence a more practical approach [53].

Finally, at Z-level=33 (top of atmosphere), the nRMSE was 12.30%, whereas the persistence model reached a significantly lower 6.32%. Even though the global model improved its performance at this height due to the more stable atmospheric conditions, it still did not outperform the persistence model. Furthermore, the improvement observed for the global model at this level was not as significant as at the lower levels, leaving its error almost twice as large as that of persistence.

The results obtained indicate that a single global model does not outperform the persistence model, at least for the optimized FNO model. This suggests that the current approach may not fully capture the dynamics of the atmosphere at different heights. Some factors may explain these results:

1. Lack of Height-Specific Modeling: The persistence model benefits from the natural continuity of the atmospheric flow, whereas the deep learning model needs to generalize across all heights. Studies such as [54] suggest that stratified modeling can significantly enhance predictive performance, where different models are trained for different atmospheric layers.
2. Boundary Layer Effects: Near-surface wind flows are influenced by topography, land use, and temperature variations, which cause strong local gradients and turbulence [55]. The inability of the global model to properly learn these effects might be a reason why it underperformed compared to persistence at Z-level 0.
3. Predictability of Atmospheric Layers: The results at Z-level=16 show that the atmosphere is more predictable at mid-altitudes, which aligns with previous findings that suggest lower turbulence and more stable wind regimes at these heights [56]. This explains why persistence and the global model performed similarly.
4. Impact of Training Data and Physical Constraints: The FNO model was trained using PINN principles, yet its ability to fully integrate the governing physics of the atmosphere might still be

limited compared to conventional numerical weather prediction (NWP) models. Hybrid approaches that combine FNO architectures with physics-based constraints have shown promise in recent literature [57].

These findings reinforce the need to refine our methodology. A promising alternative would be to apply height-dependent models, where a different FNO network is trained for each height level, ensuring that each model learns the specific characteristics of wind dynamics at its respective altitude. This will be further explored in the next section.

4.2. Specific Z-Levels Ad Hoc Models

This section presents the results of training separate models for each Z-level, which significantly improved performance compared to using a single global model for the entire 3D domain. The specialized models outperformed persistence for both the surface (Z = 0) and middle atmosphere (Z = 16). In contrast, at the top of the atmosphere (Z = 33), the results were slightly worse than persistence but still highly competitive.

Figure 4 illustrates these findings, highlighting the impact of height-specific modeling. For the surface level (Z = 0), the optimized FNO model achieved an nRMSE of 27.64%, outperforming the persistence model, which yielded 33.37%. The best-performing variable set for this level consisted of 'HGT', 'QVAPOR', 'THM', and 'P'. At the middle atmosphere (Z = 16), the FNO model further improved over persistence, with an nRMSE of 20.46% compared to 26.55%, using the variables 'HGT', 'QVAPOR', and 'THM'. At the top of the atmosphere (Z = 33), the nRMSE was 7.95% for FNO, while persistence performed slightly better at 6.32%, with the best variable set comprising only 'HGT' and 'QVAPOR'.
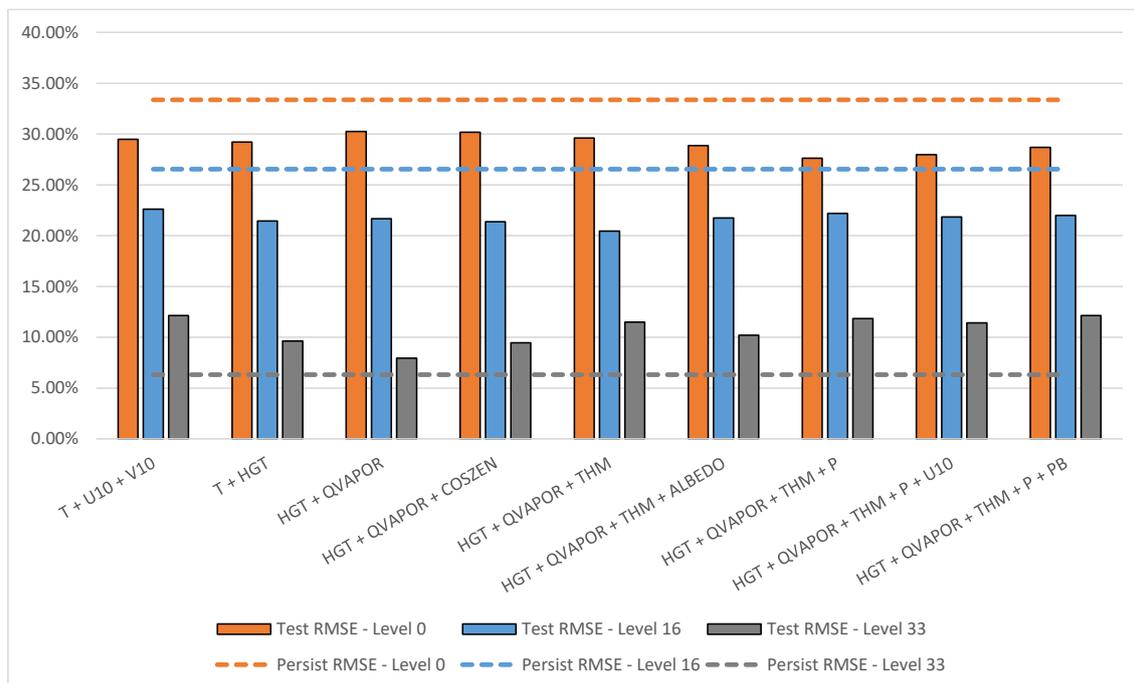
Figure 4. Normalized RMSE (nRMSE) Comparison for Z-Level Models (bars) vs. Persistence (dotted lines). The colors of the bars match the colors of the lines, for each case (Z-level).

