 doi:10.20944/preprints202205.0147.v1

Article

Predicting Climatic Variables on Local Level with SARIMA, LSTM and Hybrid Techniques

Antonios Parasyris[1,*], George Alexandrakis[1], Georgios V. Kozyrakis[1], Katerina Spanoudaki[1] and Nikolaos A. Kampanis[1]

[1] Foundation for Research and Technology-Hellas, Institute of Applied and Computational Mathematics, Heraklion, Crete, Greece
* Correspondence: antonisparasyris@iacm.forth.gr

Abstract: The choice of holiday destinations is highly depended on climate considerations. Nowadays, since the effects of climate crisis are being increasingly felt, the need of accurate weather and climate services for hotels is crucial. Such a service could be beneficial for both the future planning of tourists' activities and destinations and for hotel managers as it could help in decision making about the planning and expansion of the touristic season, due to a prediction of higher temperatures for a longer time span, thus causing increased revenue for companies in the local touristic sector. The aim of this work is to calculate predictions on climatic variables using statistical techniques as well as Artificial Intelligence (AI) for a specific area of interest utilising data from in situ meteorological station, and produce valuable and reliable localised predictions with the most cost-effective method possible. This investigation will answer the question of the most suitable prediction method for time series data from a single meteorological station that is deployed in a specific location. As a result, an accurate representation of the microclimate in a specific are is achieved. To achieve this high accuracy in situ measurements and prediction techniques are used. As prediction techniques, Seasonal Auto Regressive Integrated Moving Average (SARIMA), AI techniques like the Long-Short-Term-Memory (LSTM) Neural Network and hybrid combinations of the two are used. Variables of interest are divided in the easier to predict temperature and humidity that are more periodic and less chaotic, and the wind speed as an example of a more stochastic variable with no known seasonality and patterns. Our results show that the examined Hybrid methodology performs the best at temperature and wind speed forecasts, closely followed by the SARIMA whereas LSTM perform better overall at the humidity forecast, even after the correction of the Hybrid to the SARIMA model.

Keywords: SARIMA; Artificial Neural Networks (ANN); LSTM; hybrid methodologies; prediction

1. Introduction

Identifying and predicting the climatic characteristics can have a great impact on many facets of everyday human life and it is known to affect the mood and the choice of everyday activities. For applications like photovoltaic and wind energy prediction, climate change scenarios forecasting, extreme events, etc. various techniques have been used, see for example [1] where the authors review the abundant literature of research papers related to machine learning methods and Numerical Weather Prediction based on the application.

Weather prediction, the application that will be in the scope of our research is associated with tourists' satisfaction based on the climatic conditions. For the tourism industry, weather is a key factor [2]. The choice of a holiday destination is highly dependent on climate considerations, and so, as the climate crisis repercussions are being increasingly felt, the need of climate prediction services for hotels becomes a necessity, both for tourists and hotel managers. Existing applications of weather forecast that involves in situ information from an MS aims at current comfort indices evaluation and heritage risk analysis, like e.g., [3], but do not consider simplified weather forecasts using accurate in situ real-time datasets. This service directed to touristic comfort and satisfaction based on the climate is being overlooked, and the usual source of information that many hotels offer involve private meteorological companies like Accuweather, WeatherNet, Meteo, etc. [4]. These services typically use models with limited accuracy and low spatial resolution, they may experience downtimes, or even require subscription fees from the tourist for more premium and robust functionalities, something that can have a negative impact on a tourists' experience. An example from literature in [5] that uses Random Forests to predict the solar radiation with a high-density meteorological station grid cannot be utilized here as the plurality of data observation locations and the lack of accuracy make it a non viable option for operational purposes. The decision of which predictors that can be exploited and used with these kinds of datasets for accurate weather prediction is identified as the research gap that this paper addresses.

Our research develops and compares prediction methods that are suitable to be applied on a small and affordable scale, such that even a small scale hotel can afford the installation and maintenance for the forecasting system locally, without dependence on expensive models and subscription fees of forecasting services. Thus, the problem involves time series prediction, which poses a challenging problem on its own [6], but also focused towards climatic variables, which tend to exhibit chaotic behaviours and cannot be easily forecasted with a single point source time series [7]. Additionally, in statistical forecasting approaches, the huge downside is that stochasticity is present in the sense that there is not one single theoretical model that can outperform all methods in any given situation; see for example [8, 9]. Each method's performance is heavily reliant on the given dataset, nature of the problem, stationarity of the data, noisy data, trend, cyclicity of data etc. This also gives rise to a very active field of hybrid methodologies and new techniques being invented, that suit better some problems whilst they underperform in others. For that reason, and for the problem of short-term climatic variables prediction, we will test and employ different techniques that are proposed in the literature for our given dataset and variables.

First and foremost, the most used and researched methodology for time series forecasting is the Auto Regressive Integrated Moving Average (ARIMA) model [10, 11], where there exists abundant literature to support the use for short time prediction [[6]], even for climatic variables that tend to exhibit chaotic behaviour [12, 13]. The authors of [14] describe the prediction of solar radiation using ARIMA models, where they found good agreement, even for the high accuracy needed for photovoltaic systems' energy production. They also note that as the prediction window is enlarged (more than a day), the errors accumulate so much that the resulting accuracy is reduced significantly. Another useful remark they make about ARIMA implementation on climate datasets is that a differencing of first order is crucial to remove the non-stationarity and hence improve the accuracy of the predictions. If longer time series with multiple periodicities are available, Seasonal ARIMA (SARIMA) models are better suited and show a better fit [6].

When datasets comprising by nonlinear time series are available consistently for many years, Artificial Neural Networks (NN) are better suited for forecasting [15]. NN fall under the broader category of Machine Learning (ML) techniques and possess the ability to improve automatically through experience and by the use of data, as can be found in many ML textbooks, see e.g. [16]. NNs are seen as a part of Artificial Intelligence (AI), where the general idea is to use a sample data known as "training dataset" to make predictions or decisions without being explicitly programmed to do so. These have many applications like email filtering, computer vision, data mining and optimization and are powerful nonlinear regression techniques inspired by theories of how the brain works [16].

ARIMA and its seasonal counterpart SARIMA is found in numerous studies e.g., [6, 12, 13, 17-19] to be well suited to tackle problems like the 1-2 day ahead prediction of some of the weather variables of interest. In order to be useful in more applications with larger prediction horizons and more chaotic weather variables, combinations and hybridizations

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2022 doi:10.20944/preprints202205.0147.v1

based on ML and SARIMA are suggested by authors e.g., [19-24] and the references therein. The task of creating robust and reliable hybrid techniques that improve the accuracy of time series prediction algorithms is one and maybe the most researched question in the field and it will be discussed more in depth in the methodology and discussion section later on. Another aspect of this research is to be able to accurately depict weather and climate phenomena locally near the point of interest, which for this application is a hotel in Crete. The data sources will be discussed in the data section below.

The paper is structured as follows: The prediction methods that will be considered are introduced and discussed in the next section. Namely the more classical time-series prediction that is the SARIMA method, Artificial Intelligence techniques like the LSTM Neural Networks and Hybrid combinations as indicated by the literature. Next, we introduce our approach on a hybrid method, which is implemented for the examined dataset. In the last sections, we compare the results of the three models, for three different climatic variables and two short term prediction horizons. Finally, we conclude with our preferences and comments on other methods and different hybridization techniques that could further assist and improve the predictions.

2. Materials and Methods

In this section, we will briefly describe the methods we will utilize regarding the short-term forecasting of the weather conditions at a location of interest. We will also describe the details of each methodology that were essential for its implementation.

2.1. Data

For the accurate depiction of the weather and climate near the point of interest, which for this application is a hotel in Crete, a Meteorological Station (MS) has been installed in place and maintained by the Coastal & Marine Research Laboratory (CMRL) https://crl.iacm.forth.gr/en/ and is able to provide a continuous feed of data. The location of the hotel in which the station has been installed on is chosen as a representative example of the northern area of Crete, a highly significant socioeconomic centre which is heavily financially dependent on the touristic sector. More specifically, the MS is installed on an open field on a rooftop in an optimized position in order to assess the microclimate of the area as seen in Fig. 1. The area is also selected with respect to its representativeness of the northern Cretan microclimate.

Figure 1. The Meteorological Station installed on a hotel rooftop in Crete, Greece.

