A Case-based Explanation Method for Weather Forecasting

Moisés Fernando **Valdez-Ávila**[1,2], Gerardo Arturo **Pérez-Pérez**[1,2],
Humberto **Sarabia-Osorio**[1,2], Carlos **Bermejo-Sabbagh**[1,2] and
Mauricio G. **Orozco-del-Castillo**[1,2,*]

[1]*Tecnológico Nacional de México/IT de Mérida, Department of Systems and Computing, Merida, Mexico*
[2]*AAAIMX Student Chapter at Yucatan, Mexico (AAAIMX), Association for the Advancement of Artificial Intelligence, Mexico*

Abstract

This paper presents an explainable Artificial Intelligence of Things (AIoT) solution that combines the information provided by an environmental Internet of Thigs (IoT) sensor with the potential of artificial intelligence to obtain climate predictions using forecasting techniques. We propose a recurrent artificial neural network that provides personalized weather predictions (humidity, pressure, and temperature) based on the concrete environmental time series collected through a sensor installed in a wearable or mobile device locations of the user. However, as neural networks are black-box models which do not allow users to better understand the complex dynamics associated with climatology and the reasons that support a weather prediction, we propose a cased-based reasoning approach to explain how (future) predictions could be dependent of past time series windows.

Keywords

Explainable Artificial Intelligence, Weather Forecasting, Artificial Neural Networks, Time Series, Artificial Intelligence of Things

1. Introduction

The current rise of the Internet of Things (IoT) technology has produced a wide range of sensing solutions that are progressively being integrated into our daily life devices such as mobile phones or wearables [1]. The combination of such sensing capabilities with Artificial Intelligence (AI) is producing Artificial Intelligence of Things (AIoT) applications that provide an enhanced user experience [2].

Within the climate prediction field, a variety of efforts has been made to develop complex models to perform seasonal climate prediction [3, 4], however, this process is inevitably subject to many error sources [4]. Climate forecasting has been usually divided into two different approaches: deterministic and probabilistic forecasting. Deterministic forecasting aims to provide a quantitative estimate of the future value of climate variables. Probabilistic forecast, on

the other hand, aims to provide the probability distribution for the future state of the variables. Since a forecast of the occurrence probability for a given event can bring greater economic value than a single deterministic forecast with uncertain accuracy [5], these methods tend to be preferred.

Time series forecasting is the prediction of future data values based on some collected data [6] and has been an area of great interest in science, engineering, and business. Traditional time series forecasting is usually approached by the analysis of its internal structure: autocorrelation, trend, seasonality, etc., to capture the pattern of long-time behavior of the system [6]. These models predict future values of a target $y_i(t)$ for a given observed value i at a time t. In climatology, these observed values usually represent measurements from individual weather stations [7]. While this applies to univariate forecasting, the same components can be extended to multivariate models without loss of generality [8].

Traditional time series forecasting methods tend to focus on parametric models informed by the expertise in a given domain, including autoregressive, exponential smoothing, or structural models [7]. However, in recent years machine learning (ML) methods have provided a means to capture dynamics of time series directly from the data, instead of domain expertise information [9]. The increasing data availability and computing power have positioned ML-based time series forecasting models as the next generation of time series forecasting models [7]. One of the main techniques in ML is artificial neural networks (ANNs), which have proven to be a reliable tool for time series analysis [10], and numerous ANNs design choices have emerged given the diversity of time series problems across multiple domains [7]. Of these designs, the most commonly used one in time series forecasting is recurrent neural networks (RNNs) [11]. This is due to the natural interpretation of time series data as sequences of inputs and targets [7].

Although ML approaches have demonstrated very good prediction performance, they have a significant limitation regarding its explainability. Neural network models are considered as "black boxes", because their internal processes are difficult to interpret with respect to the predictions they produce [12]. Solving this problem is a requirement to audit the reasoning behind incorrect prediction taken by AI systems and foresee the data patterns that may led to a climate event. Assuming that ANNs achieve better performance than white-box models, in this paper we propose a post-hoc sliding-window method that enables the explanation of an ANN forecasting model for the weather prediction domain. This way, instead of using a transparent or white box model with lower performance than an ANN, we enable the use of these black-box prediction models but augmented with explanation capabilities. The post-hoc explanation system follows an explanation-by-example approach implemented through Case-based Reasoning (CBR). Here, time series are split into different time-window cases that serve as explanation cases for the outcome of the ANN prediction model.