These results align with previous atmospheric modeling and machine learning findings for weather prediction. Several studies emphasize that near-surface wind prediction is particularly challenging due to the influence of orography, boundary layer turbulence, and localized surface conditions [58,59]. The improvement seen at Z = 0 with the *ad-hoc* model can be attributed to better handling of these complex near-surface interactions, as the inclusion of pressure (P) appears to have contributed to refining the model's predictions. In contrast, at Z = 16, the atmosphere is already more

stable, which reduces the impact of localized disturbances, making the FNO model even more effective.

At the top of the atmosphere, the slight underperformance compared to persistence suggests that the model is not yet fully capturing the larger-scale atmospheric dynamics that dominate at these heights. However, the small difference in error indicates that the FNO approach remains competitive. Similar findings have been reported by [60], who showed that machine learning models tend to struggle with upper-atmosphere predictions unless explicitly trained with large-scale flow features.

Another key insight from this analysis is that, as altitude increases, fewer variables are needed to achieve optimal results. This observation is probably due to the atmosphere's increased stability and the lower influence of complex terrain interactions at higher altitudes. This trend is consistent with studies on numerical weather prediction, which indicate that lower atmospheric levels require more detailed parameterizations due to their interaction with the Earth's surface, while higher-altitude predictions benefit primarily from large-scale flow patterns [61].

Overall, the results confirm that height-specific modeling provides a significant advantage in transient wind field predictions. The FNO-based approach, when applied separately for each Z-level, successfully reduced prediction errors and demonstrated strong potential for improving atmospheric simulations over conventional persistence models.