The MS records in a minutely temporal resolution, so that 3-hour averages or maximum-minimum values have been obtained. The variables used here as inputs or outputs include combinations of the maximum-mean air temperature (Celsius), minimum-mean relative humidity (%), mean barometric pressure, total precipitation (mm) and the average wind speed (m/s) and direction. The MS has been fully operational since 2019 and therefore the amount of data was deemed insufficient for a fair comparison of the SARIMA, the LSTM and the Hybrid to be carried out. To circumvent this and achieve a methodological comparison based on the results, another dataset was provided by the Hellenic National Meteorological Service (http://www.emy.gr/emy/el/services/paroxi-ipiresion-elefthera-dedomena) which comprised by measurements from a MS similar to the one installed by authors, located in a close by area with similar climatic characteristics. This dataset comprised of the same variables, measured at a 3-hour temporal resolution, spanning the years 1975-2004. In this way, the methods investigated could be fairly compared and a proof of concept could be made which will be applied on the in-situ data from our MS in the future. The training of all methods will be carried out in a fraction of this dataset, as will be discussed on the methodology section later on.

The selection of these variables has been based on studies that associate tourism applications to climatic variables that can be found for example in [25, 26] and more specifically for the case of Greece where tourism depends on summertime beach activities, in [2] and references therein.

2.2. ARIMA methodology

The methodology for fitting a ARIMA model into our data can be easily handled in R software by the "*auto.arima*" function [27, 28]. It utilizes the Box-Jenkins algorithm [27, 28] to fit an ARIMA model into the training dataset. To do that, it searches for a triplet [p,d,q] that corresponds to [terms of Auto Regression AR, degree of differencing, terms of moving Average (MA)] respectively which will make the model fit the given training dataset with a minimized error. After the optimum triplet has been found, the model can make predictions of a specified length using the "predict" function embedded in R, and can also compare the dataset in which it has been trained on with the predicted values that the model would output. This is what is called herein the ARIMA residual, or simply a residual.

One question that this paper addresses is the relevancy and the usefulness of each predictor for the application at hand. For that reason, the next section briefly introduces

 doi:10.20944/preprints202205.0147.v1

the Artificial Neural Networks (ANN) as a predictor and later we will compare results from both methods as well as hybrid combinations of them.

2.3. *Artificial Neural Networks (ANN)- LSTMs*

Neural Networks are known to thrive in noisier and non-linear data where other methods fail, in contrast with the ARIMA method that is famously capable of handling linear time-series prediction. One issue faced for long time series is that even small errors can propagate and result in catastrophic failure of the convergence of the model [29]. Thus, we will use a special category of recurrent ANNs called Long-Short-Term Memory (LSTM) Neural Networks that are shown to be the most suited of all the Neural Networks for time series prediction [29, 30]. The suitability of the LSTMs to forecast long time series is explained as they are constructed to remember using long term memory cells, which can remember information that a simple NN cannot, and at the same time, it is able to solve vanishing gradient issues that the standard Recurrent Neural Networks suffer from and may hinder performance. For example, we see in [31] a comparison of ARIMA and LSTM applied to wind speed, that showed the superiority of LSTM, but noted that for smaller datasets, the ARIMA technique might outperform the LSTM due to the lack of training and pattern learning of the specific NN architecture.

For each of the LSTM Neural Networks shown in the results section, we conducted a grid search of parameters and architectures to determine a near optimal configuration. For some networks, the use of SGD (stochastic gradient descent) [32] optimizer has better results (standard momentum=0.9 is used), whereas in some other examples, the use of Adam optimizer [33]. The learning rate is set to the standard 0.01, with an update of 0.005 and a "patience" parameter equal to 5, meaning that the algorithm updates the learning rate every 5 epochs that the loss does not improve.

Usually, we propose 40 units in each layer but some tests had better results with 100 units, with the deterioration of the computational time. Most often we obtained good accuracy with one input layer, 2 hidden LSTM layers and one dense as output layer. All layers use a Rectified Linear Unit (ReLU) activation function, as tests with others showed deterioration in accuracy. The epochs of training ranged from 20 to 60, and a combination of the grid search and the validation error helped us to avoid over fitting. Additionally, a callback that monitored the loss and had a patience of 5 epochs reduced the learning rate starting from 0.01 by a factor of 0.005. The Neural Network is designed by having a week of previous data as input (56 3-hour samples) and either a 1-day or 2-days ahead forecast horizon (8 or 16 3-hour samples). As loss error, the mean squared error (MSE) was picked, whereas for comparison between the methods and easier translation to the real world problem, the Mean Absolute Error (MAE) was employed. For each day or 2-days ahead in the test set, a prediction was made and compared with the actual data which was not included in the training process of any algorithm. This was achieved by taking the absolute value of their difference and then the mean for each day. Then, another average over all the days of the test set correspond to the value shown in Table 1 for the three variables of interest, namely temperature, humidity and wind speed. We also use all the other correlated variables with the prediction variable in each run, and they are included as features. For example, for temperature prediction, the humidity, wind speed, and pressure were used, and similarly for the other 2 networks we created for humidity and wind speed. Also some sine and cosine waves that oscillated to show the daily and yearly frequency were also used as features. In conclusion, our LSTMs had 8 features (4 weather related and 4 combinations of sine-cosine for the daily-yearly cycle).

For the introduction of hybrid methodologies in the next section, the assumption is that ARIMA residuals exhibit this kind of nonlinear noisier behaviour, and thus a combination of the two methodologies in which Neural Networks are used to predict the future ARIMA residual and add it to the ARIMA forecast is suggested. This is expected to result in a predictor-corrector method that can in principle improve accuracy.

2.4. ARIMA-LSTM hybrid methodology

As mentioned in the introduction, a significant amount of research is spent on exploring new hybrid methods in a try to increase accuracy, but this is a problem dependent question. On similar applications, Saba et al. [23] worked towards a hybrid method for weather prediction where they averaged Multi Layered Perceptron and Radial Basis Function Neural Network outputs and succeeded in reducing the error with the simplification that their output was just a decision between a rainy or dry weather. This is much simpler than actually predicting multiple time steps of values of every single variable, and can achieve higher accuracy due to this simplification. Also it is worth mentioning that the authors used a relatively small dataset of 5 years of data for both training and testing, which could be a reason they had to resort in this kind of simplified prediction.

The Hybrid method that we will use is inspired by the work of Deng et al. [24] where ARIMA is used as a primary predictor, and subsequently a NN is used to correct the result by predicting the residual with an LSTM. Subsequently, the authors improve even further their results by using a back-propagation algorithm to optimize the way the two predictions are combined. Their results indicate an increased accuracy of the predictions on outpatient visits using their proposed technique, and it is suggested that the application of the method could be extended into more challenging datasets, such as weather. All tests were performed in R using "keras" under "tensorflow".

3. Results

We split the presentation of the results into two categories, a) predictions of temperature and humidity, which are easier to achieved and b) the more chaotic wind forecasts. The computational time for each NN epoch was in average about 3 minutes (depending on the units used) so for a 40 epoch LSTM with 40 units at each of 2 hidden LSTM layers, which was a common configuration we used, 2 hours were spent for training each LSTM. On the other hand, for the SARIMA fit we report an average time of 3 minutes for each fit, which is significantly faster than any of the Neural Networks. We note here that for the 1-day or the 2-days forecast, the same SARIMA model is used, so one does not have to compute again the SARIMA fit for bigger prediction horizons, although the accuracy is known to significantly deteriorate [15] as bigger prediction horizons are considered.

3.1. Temperature and Humidity forecast

For the 1 day prediction, we fit the ARIMA model with a seasonal component (SARIMA) as well to our 3-hour temperature data, which is the 70% of the total data (out of 61363 total measurements, 42954 are used for fitting the SARIMA which results in the ARIMA(1,0,2)(2,1,2) [8] where 8 is the time series frequency (daily), and (2,1,2) are the [P,D,Q] that correspond to the seasonal component amount of terms, again for [terms of Auto Regression AR, degree of differencing, terms of moving Average (MA)]

To obtain a fair comparison with the standard LSTM model, a standard 70% training, 15% validation and 15% training split is employed at every instance where a Neural Network is needed. Additionally, a scaling in the range of (0, 1) is performed. To get rid of the stochasticity of the results that is inherent to Neural Networks, each Mean Absolute Error (MAE) presented in Table 1 acts as a representative through means of averaging, of multiple runs with the same parameters as well as many days over the test set. Table 1 shows the MAE of the predictions in 1-day and 2-day horizons, in which we observe how small the error of the Hybrid and SARIMA model is, compared to what we accomplished with the LSTM, except in the case of Humidity, where the LSTM shows an improvement from both methods.

Table 1. MAE errors for the 3 methods for 1 and 2-day prediction of wind speed (m/s) for the three predicted variables, averaged over each day and then over the test set.