In this paper we focus on the combination of an environmental sensor for mobile devices, such as the BOSCH Sensortec BME680[1] capable of sensing several weather variables [13, 14], with the potential of AI to compute climate predictions using forecasting techniques. Our ultimate goal is to provide an AIoT solution that brings the user the possibility to obtain personalized predictions based on the historical environmental data collected through the sensors installed

---

[1]https://www.bosch-sensortec.com/products/environmental-sensors/gas-sensors/bme680/

in a wearable or mobile device. Moreover, the aim to provide explainable predictions that improves the user's acceptance of the predictions. This way, instead of receiving a generic climate prediction based on the nearby weather stations (that can be relatively far), this AIoT solution provides personalized weather predictions based on the concrete locations of the user, incorporating an explanation mechanism that lets the user understand the previous climate values that led the AI to make a concrete forecasting prediction.

Paper runs as follows. Section 2 presents the background of this work. Then, Section 3 describes our case-based explanation method. Section 4 presents the evaluation results and section 5 concludes the paper and opens lines of future work.

2. Background

In a recent work [15], scientists have used eXplainable Artificial Intelligence (XAI) to estimate precipitation from satellite images. The aim of this research was to exploit the high potential of ANNs used with satellite images to capture spatial information and inform numerical modeling and forecasting, which could help to mitigate weather risks. Using XAI techniques, insights gained from long-range planning (LRP) have given scientists confidence that the ANN generated predictions based on a physically reasonable strategy, and thus helped build more confidence in their predictions. Furthermore, if scientists want to improve the model by testing different model architectures, knowing how physically consistent the different model decision strategies are, provides a criterion for distinguishing between models [16]. This consequence of using XAI techniques goes well beyond the prediction performance obtained by using models such as ANNs by themselves without any explicability.

Some works have focused on conterfactual explanations, but not many have aimed at time series. Despite this, conterfactual explanations has become very popular, with more than 100 methods proposed, one of these is the native guide method [17]; this explainability method defines some key properties of good counterfactual explanations such as proximity, sparsity, plausibility, and diversity.

Like weather forecasting, with respect to climate prediction at subseasonal, seasonal, and decadal timescales, empirical statistical models exhibit limited predictive skill due to the complex and non-stationary nature of the relationship between large scale modes and regional climate. To address this issue, data-driven ML methods that leverage information from the entire globe have shown improvements in predictive skill [18]. A number of studies have highlighted the potential of ANNs in predicting climate across a range of scales because of their ability to capture nonlinear relations [19], while more recent studies have used XAI methods to explain these ANNs and their strategies to increase trust and generate new knowledge [20]. Mayer and Barnes [21] used XAI to show that ANNs can identify when favorable conditions that lead to enhanced predictive skill of regional climate are present in the atmosphere or not. In a decadal climate prediction application, Toms et al. [22] used simulated data from climate models to explore sources of decadal predictability in the climate system. The results showed that there are several regions where surface temperature is practically unpredictable, whereas there are also regions of high predictability. Such an analysis could motivate further mechanistic investigation to physically establish new climate teleconnections, illustrating how XAI methods can help
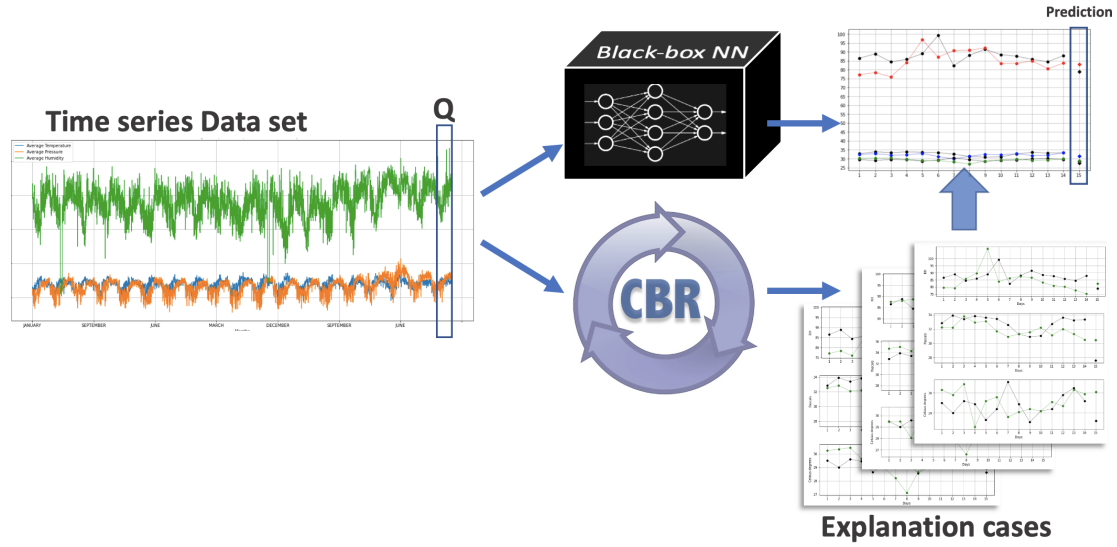
Figure 1: Global schema of the RNN-CBR twin

advance climate science [16]. Further XAI applications have focused on analyzing the impacts of climate change on the energy consumption in building, explaining the underlying reasons behind the predictions [23].