Figures 5–7 illustrate the case of a specific time step for each studied Z-level (respectively 0, 16, and 33). It visually assesses the FNO model's performance at distinct altitudes. As observed, the model's accuracy improves with increasing height, aligning with findings from similar studies. Nevertheless, it can be seen how the color patterns of the modeled fields properly reproduce the ground truth pattern.
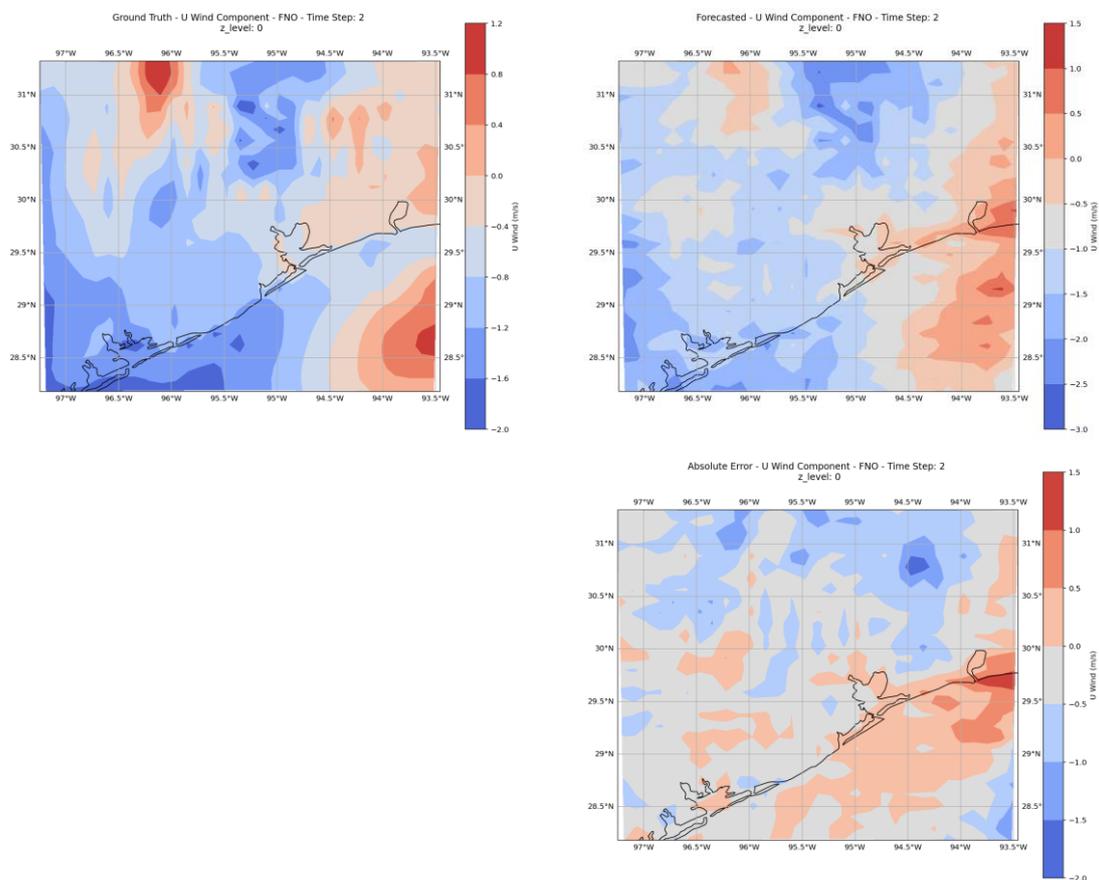
Figure 5. Example of the 'U' wind speed component field for Z-level=0 (earth's surface), specifically for the first modeled time step (2, since time steps 0 and 1 must be used as inputs). The ground truth is presented on the top-left plot, the top-right plot is the modeled field, and their difference is shown on the bottom-right plot.
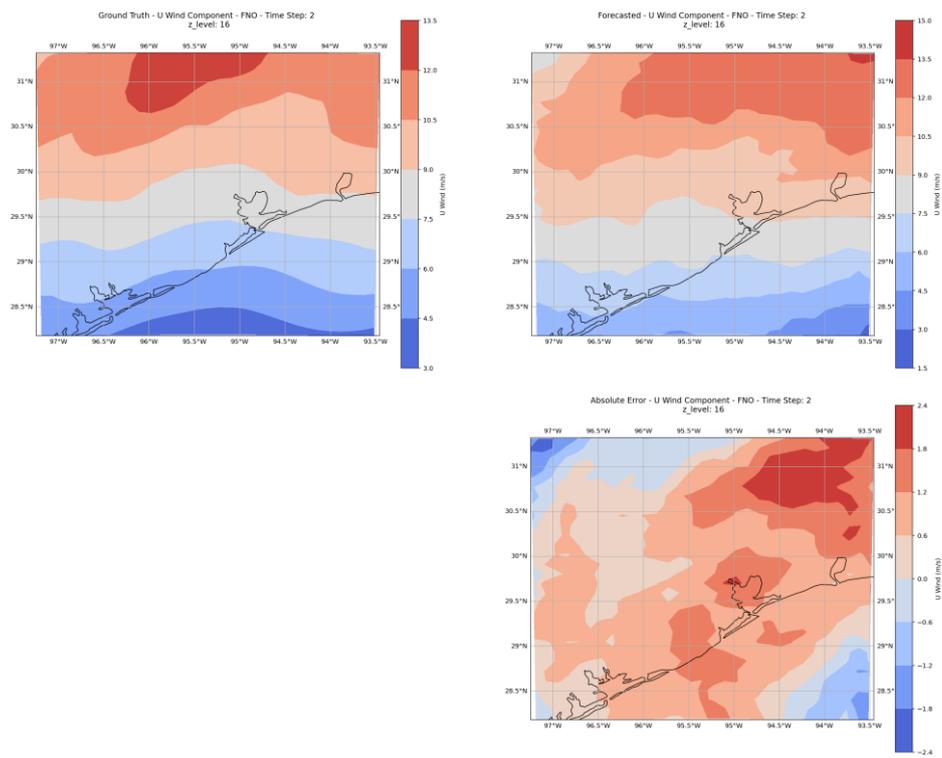
Figure 6. Same as Figure 5, for Z-level=16 (mid-atmosphere).
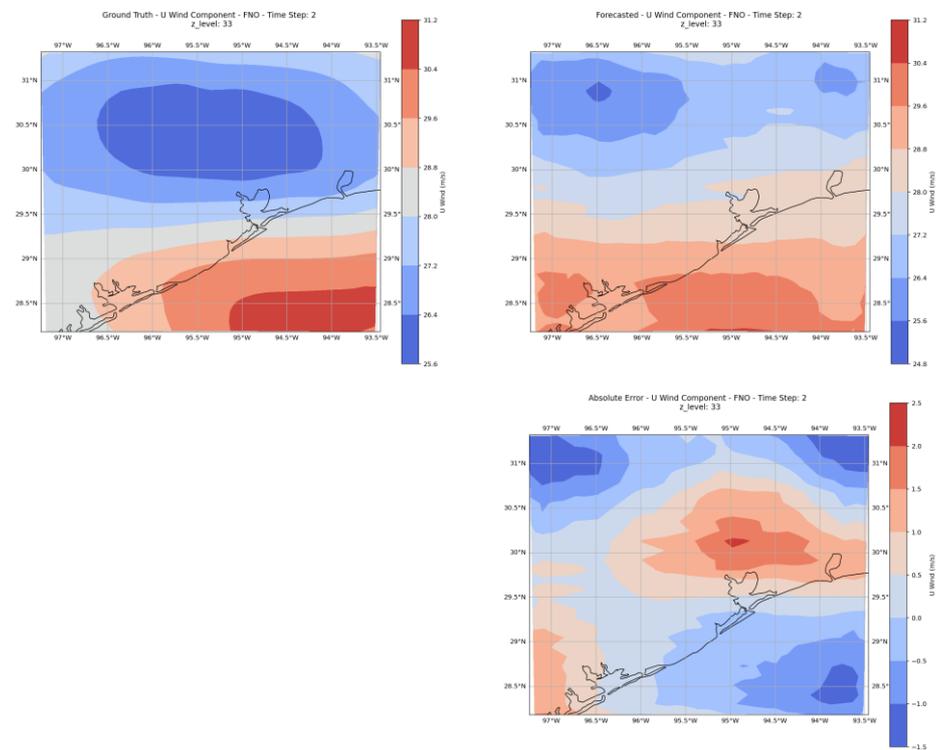
Figure 7. Same as Figure 5, for Z-level=33 (top-of-atmosphere).