Prediction Variable	Methods	MAE 1-day	MAE 2-day
Temperature (C)	SARIMA(1,0,2)(2,1,2)	1.58	1.86
	LSTM (Adam 40 units each of 2 hidden layers)	2.12	2.14
	Hybrid	1.56	1.85
Humidity (%)	SARIMA(5,1,0)(2,0,0)	10.62 %	12.33%
	LSTM (Adam-40 units each of 2 hidden layers)	9.54 %	10.01%
	Hybrid	10.30 %	12.01%
Wind Speed (m/s)	SARIMA(0,1,5)(0,0,2)	2.46	2.78
	LSTM (Adam 60 units each of 2 hidden layers)	2.73	2.79
	Hybrid	2.41	2.70

Furthermore, a small reduction of the error is achieved through hybridization, which is normal since this depends significantly on the SARIMA fit and the algorithm performs an improvement upon the residual errors of the SARIMA. Also, we note that as the prediction window increases to 2-days, the gap between the methods becomes smaller, especially in the case of Wind speed prediction. In case of Humidity, the error reduction of the Hybrid does not make it better than the LSTM, which is 2% better in the 2 day predictions. For the temperature predictions, the Hybrid shows the lowest average MAEs, but it is very closely followed by the SARIMA.

The Daily averaged MAE for the 1 and 2-day prediction horizon is presented in Figures 2 and 3, focused on the temperature in this section, for 230 days of the test set. We present MAEs for the three different methods which are color coded, and with the same color and dashed lines we plot the average of all the days in the test set as calculated in Table 1. As we can see in Fig. 2 the lower values of the Hybrid are followed closely by the SARIMA, and the worst method for the 2-day prediction is the LSTM. Another useful takeaway from Figures 2 and 3 is that there is a significant error at the 150-day mark, and since our test set begins on November, all the first 150 days are lower values of true temperature, whereas after the 150 days' mark, higher temperatures are consistently observed in the dataset as it is usual of summer months. Hence, we see all methods to fail predicting the change of weather, but when higher temperatures are inputted afterwards to the SARIMA and Hybrid model, the summer months can also be predicted well. The LSTM seems to have larger errors in those last 2 summer time months in our test set compared to the previous winter months. We will demonstrate that by plotting individual days from the dataset in the next figures.

 doi:10.20944/preprints202205.0147.v1

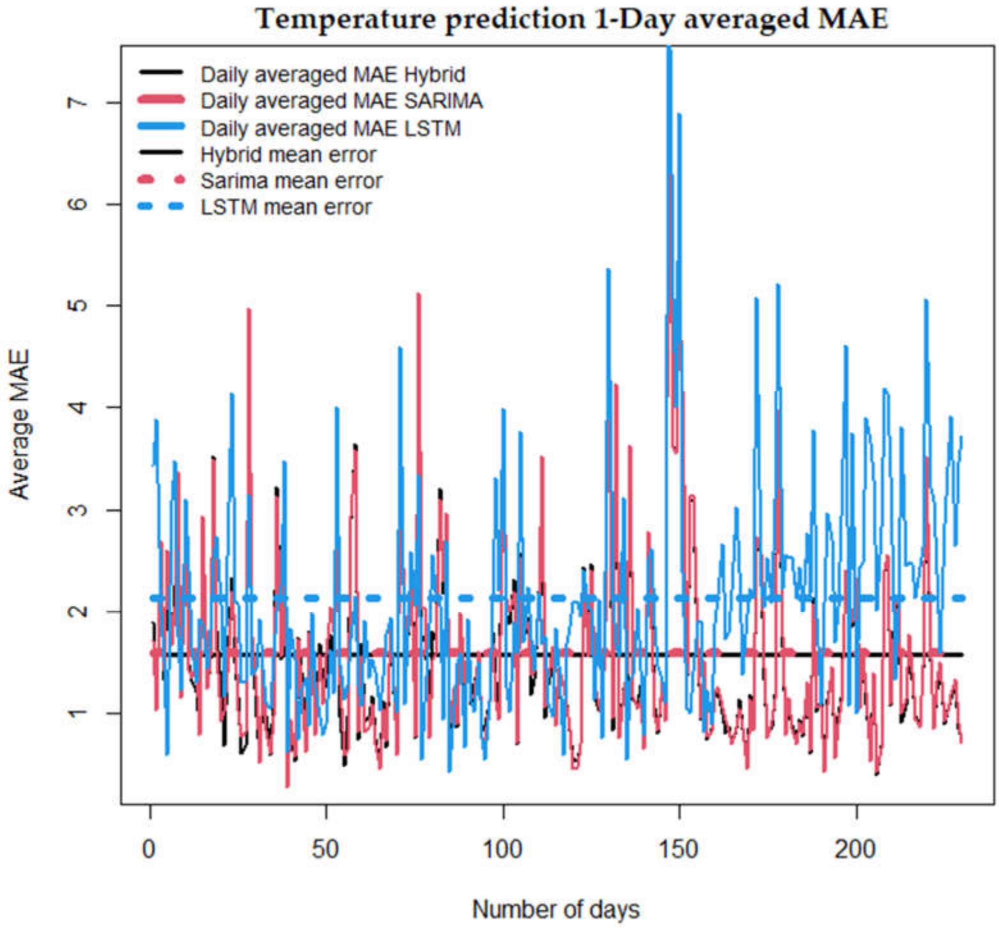

Figure 2. Comparison of SARIMA, LSTM and Hybrid using daily averaged MAE with real data for 1-day prediction horizon for 230 days of the test set. With dashed lines the average error of all the dataset is shown.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2022 doi:10.20944/preprints202205.0147.v1

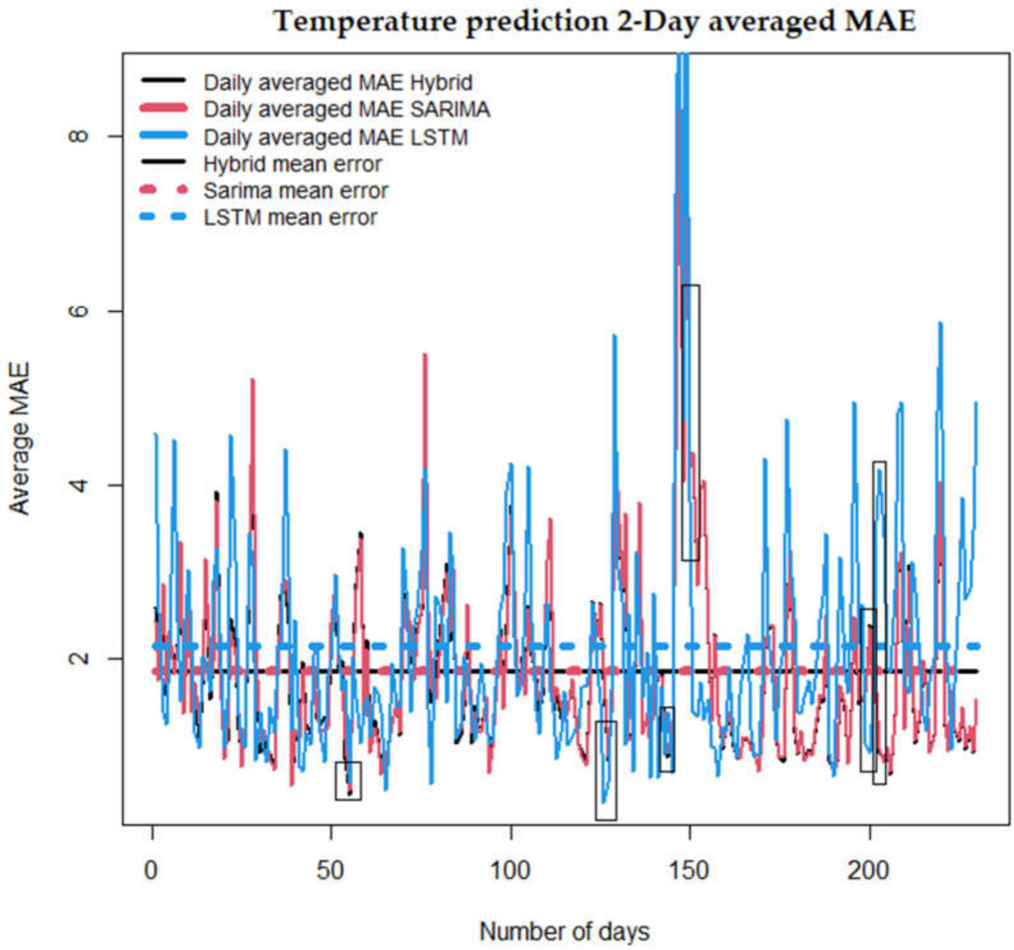

Figure 3. Comparison of SARIMA, LSTM and Hybrid using daily averaged MAE with real data for 2-day prediction horizon for 230 days of the test set. With dashed lines the average error of all the dataset is shown. The boxes indicate the predictions that we plot separately to observe their 48-hour behavior.

The days that are individually chosen to be showed in the next figures are selected based mainly on the Daily MAE of Figure 3, which are indicated with small squares in that graph. It is either extreme cases of small/big MAE errors, or instances where the preferred method is not what would be expected from the average values shown in Table 1. The goal here is to examine when the methods fail and it would be extremely useful to determine if there is a causal effect that link the weather conditions with the best predictor. Then, one could inform a posteriori, depending on the climate conditions of the present or previous days on which method could predict the future more accurately, and thus construct another Hybrid technique that would by definition lower the average MAE by choosing the method of least MAE for every days' forecast.