3. Method

The main contribution of this paper is the use of CBR for the generation of explanations associated to the prediction of the ANN. CBR systems are claimed to have a "natural" transparency as they are based on the reuse of previous experiences or examples. Therefore, we propose a particular solution for the explanation of the outcomes of the ANN to the experts, where this opaque, black-box ML system is explained by a more interpretable, white-box CBR system, following the so-called twin-systems approach [24]. This approach is illustrated in Figure 1, where the dataset is used as the input of the ANN and to create the explanatory cases provided by the CBR system. Explanatory cases are generated from the weather series using a sliding-window method over the whole time series: $C_t = \langle [t - ws, t], R_{t+1} \rangle$ where ws is the window size and R_{t+1} is the solution of the case, which corresponds to the meteorological reading for the next day. Analogously, given a query time stamp t_q, the ANN will predict the time series values for that date: $Pred(t_q)$. Then, the query for our CBR system will be the time window $Q = [t_q - 1 - ws, t_q - 1]$. Next, the prediction given by the ANN for t_q is explained by means of the most similar explanatory cases to the current time window Q, following the explanation-by-example paradigm illustrated in Figure 1.

3.1. Dataset and model

The historical dataset of the meteorological variables was recorded by the Mexican National Water Council (*Comisión Nacional del Agua, CONAGUA*) ground station located at the city of Mérida; among the data provided, the following variables were found, simulating the values that can be also obtained by the BOSCH BME680 sensor: temperature (T), vapor pressure (P), and relative humidity (H). The range dates of the records were from January 1, 2000 to September 30, 2018, where there is a daily record of temperatures with minimum, maximum, and average units, resulting in nine readings provided for each day. A RNN using LSTM cells was trained over the dataset of time windows and weather labels. 70% of the dataset was used during training, while the remaining 30% was used for testing purposes. The dataset is then divided into time windows with $ws = 14$, obtaining explanation cases with 14 days of weather evolution.

3.2. Time windows correlation

Under the assumption that time windows with similar climate will present similar outcomes for the next days, we could reuse previous time windows to explain new ones in the future, without the use of an ANN. To do so, we need to define several similarity metrics for the retrieval.

Let us define the Combined Correlation Index (CCI), which provides a way to measure how a given time window case $\mathbf{C}$ is related to a target query window $\mathbf{Q}$:

$$\mathrm{CCI}(\mathbf{C}, \mathbf{Q}) = \frac{1}{4}(\rho(\mathbf{C}, \mathbf{Q}) - 2\|(\mathbf{C}, \mathbf{Q})\| + 3), \tag{1}$$

where ρ is the function that calculates the Pearson correlation coefficient, and the double bars represent the normalized Euclidean distance between those vectors. In our particular case, we define CCI_f as the CCI of a given meteorological feature f (temperature, vapor pressure, or humidity), $\mathbf{C}_f$ represents a fixed window for comparison of different (sliding) windows $\mathbf{Q}_f$, vectors of 14 values registered daily for the same aforementioned meteorological feature. The normalized Euclidean distance is simply the Euclidean distance between all possible vectors (windows) in the time series, but normalized so it is defined in the range $[0, 1]$. CCI is defined in the range $[0, 1]$, where higher values indicate a greater similarity between time windows, and viceversa. The correlation component deals with the morphological similarity of the time windows, while the Euclidean distance component deals with the proximity between the time series in the given time windows.

An enhanced CCI can be obtained using the original CCI values, so that:

$$\mathrm{CCI}_f{}^{+}(\mathbf{C}_f, \mathbf{Q}_f) = (\mathrm{CCI}_f(\mathbf{C}_f, \mathbf{Q}_f) + h)^2, \tag{2}$$

where h corresponds to an "enhancement value" which increases higher CCI values (those over 0.5), while decreasing lower ones. In our paper, since CCI is in the range $[0, 1]$, $h = 0.5$, and $\mathrm{CCI}_f{}^{+}$ is in the range $[0.25, 2.25]$.