To close the comparison, Figure 8 illustrates the variation of the nRMSE across the WRF atmospheric levels, comparing the performance of the persistence model with the ad hoc models developed for each level. The results indicate that the ad hoc models consistently outperform the persistence approach from the surface up to the mid-atmosphere. Beyond level 20, the performance of both approaches becomes comparable, and at the top of the atmosphere, the discrepancies are

negligible. The proposed methodology is particularly advantageous as it significantly reduces the error in most of the atmosphere, especially in regions with more intense variations, where CFD models will have a higher computational load to achieve convergence. This further highlights the importance of employing the developed models instead of relying on a simple persistence approach.
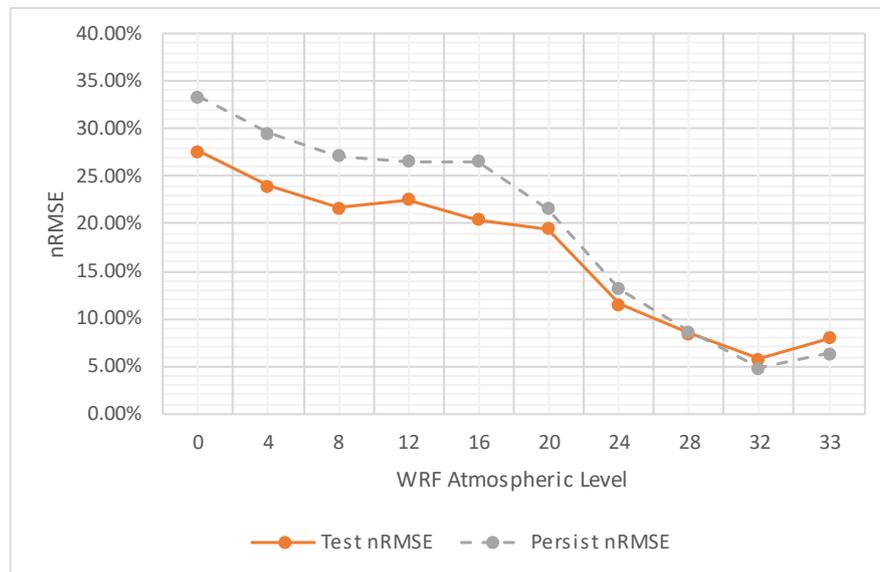
Figure 8. Comparison of the performance of the developed models against Persistence, for several Z-levels, from the surface to the top-of-atmosphere.

Similar results may be found in the literature. For instance, in [62] the FNO model demonstrated varying performance across distinct urban layouts and wind scenarios. While the results indicated that the FNO model could offer accurate prediction and achieve a remarkable 99% reduction in the required computational time, its ability to generalize to new wind directions proved challenging. This suggests that FNO model's performance is shaped by the complexity of the environment and the specific conditions it is applied.

Similarly, in [63], the authors investigated the application of machine learning models to predict wind speeds at higher altitudes based on data from lower levels. Their findings revealed that wind speed value at higher altitudes are directly influenced by the corresponding value at lower altitudes for a same geographical location. This indicates that machine learning models can properly capture altitude-dependent relationships, leading to accurate predictions at higher levels.

These studies show that the FNO model's performance improves with altitude, likely due to the decreasing complexity and increased predictability of atmospheric conditions at higher levels. The visual evidence provided in Figure 5 underscores this trend, highlighting the model's enhanced accuracy as height increases.

Other studies have shown the feasibility of Fourier-based approaches for atmospheric modeling. The model proposed by Pathak et al. [64] was built upon a Fourier Neural Network, used reanalysis data from the ERA5 dataset for training and validation, and successfully predicted wind speed for both land and sea regions. Differently from our study, however, their work found that increasing the number of attributes returned improved results by the model. This might be attributed to the fact that their model seeks a more general approach than ours and would benefit further from more spatiotemporal information provided by the input attributes. Contrarywise, the ad hoc models proposed in our study are tailored for specific atmospheric interactions, thus requiring different input attributes for each scenario, as the scenario's complexity decreases with higher altitudes.

In another study [65], the authors developed the Local-FNO model for modeling the microclimate of their study site. Their results highlighted that the model using local information for the location reached better outcomes than the vanilla FNO approach for wind dynamics modeling

for different building geometries. Their findings corroborate the superior efficiency of the ad hoc methodologies presented in this work on capturing local patterns.

5. Conclusions

This study developed and investigated a proof of concept on applying deep learning techniques for wind speed prediction using a Fourier Neural Operator (FNO) and a Multilayer Perceptron (MLP). The models were trained and tested with WRF-generated data for Houston from 2018 to 2022, focusing on predicting the X-direction wind speed component. The MLP was used primarily as a baseline, achieving a testing RMSE of 4.75 m/s, while the optimized FNO significantly outperformed it, reducing the RMSE to 1.25 m/s.

We initially trained a single global model across all Z-levels, which, despite its advanced architecture, still did not outperform the persistence model for wind prediction. Normalized RMSE values revealed that the global model reached an nRMSE of 17.15%, compared to 13.69% for the persistence model

To improve upon these results, we implemented and tested Z-level-specific models, significantly enhancing the performance at lower altitudes. At the surface (Z=0), our optimized FNO model achieved an nRMSE of 27.64%, outperforming persistence at 33.37%. In the middle atmosphere (Z=16), the nRMSE improved to 20.46%, compared to 26.55% for persistence. For the top of the atmosphere (Z=33), although persistence still outperformed our model (6.32% vs. 7.95%), the results remained highly competitive, which can help implement a fully automated process. This process shall benefit from using the model output to start the iterative process of the transient CFD solution, which is slow and has a high computational demand. This will require fewer iterations, optimizing processing time, and its associated financial costs.