In Figure 4 we present an example of a prediction day from the test set, which is in January, characterized by low temperature of 9 degrees night temperature, and low range that reaches 11.2 degrees during the day. This is not unusual during the winter months, and we see great agreement between all methods. For that example of a 2-day prediction horizon, the calculated MAEs are: LSTM=0.87, SARIMA=0.48 and Hybrid=0.43, which is consistent with our average results that indicate the Hybrid has a better fit in the test dataset overall. Nevertheless, we see in Fig. 4 that the highest values of the second day are very well approximated by both the SARIMA and the Hybrid, with the hybrid being a correction to the SARIMA during the first day, and under predict during the higher values of the second day. Other days the correction of the hybrid is not so pronounced, but on average, as the Table 1 suggests, it is a useful improvement over the standard SARIMA

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2022 doi:10.20944/preprints202205.0147.v1

model. Another meaningful observation of Fig. 4 is that the LSTM might have severely under predicted the low of the first day, but reached the higher values that occurred during noon (at 12 and 39 hours ahead in Fig 4). It is worth commenting that this example has not a typical parabolic shape that daily temperature profiles usually exhibit, and that the small range plays a big role in the low MAE that is found. As we see in Figs. 2-3, in the summer months when the daily range is much higher and the maximum daily temperatures can exceed 40 degrees, the average MAE increases, especially with the LSTM method.

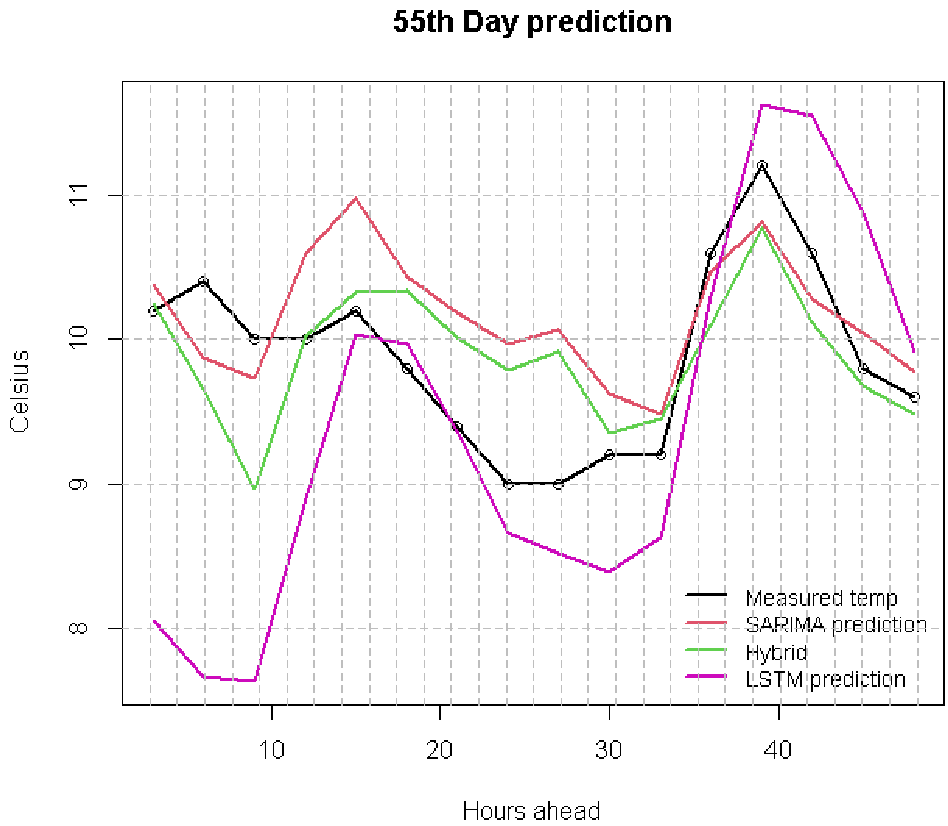

Figure 4. Comparison of SARIMA, LSTM and Hybrid with real data on temperature example for 2-day prediction horizon in January.

In the following figure (Fig. 5), that is an example during March, we choose to show another 2-days forecasting, which lie in the second box of Figure 3, that show a very good agreement with the LSTM method, the inverse result of the averaged outcome of Table 1. Indeed, looking at Fig. 5 and evaluating that the MAE in that day we find LSTM=0.75, SARIMA 1.87 and Hybrid 1.76. The range here is normal, from a low of 11.3 to a high of 19 degrees and the parabolic pattern of temperature is stable between the two days, with consistent extremes. Between the SARIMA and the Hybrid, we see that the Hybrid prediction increases the under predicted values of the SARIMA by a small margin, which constitutes an improvement.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2022 doi:10.20944/preprints202205.0147.v1

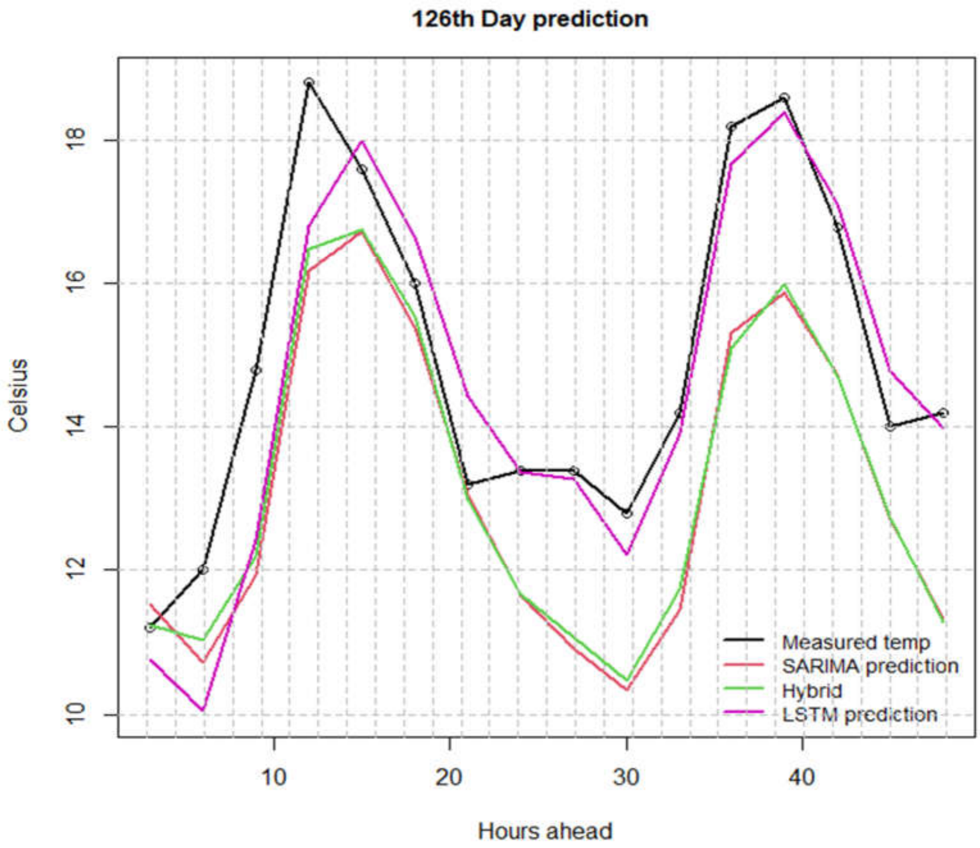

Figure 5. Comparison of SARIMA, LSTM and Hybrid with real data on temperature example for 2-day prediction horizon during March; stable parabolic temperature profile.

In Figure 6, which is during the same month as Fig. 5, we observe again that having a stable temperature profile, all three methods give accurate results with a reported MAE of LSTM=1.02, SARIMA=0.82 and Hybrid=0.87. In this case, the LSTM is worse than the other two methods, but the reason this example is interesting is that the Hybrid is not a correction of the SARIMA anymore. The residuals prediction informs that the SARIMA has over predicted, where the SARIMA in reality under predicted the measured temperature.