With this in mind, we define the full combined correlation index, FCCI, between two windows $\mathbf{C}_f$ and $\mathbf{Q}_f$ as the sum of the $\mathrm{CCI}_f{}^{+}$ computed for every meteorological feature (humidity, temperature and vapor pressure: H, T and P, respectively). This is mathematically defined as follows:
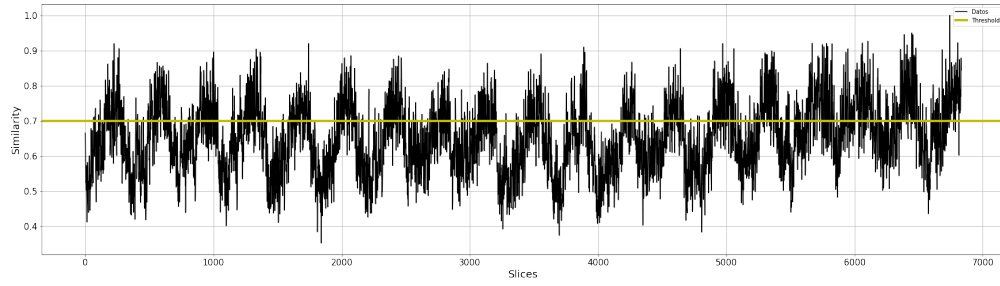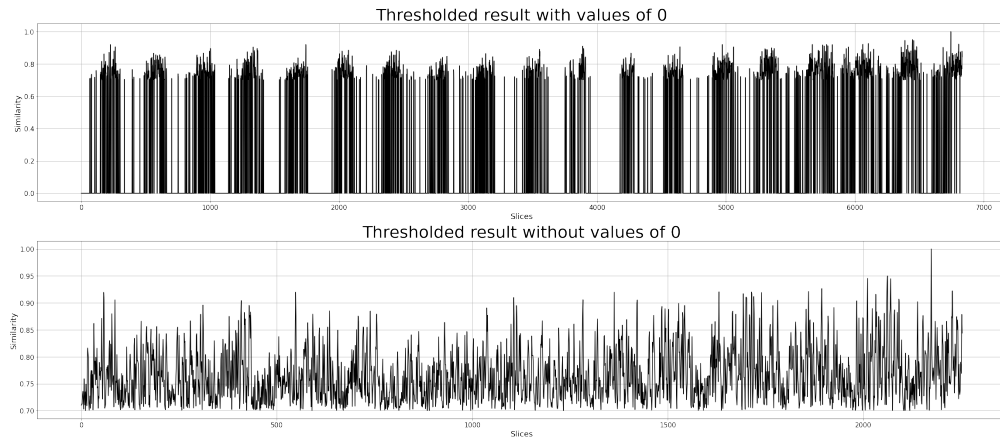
Figure 2: threshold being applying to the data.

Figure 3: Results of applying the threshold.

$$\text{FCCI}(\mathbf{C}_f, \mathbf{Q}_f) = \text{CCI}_H{}^+(\mathbf{C}_H, \mathbf{Q}_H) + \text{CCI}_T{}^+(\mathbf{C}_T, \mathbf{Q}_T) + \text{CCI}_P{}^+(\mathbf{C}_P, \mathbf{Q}_P). \qquad (3)$$

This value can also be normalized, dividing each FCCI result between the highest FCCI encountered, therefore FCCI is defined in the range $[0, 1]$.

3.3. Retrieval process

The retrieval process begins with the thresholding of the FCCI values to filter the windows with the greatest similarity results. To threshold the correlations results we define a minimum similarity value s_{min} which was empirically defined a value of 0.7. The values of FCCI for all the time series are shown in Figure 2, with the s_{min} threshold overlaid as an horizontal yellow line. The resulting data after the threshold was applied is shown in Figure 3.