Variable selection played a relevant role in model optimization. While all models always used latitude-longitude data and wind components, we tested auxiliary variables to pinpoint the most pertinent attributes for each height. The optimal feature sets varied across Z-levels, suggesting that different physical phenomena dominate wind behavior at different heights. At lower levels, the inclusion of HGT, QVAPOR, THM, and pressure yielded the best results, while higher-altitude models required fewer input variables, likely due to the increasing stability of the atmosphere.

The study case provided visual confirmation of model performance, reinforcing that FNO's effectiveness improves with altitude. This aligns with previous studies demonstrating that machine learning-based wind forecasting benefits from stratifying models based on atmospheric layers.

Overall, this research highlights the potential of deep learning for atmospheric forecasting, emphasizing the need for height-specific models. Future work may explore hybrid approaches combining deep learning with domain-specific physics constraints and extend the methodology to more diverse atmospheric conditions and locations. This type of result has the potential to lead to cheaper and faster weather simulations, which is a pressing demand in our modern society.

References

1. Cabrera, D.; Quinteros, M.; Cerrada, M.; Sánchez, R.-V.; Guallpa, M.; Sancho, F.; Li, C. Rainfall Forecasting Using a Bayesian Framework and Long Short-Term Memory Multi-Model Estimation Based on an Hourly Meteorological Monitoring Network. Case of Study: Andean Ecuadorian Tropical City. *Earth Sci. Inform.* **2023**, *16*, 1373–1388, doi:10.1007/s12145-023-00958-0.

2. Tahir Bahadur, F.; Rasool Shah, S.; Rao Nidamanuri, R. Air Pollution Monitoring, and Modelling: An Overview. *Environ. Forensics* **2024**, *25*, 309–336, doi:10.1080/15275922.2023.2297437.

3. Roy, C.; Rahman, Md.R.; Ghosh, M.K.; Biswas, S. Tropical Cyclone Intensity Forecasting in the Bay of Bengal Using a Biologically Inspired Computational Model. *Model. Earth Syst. Environ.* **2024**, *10*, 523–537, doi:10.1007/s40808-023-01786-3.

4. Singh, H.; Ang, L.-M.; Lewis, T.; Paudyal, D.; Acuna, M.; Srivastava, P.K.; Srivastava, S.K. Trending and Emerging Prospects of Physics-Based and ML-Based Wildfire Spread Models: A Comprehensive Review. *J. For. Res.* **2024**, *35*, 135, doi:10.1007/s11676-024-01783-x.

5. Kumar, V.; Sharma, K.V.; Caloiero, T.; Mehta, D.J.; Singh, K. Comprehensive Overview of Flood Modeling Approaches: A Review of Recent Advances. **2023**.

6. Wu, J.; Wang, Z.; Dong, J.; Cui, X.; Tao, S.; Chen, X. Robust Runoff Prediction With Explainable Artificial Intelligence and Meteorological Variables From Deep Learning Ensemble Model. *Water Resour. Res.* **2023**, *59*, e2023WR035676, doi:10.1029/2023WR035676.

7. Mitreska Jovanovska, E.; Batz, V.; Lameski, P.; Zdravevski, E.; Herzog, M.A.; Trajkovik, V. Methods for Urban Air Pollution Measurement and Forecasting: Challenges, Opportunities, and Solutions. *Atmosphere* **2023**, *14*, 1441, doi:10.3390/atmos14091441.

8. Kalajdjieski, J.; Trivodaliev, K.; Mirceva, G.; Kalajdziski, S.; Gievska, S. A Complete Air Pollution Monitoring and Prediction Framework. *IEEE Access* **2023**, *11*, 88730–88744, doi:10.1109/ACCESS.2023.3251346.

9. Federico, S.; Torcasio, R.C.; Popova, J.; Sokol, Z.; Pop, L.; Lagasio, M.; Lynn, B.H.; Puca, S.; Dietrich, S. Improving the Lightning Forecast with the WRF Model and Lightning Data Assimilation: Results of a Two-Seasons Numerical Experiment over Italy. *Atmospheric Res.* **2024**, *304*, 107382, doi:10.1016/j.atmosres.2024.107382.

10. Paul, D.J.; Janani, S.P.; J, A.J. Advanced Weather Monitoring and Disaster Mitigation System. In Proceedings of the 2024 International Conference on Cognitive Robotics and Intelligent Systems (ICC - ROBINS); April 2024; pp. 540–546.

11. Leung, X.Y.; Islam, R.M.; Adhami, M.; Ilic, D.; McDonald, L.; Palawaththa, S.; Diug, B.; Munshi, S.U.; Karim, M.N. A Systematic Review of Dengue Outbreak Prediction Models: Current Scenario and Future Directions. *PLoS Negl. Trop. Dis.* **2023**, *17*, e0010631, doi:10.1371/journal.pntd.0010631.