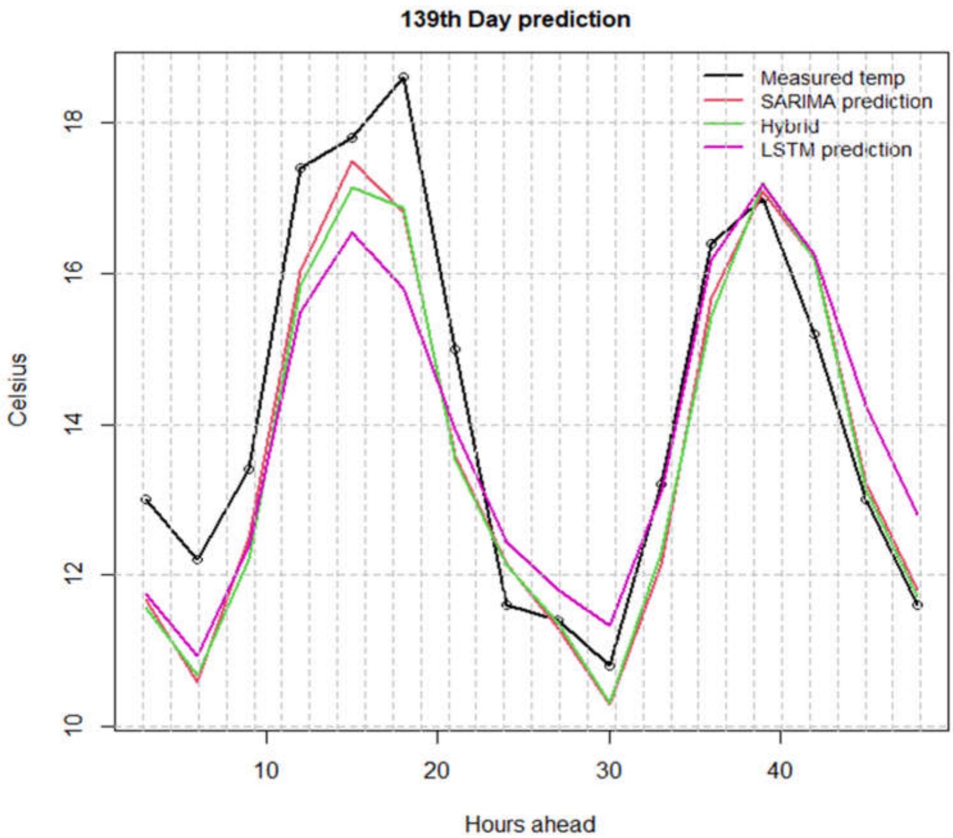

Figure 6. Comparison of SARIMA, LSTM and Hybrid with real data on temperature example for 2-day prediction horizon in March; stable temperature example where Hybrid is not achieving a correction on the SARIMA.

In Figure 7 we present a case of during April, which as seen in Figure 3 at the 150th day mark, is an area where the biggest MAE are found. Thus, the worst predictions occur not randomly, but at periods where big ranges exist, and huge shift between 2-3 days can occur. In this case, the highest value at 34.2 degrees followed by a low of 21 at the same day, and a low of 18 the next day renders all three methods unreliable for the specific example. The calculated MAE for the example is LSTM=5.02, SARIMA=4.19 and Hybrid=4.18, which does not represent an acceptable forecast.

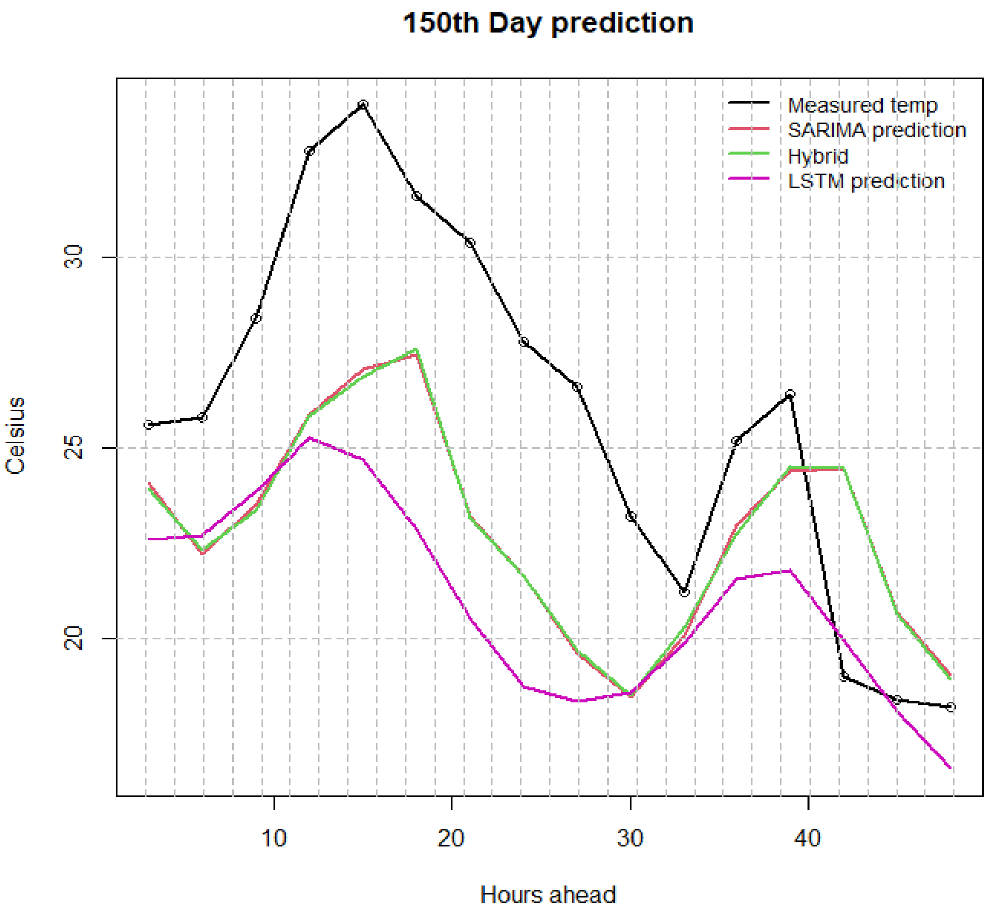

Figure 7. Comparison of SARIMA, LSTM and Hybrid with real data on temperature example for 2-day prediction horizon in April; Example with very big MAE.

In Figure 8, a day in June is presented (5th box in Figure 3) where the temperature range is also high, and the first days' high is not well predicted, but the second day is very well predicted, especially by the SARIMA and the Hybrid. More specifically, the 2 day MAE for Fig. 8 are LSTM=3.25, SARIMA=1.53 and Hybrid=1.48. The Hybrid increases accuracy upon the SARIMA, and the poor LSTM fit is also reflected by the MAE error which is more than double of the Hybrid MAE. These are the summer months we referred to earlier, where the daily temperature range is more than 10 degrees Celsius and that we see a clear advantage of the SARIMA, and its corrected version in the Hybrid method.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2022 doi:10.20944/preprints202205.0147.v1

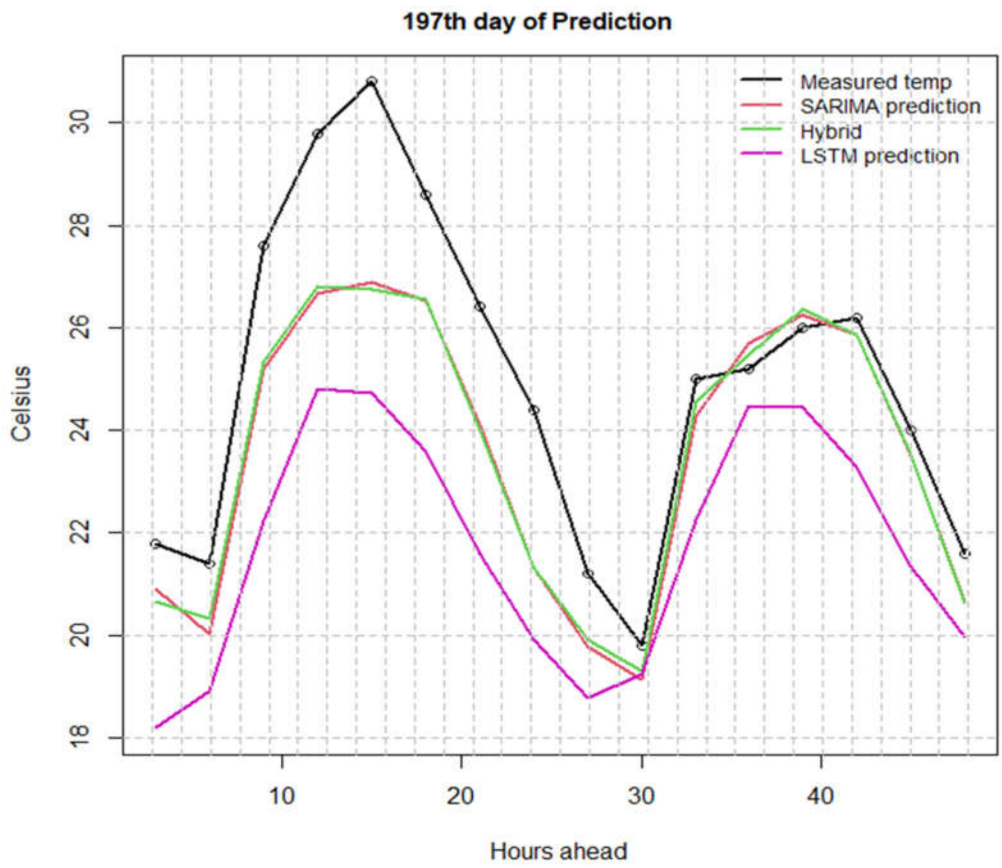

Figure 8. Comparison of SARIMA, LSTM and Hybrid with real data on temperature example for 2-day prediction horizon in June; Example with very big MAE.