However, the calculated FCCI values yield undesirable high frequency components as shown in Figure 3. To solve this problem, a low pass filter can be applied over the FCCI time series to smooth the readings. In our case, we performed a filtering phase which consisted of the multiple application of a simple moving average filter (MAF). After iteratively applying this filter until the resulting signal no longer changes, we obtain a smoothed time series for the FFCI.
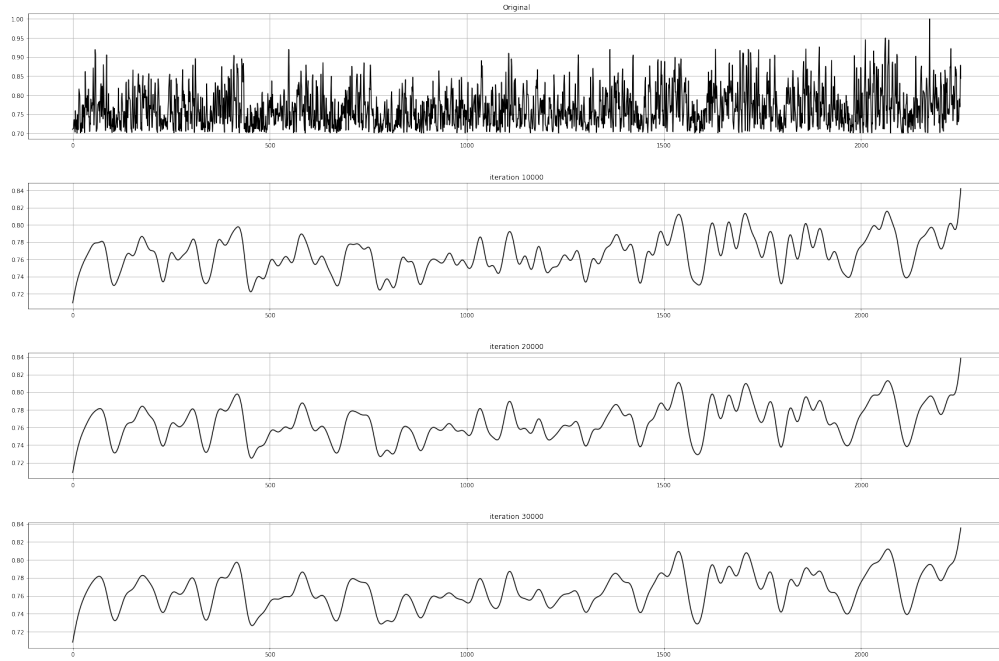
Figure 4: FFCI time series after several filtering iterations. After successive iterations, the time series converges into a smooth signal.

Smoothed time series can be seen in Figure 4.

The first numerical derivative of this time series is also obtained, helping in the identification of peaks and valleys in the FFCI, which will be used to segment it into groups of ascending and decreasing curves to obtain the maximum and minimum values.

The highest FFCI values are used to identify the k most similar explanation cases to the time-window query Q. These are the cases that are presented to the user in order to explain the prediction given by the ANN model. An example showing the windows of humidity, pressure, and temperature, for the highest FFCI value with respect to the time-window query, is presented in the left-hand side of Figure 5. For visual comparison purposes, we also identified the window corresponding to the lowest FFCI value, and is shown on the right-hand side of this same figure. Notice how we also included the prediction of the ANN for the 15th day, as well as the real value of the 15th day, the following value of these two windows for each climate feature, respectively.

4. Evaluation

Even though a visual inspection of the plots of the windows shows an evident difference between the highest and the lowest FCCIs, respectively, we proposed a quantitative approach to measure this difference. When a window is retrieved from any of the absolute highest peaks (maximum or minimum) in the FFCI signal, we calculated the mean absolute error (MAE) between the ANN prediction for the t_q date and the actual values in the solution of the case, R_t, containing the readings of the day represented by the explanation case, particularly:
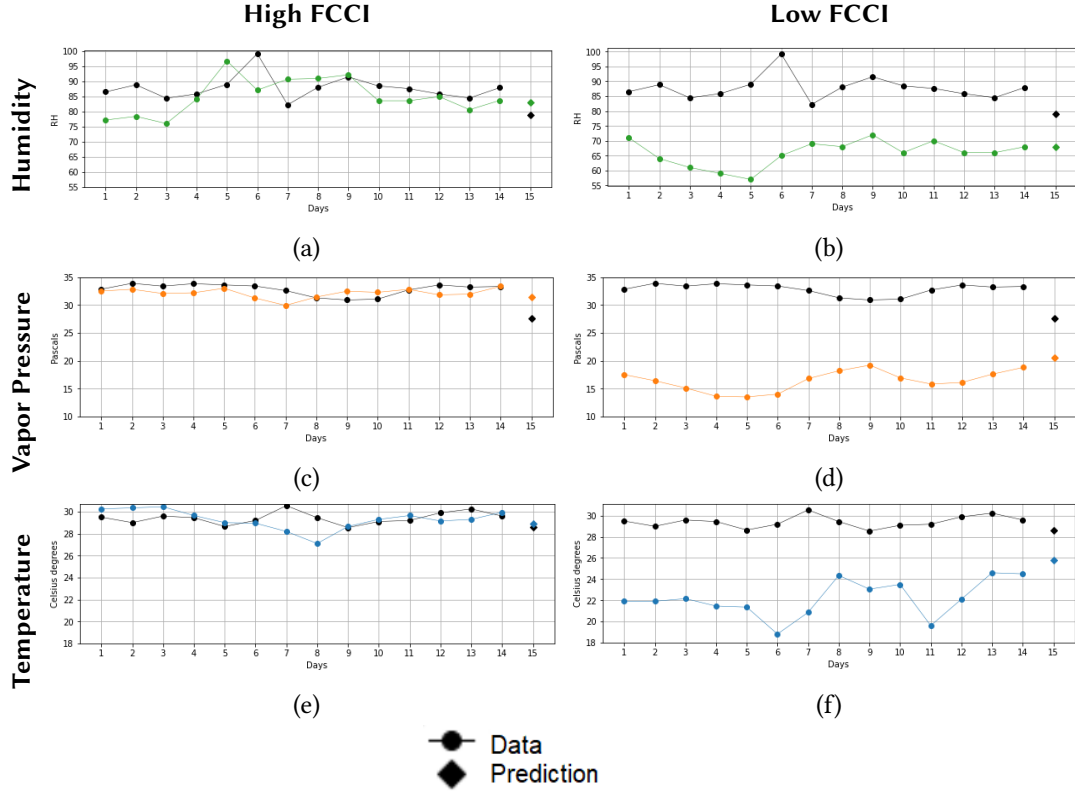
Figure 5: Example of explanation through the comparison of meteorological features between the **Q** and the most similar explanation case (left column). Right column compares with the most dissimilar one (right column). Both according to the FFCI time readings).