12. Dieng, M.D.B.; Tompkins, A.M.; Arnault, J.; Sié, A.; Fersch, B.; Laux, P.; Schwarz, M.; Zabré, P.; Munga, S.; Khagayi, S.; et al. Process-Based Atmosphere-Hydrology-Malaria Modeling: Performance for Spatio-Temporal Malaria Transmission Dynamics in Sub-Saharan Africa. *Water Resour. Res.* **2024**, *60*, e2023WR034975, doi:10.1029/2023WR034975.

13. Abrego-Perez, A.L.; Pacheco-Carvajal, N.; Diaz-Jimenez, M.C. Forecasting Agricultural Financial Weather Risk Using PCA and SSA in an Index Insurance Model in Low-Income Economies. *Appl. Sci.* **2023**, *13*, 2425, doi:10.3390/app13042425.

14. Costa Rocha, P.A.; Oliveira Santos, V.; Scott, J.; Van Griensven Thé, J.; Gharabaghi, B. Application of Graph Neural Networks to Forecast Urban Flood Events: The Case Study of the 2013 Flood of the Bow River, Calgary, Canada. *Int. J. River Basin Manag.* **2024**, 1–18, doi:10.1080/15715124.2024.2329243.

15. Luo, Z.; Liu, J.; Zhang, S.; Shao, W.; Zhang, L. Research on Climate Change in Qinghai Lake Basin Based on WRF and CMIP6. *Remote Sens.* **2023**, *15*, 4379, doi:10.3390/rs15184379.

16. Kadaverugu, R. A Comparison between WRF-Simulated and Observed Surface Meteorological Variables across Varying Land Cover and Urbanization in South-Central India. *Earth Sci. Inform.* **2023**, *16*, 147–163, doi:10.1007/s12145-022-00927-z.

17. Chen, L.; Chen, Z.; Zhang, Y.; Liu, Y.; Osman, A.I.; Farghali, M.; Hua, J.; Al-Fatesh, A.; Ihara, I.; Rooney, D.W.; et al. Artificial Intelligence-Based Solutions for Climate Change: A Review. *Environ. Chem. Lett.* **2023**, *21*, 2525–2557, doi:10.1007/s10311-023-01617-y.

18. Powers, J.G.; Klemp, J.B.; Skamarock, W.C.; Davis, C.A.; Dudhia, J.; Gill, D.O.; Coen, J.L.; Gochis, D.J.; Ahmadov, R.; Peckham, S.E.; et al. The Weather Research and Forecasting Model: Overview, System Efforts, and Future Directions. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 1717–1737, doi:10.1175/BAMS-D-15-00308.1.

19. Skamarock, W.C.; Klemp, J.B.; Dudhia, J.; Gill, D.O.; Barker, D.M.; Duda, M.G.; Huang, X.-Y.; Wang, W.; Powers, J.G. A Description of the Advanced Research WRF Version 3. *NCAR Tech. Note* **2008**, doi:10.5065/D68S4MVH.

20. Sosa-Tinoco, I.; Prósper, M.A.; Miguez-Macho, G. Development of a Solar Energy Forecasting System for Two Real Solar Plants Based on WRF Solar with Aerosol Input and a Solar Plant Model. *Sol. Energy* **2022**, *240*, 329–341, doi:10.1016/j.solener.2022.05.049.

21. Sarvestan, R.; Karami, M.; Javidi Sabbaghian, R. Evaluation of Meteorological Microphysical Schemas Based on the WRF Model for Simulation of Rainfall in the Northeastern Region of Iran. *J. Hydrol. Reg. Stud.* **2023**, *50*, 101524, doi:10.1016/j.ejrh.2023.101524.

22. Stathopoulos, C.; Chaniotis, I.; Patlakas, P. Assimilating Aeolus Satellite Wind Data on a Regional Level: Application in a Mediterranean Cyclone Using the WRF Model. *Atmosphere* **2023**, *14*, 1811, doi:10.3390/atmos14121811.

23. Hallaji, H.; Bohloul, M.R.; Peyghambarzadeh, S.M.; Azizi, S. Measurement of Air Pollutants Concentrations from Stacks of Petrochemical Company and Dispersion Modeling by AERMOD Coupled with WRF Model. *Int. J. Environ. Sci. Technol.* **2023**, *20*, 7217–7236, doi:10.1007/s13762-023-04959-w.

24. López-Noreña, A.I.; Berná, L.; Tames, M.F.; Millán, E.N.; Puliafito, S.E.; Fernandez, R.P. Influence of Emission Inventory Resolution on the Modeled Spatio-Temporal Distribution of Air Pollutants in Buenos Aires, Argentina, Using WRF-Chem. *Atmos. Environ.* **2022**, *269*, 118839, doi:10.1016/j.atmosenv.2021.118839.

25. Qi, Q.; Wang, S.; Zhao, H.; Kota, S.H.; Zhang, H. Rice Yield Losses Due to O3 Pollution in China from 2013 to 2020 Based on the WRF-CMAQ Model. *J. Clean. Prod.* **2023**, *401*, 136801, doi:10.1016/j.jclepro.2023.136801.

26. Carneiro, F.O.M.; Moura, L.F.M.; Costa Rocha, P.A.; Pontes Lima, R.J.; Ismail, K.A.R. Application and Analysis of the Moving Mesh Algorithm AMI in a Small Scale HAWT: Validation with Field Test's Results against the Frozen Rotor Approach. *Energy* **2019**, *171*, 819–829, doi:10.1016/j.energy.2019.01.088.