In summary, in Figures 4-8 we see that both the ARIMA and the LSTM can predict relatively well the temperature of the next two days, similar to the results of Table 1. The interesting outcome that is not reflected in the averaged values of Table 1 is that the question of the best method for each day has not a clear answer every day, since for most days the SARIMA and Hybrid are best suited, but also some other days the LSTM has lower MAE. This result raises the question of the weather characteristics that may or may not affect the comparison between the predictors for different days. Our analysis and explanation of the individual results of Figures 4-8 indicate that the range of the temperature might be a parameter that could inform such a decision.

For the relative humidity, we see in Table 1 we see that the LSTM outperformed both methods, followed by the Hybrid, which again corrected the SARIMA by a small margin. The fact that LSTM outperformed and decreased the margin in all examples in the 2 day forecast is consistent with the authors of [15] that note that the more days we try to forecast, the worse the SARIMA method performs and the better predictors are the deep Neural Networks. This affects the performance of the hybrid method, which is strongly related to the performance of the SARIMA.

In Figure 9 we repeat the errors computed in Figure 3 for humidity, and we see in this case that there is a peak at 150-day mark, similar to the temperature example, but not the MAE after that threshold do not increase for the LSTM model, which keeps a more consistent error magnitude throughout.

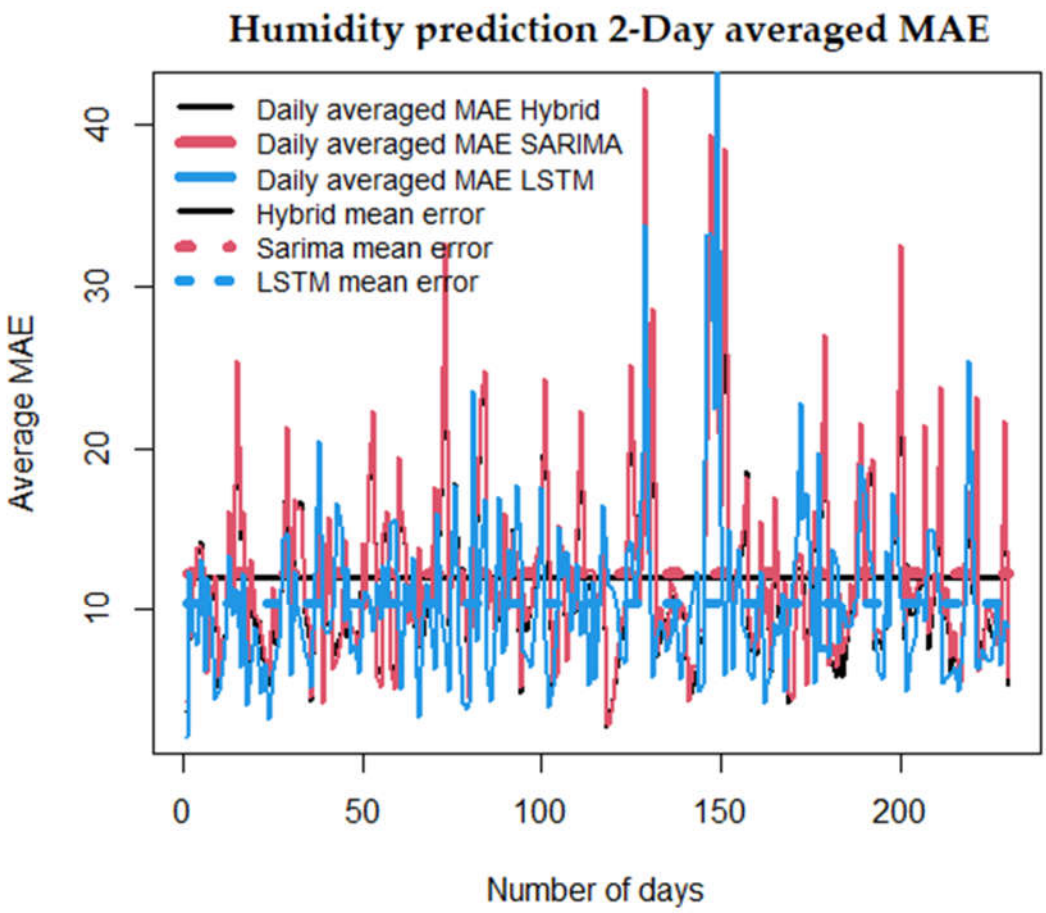

Figure 9. Comparison of SARIMA, LSTM and Hybrid using daily averaged MAE with real data for 2-day prediction horizon on humidity prediction for 230 days of the test set. With dashed lines the average error of all the dataset is shown.

Finally, in Figure 10 we show an example in June with predictions that achieved an adequate performance from all three methods. Indeed, we calculate the 2 day MAE for LSTM=8.09%, SARIMA=9.19% and Hybrid=8.01% for that example.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2022 doi:10.20944/preprints202205.0147.v1

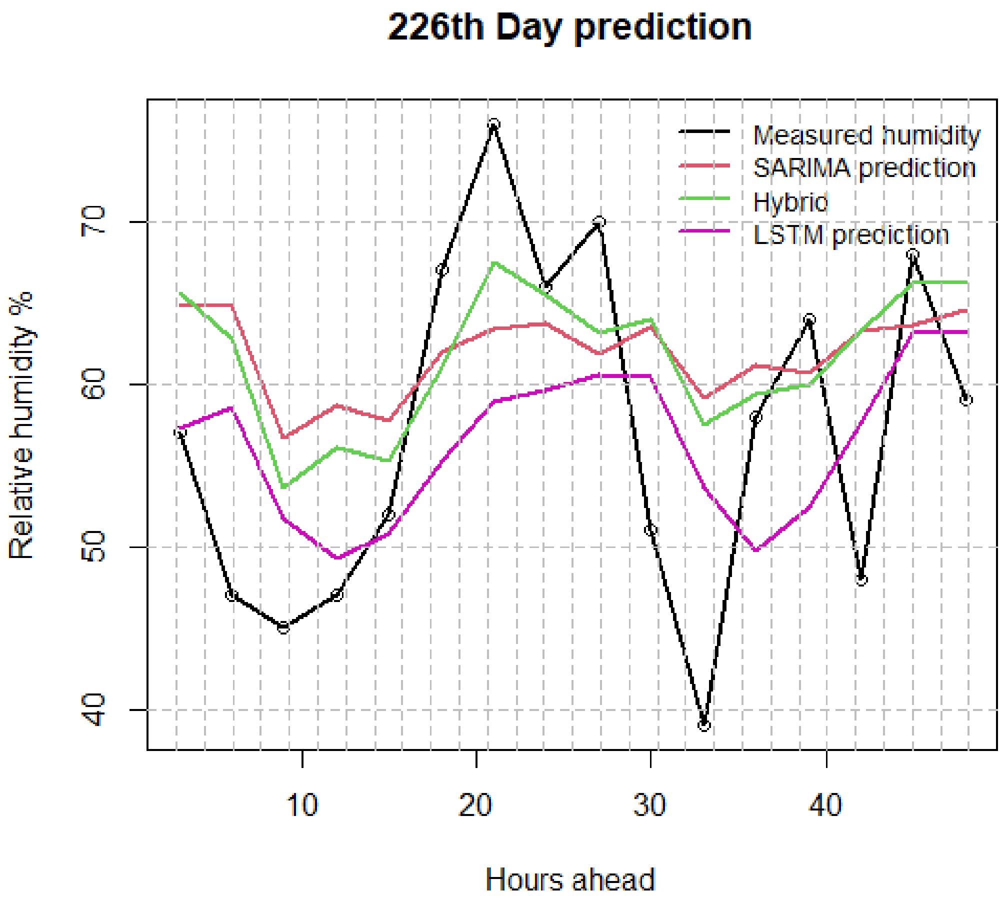

Figure 10. Comparison of SARIMA, LSTM and Hybrid with real data on humidity example for 2-day prediction horizon in June; Example with very big MAE.

Due to similarity of the discussions and conclusions of humidity predictions with the temperature prediction in figure form, additional figures are omitted.

3.2. Wind speed forecast

Lastly, a presentation analogous of the previous section is being done here for the wind speed (m/s), which is considered the more chaotic variable of the three, with no obvious daily pattern like the temperature/humidity pair.