$$\text{MAE}(\mathbf{pred}(t_q), R_t) = \frac{1}{|\text{TS}|} \sum_{V \in \{\text{TS}\}} |\mathbf{pred}(t_q)[V] - \mathbf{R}_t[V]|, \tag{4}$$

where $\mathbf{pred}(t_q)$ is a vector containing the outputs of the ANN for each climate variable V represented by the time series TS or the case, concretely: $\{\text{H}, \text{P}, \text{T}\}$, representing humidity, pressure, and temperature, respectively. R_t is another vector containing the three actual values of a given time window for each climate feature obtained from the explanation case. We calculated the MAEs of the $k = 20$ highest and lowest FFCI values, respectively, which are shown in Table 1.

To determine if the MAEs of the prediction of the ANN against the actual value for each window corresponding to the highest FCCIs are indeed statistically lower than those corresponding to the lowest FCCIs, i.e., if there is a high contrast of the central tendencies of both samples implying that they come from truly distinct populations, we proposed to use the Mann-Whitney U-test. Also called the Mann-Whitney-Wilcoxon, Wilcoxon rank-sum test or simply U-test [25], the Mann-Whitney U-test is appropriate for samples which do not follow a normal distribution. The result of the U-test is the z statistic, which represents a measure of how large the contrast

	Highest FCCIs			Lowest FCCIs	
W	FCCI	MAE	W	FCCI	MAE
6744	1.000	.2923	3996	.4179	.3640
6447	.9500	.2893	1507	.4170	.3982
6388	.9455	.2956	3715	.4170	.4023
6456	.9448	.2936	4685	.4169	.3920
6446	.9380	.3061	3696	.4169	.3909
6104	.9264	.3063	3714	.4163	.3661
6805	.9222	.3118	3246	.4155	.4018
6062	.9211	.2647	14	.4121	.3471
5670	.9204	.2954	1506	.4107	.4454
5781	.9201	.3232	1101	.4105	.3890
1741	.9198	.2888	4011	.4100	.4076
4972	.9198	.2867	3687	.4084	.4004
225	.9195	.3070	3995	.4079	.3385
5825	.9192	.3257	4349	.4030	.3213
5782	.9170	.3302	1099	.4016	.3947
5749	.9170	.3108	3257	.3921	.3761
6410	.9159	.2863	1811	.3840	.3799
5790	.9117	.3262	4809	.3834	.4322
5761	.9105	.2974	3697	.3741	.4167
5762	.9104	.3113	1841	.3521	.3711

Table 1
The values of the $k = 20$ highest and lowest FCCI values, including the corresponding window number (W) and MAE, according to Equation 4.

between the central tendencies of two samples is [26]. Usually z would show a significant difference to the 0.05 level if it had a value of 1.96 and any greater value corresponds to an even greater difference. The z statistic resulting from comparing both the 20 highest and 20 lowest values of FCCI was 5.28, showing a clear difference between these two sets.

5. Conclusions

This paper proposes a RNN for weather forecasting of three climatological variables: humidity, vapor pressure, and temperature. Using 14-day windows extracted from the time series of these variables, the RNN forecasts the corresponding values for the 15th day. More importantly, we also propose a XAI CBR-based methodology to illustrate how in similar cases the RNN yields similar predictions to the actual values of the climatological variables.

The similarity between two different time windows, a query and a fixed case, was calculated with a proposed index, the FFCI, also proposed in this work, which takes into account both 1) the Euclidean distance between the time series and 2) their morphological aspect through the Pearson correlation coefficient.