27. Ajani, C.K.; Zhu, Z.; Sun, D.-W. Recent Advances in Multiscale CFD Modelling of Cooling Processes and Systems for the Agrifood Industry. *Crit. Rev. Food Sci. Nutr.* **2021**, *61*, 2455–2470, doi:10.1080/10408398.2020.1809992.

28. Costa Rocha, P.A.; Johnston, S.J.; Oliveira Santos, V.; Aliabadi, A.A.; Thé, J.V.G.; Gharabaghi, B. Deep Neural Network Modeling for CFD Simulations: Benchmarking the Fourier Neural Operator on the Lid-Driven Cavity Case. *Appl. Sci.* **2023**, *13*, 3165, doi:10.3390/app13053165.

29. Kehrein, P.; Van Loosdrecht, M.; Osseweijer, P.; Posada, J.; Dewulf, J. The SPPD-WRF Framework: A Novel and Holistic Methodology for Strategical Planning and Process Design of Water Resource Factories. *Sustainability* **2020**, *12*, 4168, doi:10.3390/su12104168.

30. Cai, S.; Wang, Z.; Wang, S.; Perdikaris, P.; Karniadakis, G.E. Physics-Informed Neural Networks for Heat Transfer Problems. *J. Heat Transf.* **2021**, *143*, 060801, doi:10.1115/1.4050542.

31. Osorio, J.D.; Florio, M.D.; Hovsapian, R.; Chryssostomidis, C.; Karniadakis, G.E. Physics-Informed Machine Learning for Solar-Thermal Power Systems. *Energy Convers. Manag.* **2025**, *327*, 119542, doi:10.1016/j.enconman.2025.119542.

32. McClenny, L.D.; Braga-Neto, U.M. Self-Adaptive Physics-Informed Neural Networks. *J. Comput. Phys.* **2023**, *474*, 111722, doi:10.1016/j.jcp.2022.111722.

33. Zou, Z.; Meng, X.; Karniadakis, G.E. Uncertainty Quantification for Noisy Inputs–Outputs in Physics-Informed Neural Networks and Neural Operators. *Comput. Methods Appl. Mech. Eng.* **2025**, *433*, 117479, doi:10.1016/j.cma.2024.117479.

34. Fang, Z. A High-Efficient Hybrid Physics-Informed Neural Networks Based on Convolutional Neural Network. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 5514–5526, doi:10.1109/TNNLS.2021.3070878.

35. Beck, H.E.; Zimmermann, N.E.; McVicar, T.R.; Vergopolan, N.; Berg, A.; Wood, E.F. Present and Future Köppen-Geiger Climate Classification Maps at 1-Km Resolution. *Sci. Data* **2018**, *5*, 180214, doi:10.1038/sdata.2018.214.

36. Kyropoulou, M. Thermal Comfort and Indoor Air Quality in Higher Education: A Case Study in Houston, TX, during Mid-Season. *J. Phys. Conf. Ser.* **2023**, *2600*, 102022, doi:10.1088/1742-6596/2600/10/102022.

37. Google LLC Google Earth Available online: https://earth.google.com/earth/d/1QR_q78kghHqoVdYX6M2lsY-F3mK0wmkp?usp=sharing (accessed on 28 January 2025).

38. Li, X.; Fu, D.; Nielsen-Gammon, J.; Gangrade, S.; Kao, S.-C.; Chang, P.; Morales Hernández, M.; Voisin, N.; Zhang, Z.; Gao, H. Impacts of Climate Change on Future Hurricane Induced Rainfall and Flooding in a Coastal Watershed: A Case Study on Hurricane Harvey. *J. Hydrol.* **2023**, *616*, 128774, doi:10.1016/j.jhydrol.2022.128774.

39. Chang, H.; Ross, A.R. *Climate Change, Urbanization, and Water Resources: Towards Resilient Urban Water Resource Management*; Springer International Publishing: Cham, 2024; ISBN 978-3-031-49629-5.

40. Chakraborty, T.; Hsu, A.; Manya, D.; Sheriff, G. Disproportionately Higher Exposure to Urban Heat in Lower-Income Neighborhoods: A Multi-City Perspective. *Environ. Res. Lett.* **2019**, *14*, 105003, doi:10.1088/1748-9326/ab3b99.

41. Mohammad Harmay, N.S.; Choi, M. The Urban Heat Island and Thermal Heat Stress Correlate with Climate Dynamics and Energy Budget Variations in Multiple Urban Environments. *Sustain. Cities Soc.* **2023**, *91*, 104422, doi:10.1016/j.scs.2023.104422.

42. Oliveira Santos, V.; Costa Rocha, P.A.; Scott, J.; Van Griensven Thé, J.; Gharabaghi, B. Spatiotemporal Air Pollution Forecasting in Houston-TX: A Case Study for Ozone Using Deep Graph Neural Networks. *Atmosphere* **2023**, *14*, 308, doi:10.3390/atmos14020308.

43. Patel, Y. Levelized Cost of Repurposing Oil and Gas Infrastructure for Clean Energy in the Gulf of Mexico. **2025**.

44. Lamer, K.; Mages, Z.; Treserras, B.P.; Walter, P.; Zhu, Z.; Rapp, A.D.; Nowotarski, C.J.; Brooks, S.D.; Flynn, J.; Sharma, M.; et al. Spatially Distributed Atmospheric Boundary Layer Properties in Houston – A Value-Added Observational Dataset. *Sci. Data* **2024**, *11*, 661, doi:10.1038/s41597-024-03477-9.