As we see in Table 1 in the wind speed section, the Hybrid has the best average MAE among both SARIMA and the LSTM, but the correction of the Hybrid over the SARIMA was by a small percentage of 3%. Additionally, the 2-day forecast showed a very small margin between the errors compared to the 1-day forecast horizon.

Similar to before, we see in Figure 11 the daily averaged MAE for the 2-day wind prediction of the first 230 days in the test set. Unlike before, we do not see an increase at the 150th day mark and afterwards, the error of the tail is significantly smaller than the first part. This can be related to the higher wind speeds and weather changes which are more prominent in the first part than in the latter.

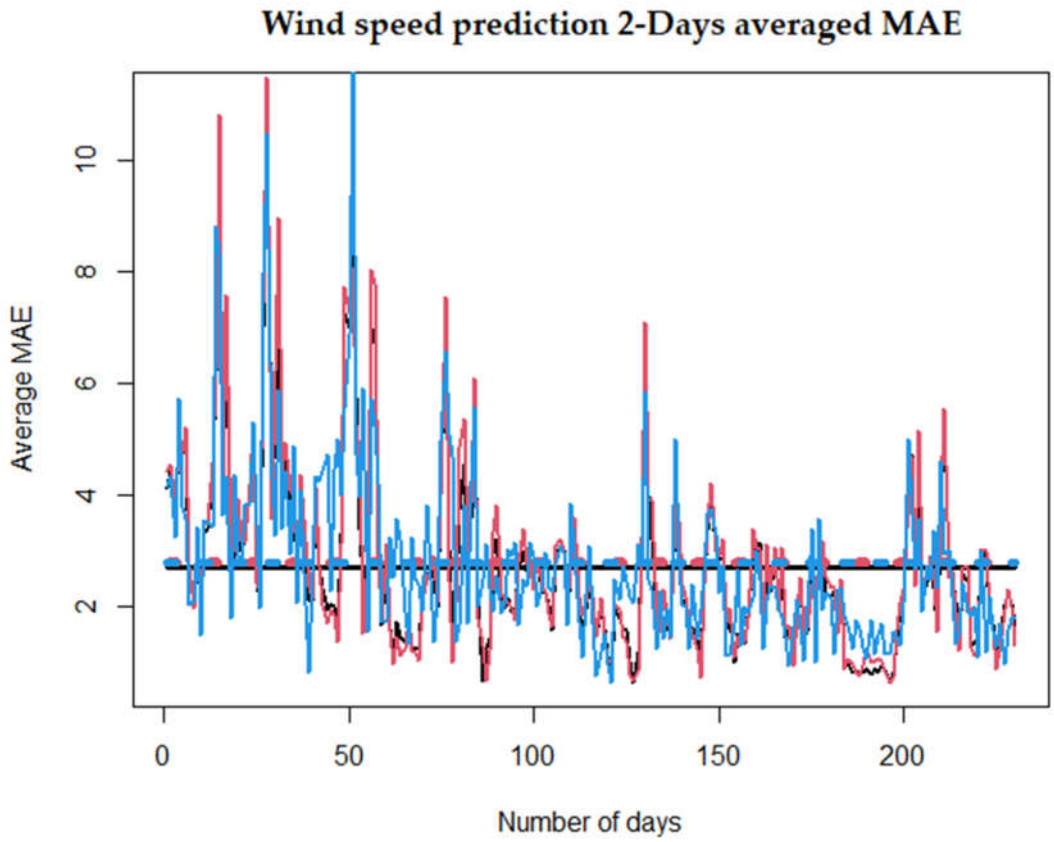

Figure 11. Comparison of SARIMA, LSTM and Hybrid using daily averaged MAE with real data for 2-day prediction horizon on wind speed prediction for 230 days of the test set. With dashed lines the average error of all the dataset is shown.

Additionally, we see in Figure 12 that all methods have a difficulty at capturing the zero speed wind that is a possibility in wind prediction, and also they fail to predict the increase of the second day wind speeds at an extreme phenomenon, but the methods provide an estimate on the next 2 days' wind, with a clear advantage on the LSTM. Indeed, we calculate the 2 Day MAE of Fig. 12 to be LSTM= 2.18, SARIMA=4.34 and Hybrid=3.97. The hybrid is a correction over the SARIMA in this case, although it is not better than the LSTM, since the Hybrid is heavily reliant on the SARIMA prediction. Lastly, we note that the case of Fig. 12 is an example where low wind speeds have been reported during the previous days, which explains why the SARIMA and Hybrid are extremely low and the LSTM has a zero prediction at 6 hours ahead, although it is significantly better at predicting the rest of the hours.

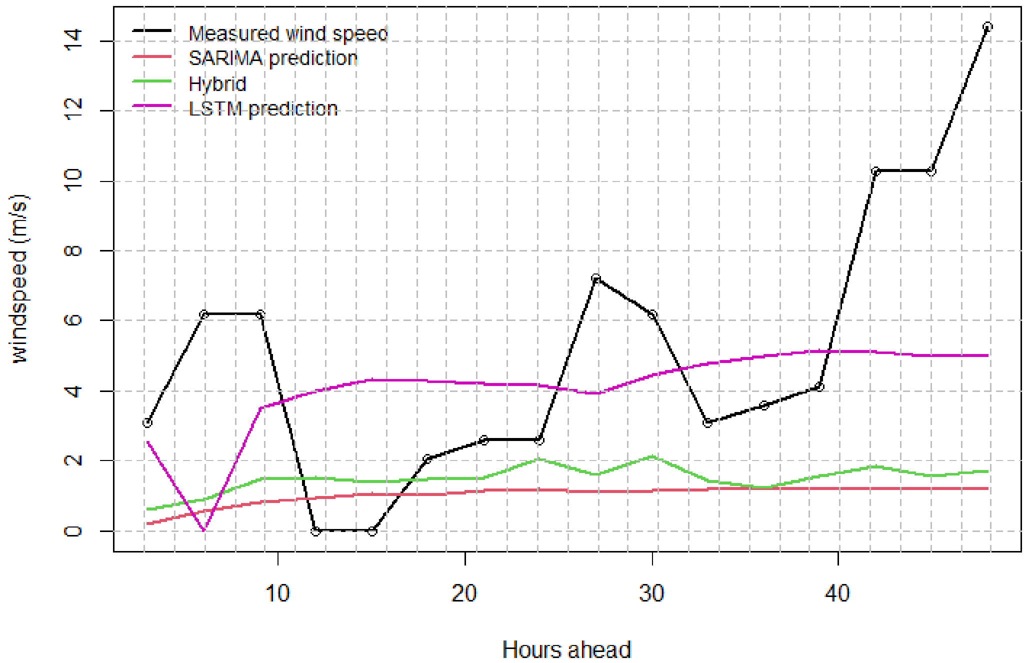

Figure 12. Comparison of SARIMA, LSTM and Hybrid with real data for 2-day prediction horizon during December; wind speed example; low wind speed during previous days.

Another case of 2-day wind speed forecasting is presented at Fig.13 where the wind speed that was recorded during the previous hours had a big magnitude of approximately 10 m/s. In this case, we find the best method to be the Hybrid by comparing their MAE, which are LSTM=2.81, SARIMA=2.43 and Hybrid=2.28. We also note that the SARIMA tends to predict around the mean value of the measured wind speed, whereas the Hybrid to predict correctly the added residual to the SARIMA. Judging by Figs 12-13, a different best method outcome can be made for each. This adds to the ambiguity on the choice of the better method, and points towards more research and deeper understanding of such methods, and the inherent chaotic nature that the wind speed time series exhibits.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 May 2022 doi:10.20944/preprints202205.0147.v1

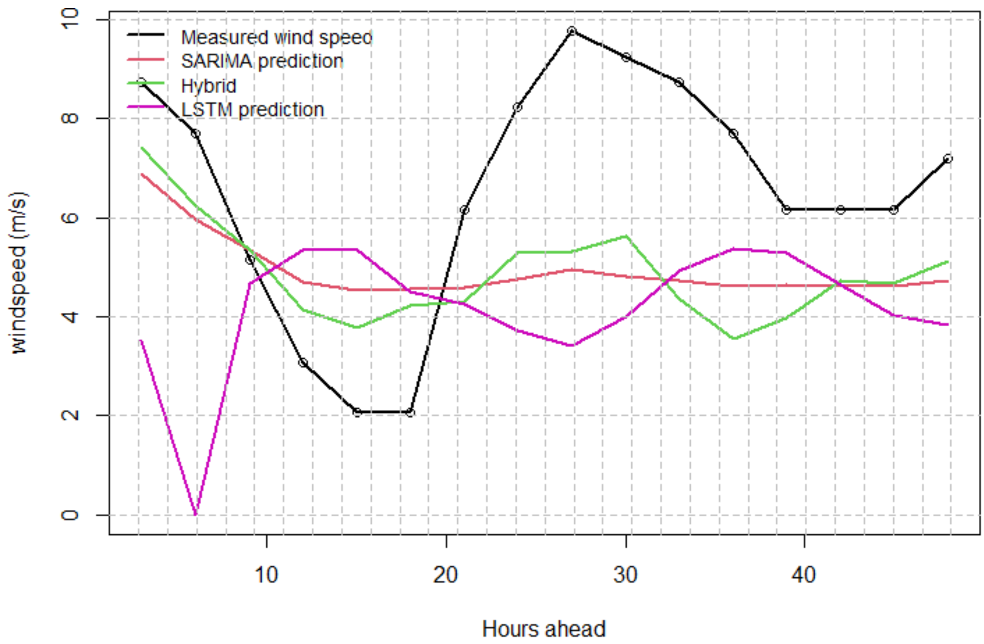

Figure 13. Comparison of SARIMA, LSTM and Hybrid with real data for 2-day prediction horizon during May; wind speed example; high wind speed during previous days.