The 20 cases corresponding to the highest and lowest values of the FCCI, and their respective

15th value, were statistically compared using the Mann-Whitney U-Test, yielding a z statistic of 5.28, well over the 1.96 value commonly associated with a 95% confidence. This means that the 14-day time period prior to a given date has a clear impact on the forecast of the climatological variables of the following day. As future work, an optimization method could be applied to determine the number of days which most affect the forecasting of future variables of climatological variables.

The main conclusion of this work is that XAI CBR-based techniques can be used to determine if there is a statistically significant impact of the past values of climatological variables when trying to forecast future ones. Furthermore, these methodologies could be used to determine the optimal number of previous days which affect the behavior of climatological variables in the future. This can eventually help us to better understand complex processes related to the dynamics of climatology.

Acknowledgments

References

[1] A. J. Jara, Wearable internet: Powering personal devices with the internet of things capabilities, in: 2014 International Conference on Identification, Information and Knowledge in the Internet of Things, 2014, pp. 7–7. doi:10.1109/IIKI.2014.9.

[2] J. Zhang, D. Tao, Empowering things with intelligence: A survey of the progress, challenges, and opportunities in artificial intelligence of things, IEEE Internet of Things Journal 8 (2021) 7789–7817. doi:10.1109/JIOT.2020.3039359.

[3] L. Jia, X. Yang, G. A. Vecchi, R. G. Gudgel, T. L. Delworth, A. Rosati, W. F. Stern, A. T. Wittenberg, L. Krishnamurthy, S. Zhang, R. Msadek, S. Kapnick, S. Underwood, F. Zeng, W. G. Anderson, V. Balaji, K. Dixon, Improved Seasonal Prediction of Temperature and Precipitation over Land in a High-Resolution GFDL Climate Model, Journal of Climate 28 (2015) 2044–2062. URL: https://journals.ametsoc.org/view/journals/clim/28/5/jcli-d-14-00112.1.xml. doi:10.1175/JCLI-D-14-00112.1.

[4] D. Yang, X. Q. Yang, D. Ye, X. Sun, J. Fang, C. Chu, T. Feng, Y. Jiang, J. Liang, X. Ren, Y. Zhang, Y. Tang, On the Relationship Between Probabilistic and Deterministic Skills in Dynamical Seasonal Climate Prediction, Journal of Geophysical Research: Atmospheres 123 (2018) 5261–5283. doi:10.1029/2017JD028002.

[5] D. S. Richardson, Predictability and economic value, Predictability of Weather and Climate 9780521848824 (2006) 628–644. URL: https://www.cambridge.org/core/books/predictability-of-weather-and-climate/predictability-and-economic-value/97F53556B8432DFE6CEECDE1E6E537CF. doi:10.1017/CBO9780511617652.026.

[6] A. K. Palit, D. Popovic, Computational Intelligence in Time Series Forecasting, Advances in Industrial Control, 1 ed., Springer London, London, 2005. URL: http://link.springer.com/10.1007/978-3-319-08413-8. doi:10.1007/1-84628-184-9. arXiv:arXiv:1011.1669v3.

[7] B. Lim, S. Zohren, Time-series forecasting with deep learning: A survey, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 379 (2021). doi:10.1098/rsta.2020.0209. arXiv:2004.13408.

[8] R. Sen, H. F. Yu, I. Dhillon, Think globally, act locally: A deep neural network approach to high-dimensional time series forecasting, Advances in Neural Information Processing Systems 32 (2019) 1–10. arXiv:1905.03806.

[9] N. K. Ahmed, A. F. Atiya, N. El Gayar, H. El-Shishiny, An empirical comparison of machine learning models for time series forecasting, Econometric Reviews 29 (2010) 594–621. doi:10.1080/07474938.2010.481556.

[10] M. G. Orozco-del Castillo, J. Ortiz-Alemán, C. Couder-Castañeda, J. Hernández-Gómez, A. Solís-Santomé, High solar activity predictions through an artificial neural network, International Journal of Modern Physics C 28 (2017). doi:10.1142/S0129183117500759.

[11] B. Lim, S. Zohren, S. Roberts, Recurrent Neural Filters: Learning Independent Bayesian Filtering Steps for Time Series Prediction, Proceedings of the International Joint Conference on Neural Networks (2020). doi:10.1109/IJCNN48605.2020.9206906. arXiv:1901.08096.

[12] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi, A survey of methods for explaining black box models, ACM computing surveys (CSUR) 51 (2018) 1–42.

[13] T. Newton, J. T. Meech, P. Stanley-Marbell, Machine learning for sensor transducer conversion routines, IEEE Embedded Systems Letters 14 (2022) 75–78. doi:10.1109/LES.2021.3129892.