45. Oliveira Santos, V.; Costa Rocha, P.A.; Thé, J.V.G.; Gharabaghi, B. Optimizing the Architecture of a Quantum–Classical Hybrid Machine Learning Model for Forecasting Ozone Concentrations: Air Quality Management Tool for Houston, Texas. *Atmosphere* **2025**, *16*, 255, doi:10.3390/atmos16030255.

46. Baïle, R.; Muzy, J.-F. Leveraging Data from Nearby Stations to Improve Short-Term Wind Speed Forecasts. *Energy* **2023**, *263*, 125644, doi:10.1016/j.energy.2022.125644.

47. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*; Springer Science & Business Media, 2009; ISBN 978-0-387-84858-7.

48. Deisenroth, M.P.; Faisal, A.A.; Ong, C.S. *Mathematics for Machine Learning*; Cambridge University Press, 2020; ISBN 978-1-108-47004-9.

49. Bodnar, C.; Bruinsma, W.P.; Lucic, A.; Stanley, M.; Vaughan, A.; Brandstetter, J.; Garvan, P.; Riechert, M.; Weyn, J.A.; Dong, H.; et al. A Foundation Model for the Earth System. **2024**, doi:10.48550/arXiv.2405.13063.

50. James, G.; Witten, D.; Hastie, T.; Tibshirani, R.; Taylor, J. *An Introduction to Statistical Learning: With Applications in Python*; Springer International Publishing, 2023; ISBN 978-3-031-38746-3.

51. HOLTON, J.R. An Introduction to Dynamic Meteorology. *Introd. Dyn. Meteorol.* **1992**, *48*, 1–497.

52. Johnson, P.L.; Wilczek, M. Multiscale Velocity Gradients in Turbulence. *Annu. Rev. Fluid Mech.* **2024**, *56*, 463–490, doi:10.1146/annurev-fluid-121021-031431.

53. Holton, J.R. *An Introduction to Dynamic Meteorology*; Academic Press, 2004; ISBN 978-0-12-354015-7.

54. Ranade, R.; Hill, C.; Pathak, J. DiscretizationNet: A Machine-Learning Based Solver for Navier–Stokes Equations Using Finite Volume Discretization. *Comput. Methods Appl. Mech. Eng.* **2021**, *378*, 113722, doi:10.1016/j.cma.2021.113722.

55. Cant, S. S. B. Pope, Turbulent Flows, Cambridge University Press, Cambridge, U.K., 2000, 771 Pp. *Combust. Flame* **2001**, *125*, 1361–1362, doi:10.1016/S0010-2180(01)00244-9.

56. Stevens, B.; Bony, S. What Are Climate Models Missing? *Science* **2013**, *340*, 1053–1054, doi:10.1126/science.1237554.

57. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-Informed Machine Learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440, doi:10.1038/s42254-021-00314-5.

58. Bochenek, B.; Ustrnul, Z. Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives. *Atmosphere* **2022**, *13*, 180, doi:10.3390/atmos13020180.

59. Hoolohan, V.; Tomlin, A.S.; Cockerill, T. Improved near Surface Wind Speed Predictions Using Gaussian Process Regression Combined with Numerical Weather Predictions and Observed Meteorological Data. *Renew. Energy* **2018**, *126*, 1043–1054, doi:10.1016/j.renene.2018.04.019.

60. Zhang, H.; Liu, Y.; Zhang, C.; Li, N. Machine Learning Methods for Weather Forecasting: A Survey. *Atmosphere* **2025**, *16*, 82, doi:10.3390/atmos16010082.

61. Gentine, P.; Pritchard, M.; Rasp, S.; Reinaudi, G.; Yacalis, G. Could Machine Learning Break the Convection Parameterization Deadlock? *Geophys. Res. Lett.* **2018**, *45*, 5742–5751, doi:10.1029/2018GL078202.

62. Chen, C.; Tian, G.; Qin, S.; Yang, S.; Geng, D.; Zhan, D.; Yang, J.; Vidal, D.; Wang, L.L. Generalization of Urban Wind Environment Using Fourier Neural Operator Across Different Wind Directions and Cities 2025.

63. P., V.; Thumba, D.A.; Kumar, S. Machine Learning-Based Simplified Methods Using Shorter Wind Measuring Masts for the Time Ahead Wind Forecasting at Higher Altitude for Wind Energy Applications. *Renew. Energy Environ. Sustain.* **2022**, *7*, 24, doi:10.1051/rees/2022012.

64. Pathak, J.; Subramanian, S.; Harrington, P.; Raja, S.; Chattopadhyay, A.; Mardani, M.; Kurth, T.; Hall, D.; Li, Z.; Azizzadenesheli, K.; et al. FourCastNet: A Global Data-Driven High-Resolution Weather Model Using Adaptive Fourier Neural Operators 2022.

65. Qin, S.; Zhan, D.; Geng, D.; Peng, W.; Tian, G.; Shi, Y.; Gao, N.; Liu, X.; Wang, L.L. Modeling Multivariable High-Resolution 3D Urban Microclimate Using Localized Fourier Neural Operator. *Build. Environ.* **2025**, *273*, 112668, doi:10.1016/j.buildenv.2025.112668.