[14] A. Vishnubhatla, Iot based air pollution monitoring through telit bravo kit, in: 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 2022, pp. 1751–1755. doi:10.1109/ICAAIC53929.2022.9793252.

[15] K. A. Hilburn, I. Ebert-Uphoff, S. D. Miller, Development and Interpretation of a Neural-Network-Based Synthetic Radar Reflectivity Estimator Using GOES-R Satellite Observations, Journal of Applied Meteorology and Climatology 60 (2021) 3–21. doi:10.1175/JAMC-D-20-0084.1.

[16] A. Mamalakis, I. Ebert-Uphoff, E. A. Barnes, Explainable Artificial Intelligence in Meteorology and Climate Science: Model Fine-Tuning, Calibrating Trust and Learning New Science, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 13200 LNAI, Springer Science and Business Media Deutschland GmbH, 2022, pp. 315–339. URL: https://link.springer.com/chapter/10.1007/978-3-031-04083-2_16. doi:10.1007/978-3-031-04083-2_16.

[17] E. Delaney, D. Greene, M. T. Keane, Instance-based counterfactual explanations for time series classification, volume 12877 LNAI, Springer Science and Business Media Deutschland GmbH, 2021, pp. 32–47. doi:10.1007/978-3-030-86957-1_3.

[18] T. DelSole, A. Banerjee, Statistical seasonal prediction based on regularized regression, Journal of Climate 30 (2017) 1345–1361. doi:10.1175/JCLI-D-16-0249.1.

[19] Y.-G. Ham, J.-H. Kim, J.-J. Luo, Deep learning for multi-year ENSO forecasts, Nature 573

(2019) 568–572. URL: http://www.nature.com/articles/s41586-019-1559-7. doi:10.1038/s41586-019-1559-7.

[20] B. A. Toms, E. A. Barnes, I. Ebert-Uphoff, Physically Interpretable Neural Networks for the Geosciences: Applications to Earth System Variability, Journal of Advances in Modeling Earth Systems 12 (2020) e2019MS002002. URL: https://onlinelibrary.wiley.com/doi/full/10.1029/2019MS002002https://onlinelibrary.wiley.com/doi/abs/10.1029/2019MS002002https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2019MS002002. doi:10.1029/2019MS002002. arXiv:1912.01752.

[21] K. J. Mayer, E. A. Barnes, Subseasonal Forecasts of Opportunity Identified by an Explainable Neural Network, Geophysical Research Letters 48 (2021) e2020GL092092. URL: https://onlinelibrary.wiley.com/doi/full/10.1029/2020GL092092https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL092092https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020GL092092. doi:10.1029/2020GL092092.

[22] B. A. Toms, E. A. Barnes, J. W. Hurrell, Assessing Decadal Predictability in an Earth-System Model Using Explainable Neural Networks, Geophysical Research Letters 48 (2021) e2021GL093842. URL: https://onlinelibrary.wiley.com/doi/full/10.1029/2021GL093842https://onlinelibrary.wiley.com/doi/abs/10.1029/2021GL093842https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021GL093842. doi:10.1029/2021GL093842.

[23] D. Chakraborty, A. Alam, S. Chaudhuri, H. Başağaoğlu, T. Sulbaran, S. Langar, Scenario-based prediction of climate change impacts on building cooling energy consumption with explainable artificial intelligence, Applied Energy 291 (2021) 116807. doi:10.1016/j.apenergy.2021.116807.

[24] M. T. Keane, E. M. Kenny, How case-based reasoning explains neural networks: A theoretical analysis of xai using post-hoc explanation-by-example from a survey of ann-cbr twin-systems, in: Case-Based Reasoning Research and Development: 27th International Conference, ICCBR 2019, Otzenhausen, Germany, September 8–12, 2019, Proceedings, Springer-Verlag, Berlin, Heidelberg, 2019, p. 155–171. URL: https://doi.org/10.1007/978-3-030-29249-2_11. doi:10.1007/978-3-030-29249-2_11.

[25] H. B. Mann, D. R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, The annals of mathematical statistics 18 (1947) 50–60. URL: http://projecteuclid.org/euclid.aoms/1177730491.

[26] M. G. Orozco-del Castillo, C. Ortiz-Aleman, J. Urrutia-Fucugauchi, R. Martin, A. Rodríguez-Castellanos, P. Villaseñor-Rojas, A genetic algorithm for filter design to enhance features in seismic images, Geophysical Prospecting 62 (2013) 210–222. URL: http://doi.wiley.com/10.1111/1365-2478.12026. doi:10.1111/1365-2478.12026.