4. Discussion – Future Work

In this work we introduced the Hybrid method and we showed its superiority over the SARIMA and that it is a valid predicting approach for the climatic variables considered here. The Hybrid was proven to be superior at the examples of temperature and wind speed prediction, whereas the LSTM compared well and had smaller errors on average at predicting the relative humidity.

The results indicate that in cases of extreme weather conditions and when the weather changes abruptly, the predictions are heavily unreliable. This is the biggest source of the errors that are emphasized by Figures 7, 8 and 13 and increase our averaged errors in table 1. With that in mind, these situations can be considered as an extreme case and not the norm for the dataset under consideration as Figures 2, 3, 9 and 11 indicate, where most of the days have low MAE and there are spikes of high error in some circumstances. Thus, the resulting errors in the table (that include both the stable and the few extreme weather change incidents) are at an acceptable scale. Of course, the reduction of error is always a goal, which is the reason why we introduced our Hybrid method that achieved up to 10% MAE reduction in some cases.

Our Hybrid technique can be further improved by combining another Back Propagation Network, so that instead of adding the SARIMA forecast and the Residual forecast, one can search for the functional relationship between these two quantities. Similar work for much less chaotic dataset is found in [24]. The LSTM and the SARIMA model performed comparably well and can both be utilized to predict the weather variables considered here adequately. Another Hybrid approach with promising results that could be implemented for our problem is found in [34], where the time series is firstly decomposed into a seasonal, a trend and a residual component, and then, different predictors are imposed to the different components. This could be combined with our Hybrid approach or the previously mentioned one of [24].

One could argue that all this discussion about Hybrid method implementation is insignificant since the LSTM model we presented here as a base model has better results most of the time than both the SARIMA and our proposed hybridization. We argue that the amount of work that it takes to optimize a LSTM model, fine tune the parameters, find the best working optimizers, activation functions, units for the layers etc. becomes a significant downside for practical applications. Parameter tuning is necessary for the research setting and methodological comparison, but it is not ideal for real world applications where multiple networks must be fitted for different locations, datasets and variables. Furthermore, the fact that it is not clear which days are predicted better with the LSTM and the SARIMA-Hybrid combination is a clear indication that more research need to be done towards developing more accurate methods.

Another way accuracy could be improved is by reframing the problem itself. To demonstrate, we will entertain an approach that will be implemented and tested in the future, and would eliminate the daily cyclic behaviour of temperature, which, as our tests have shown us, it is a quite difficult characteristic to predict on top of the yearly frequency. This method can be applied to variables with periodic characteristics and simplify the datasets beforehand. By applying daily max-min-mean functions to our data, instead of predicting 8 3-hour intervals for 1-day ahead, one could focus on predicting just one value using the daily averaged time series of maximum daily temperature, and analogously, train a different NN or SARIMA model on the minimum daily temperature time series. Having a prediction of the maximum and minimum temperature of the next day is of great importance. Furthermore, we can extend this to 3-hour interval predictions by invoking interpolating polynomials, since the general pattern of the temperature variation is known, within certain tolerances. This methodology can be applied also to humidity and any other periodic variable, utilizing the usual parabolic shape of most days. On the downside, it might impose additional errors due to the interpolation of the polynomial, but it is expected to have better accuracy on the predictor side since the dataset is less noisy, smaller and thus easier to handle computationally. Additionally, the amount of prediction time steps decreases significantly to 1 (or 2 for the 2-day forecast). In our future work, it remains to be seen if this method or the other hybrid methods mentioned here provide improvements in the accuracy of the predictions. Lastly, following [5] which focuses on solar radiation prediction for a high density grid of Meteorological stations, one could add different data sources and compare a model like the Random Forests that those authors implement, with the methods we presented here.

5. Conclusions

SARIMA and LSTM perform adequately well at predicting weather data for one to two days, by using localised time series. The SARIMA is a really good fit, especially for the one-day forecast and the Hybrid method having SARIMA as its primary predictor performed better. The improvement was not so pronounced, especially considering the added computational cost and complexity (2 hours of computing). One can argue that this is “offline” computational cost, meaning that the training of the methods happens once, and then the model can predict each day without too much computational effort, so the computational time is an insignificant concern. On the other hand, the LSTM with 2 hidden layers and a lot of fine tuning sometimes outperformed the other methods, especially in the 2-day forecasts in which it is clearly preferred. The future challenge is to further increase accuracy, and as proposed in the discussion section, use a combination of these techniques, apply new hybrid techniques or even reframe the problem in order to achieve it. Finally, the methods would preferably require minimal expert knowledge, testing time from scientific personnel and cost of computing capabilities so that these methods can be utilized in the field and have a wider range of applications in everyday life.

Funding: This work was financially supported by the project “Meteorological, Bioclimatic and oceanographic information display system for tourist applications (SmarTourism)” under the code MIS 5029194 under the RIS3Crete co-funded by Greece and European Union.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Additional data to our MS were provided by the Hellenic National Meteorological Service and can be downloaded freely from (http://www.emy.gr/emy/el/services/paroxi-ipiresion-elefthera-dedomena). Our MS data are available upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bochenek, B. and Z. Ustrnul, *Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives.* Atmosphere, 2022. **13**(2).
2. Georgopoulou, E., et al., *Climatic preferences for beach tourism: an empirical study on Greek islands.* Theoretical and Applied Climatology, 2019. **137**(1): p. 667-691.
3. Elmira, J. and O. Dilshan Remaz, *Intra Urban Air Temperature Distributions in Historic Urban Center.* American Journal of Environmental Sciences, 2012. **8**(5).
4. Scott, D. and C. Lemieux, *Weather and Climate Information for Tourism.* Procedia Environmental Sciences, 2010. **1**: p. 146-183.
5. Zeng, Z., et al., *Daily Global Solar Radiation in China Estimated From High-Density Meteorological Observations: A Random Forest Model Framework.* Earth and Space Science, 2020. **7**(2): p. e2019EA001058.
6. Shih, H. and S. Rajendran, *Comparison of Time Series Methods and Machine Learning Algorithms for Forecasting Taiwan Blood Services Foundation's Blood Supply.* Journal of Healthcare Engineering, 2019. **2019**: p. 6123745.
7. Sillmann, J., et al., *Understanding, modeling and predicting weather and climate extremes: Challenges and opportunities.* Weather and Climate Extremes, 2017. **18**: p. 65-74.
8. Wilks, D.S. and R.L. Wilby, *The weather generation game: a review of stochastic weather models.* Progress in Physical Geography: Earth and Environment, 1999. **23**(3): p. 329-357.
9. Kellengere Shankarnarayan, V. and H. Ramakrishna, *Comparative study of three stochastic future weather forecast approaches: a case study.* Data Science and Management, 2021. **3**: p. 3-12.
10. Box, G., *Box and Jenkins: Time Series Analysis, Forecasting and Control*, in *A Very British Affair: Six Britons and the Development of Time Series Analysis During the 20th Century*, T.C. Mills, Editor. 2013, Palgrave Macmillan UK: London. p. 161-215.
11. Hyndman, R.J. and G. Athanasopoulos, *Forecasting: principles and practice*. 2014: OTexts.
12. Grigonytė, E. and E. Butkevičiūtė, *Short-term wind speed forecasting using ARIMA model.* Energetika, 2016. **62**.
13. Lai, Y. and D.A. Dzombak, *Use of the Autoregressive Integrated Moving Average (ARIMA) Model to Forecast Near-Term Regional Temperature and Precipitation.* Weather and Forecasting, 2020. **35**(3): p. 959-976.
14. Paulescu, M., et al., *Time Series Forecasting*, in *Weather Modeling and Forecasting of PV Systems Operation*, M. Paulescu, et al., Editors. 2013, Springer London: London.
15. Kreuzer, D., M. Munz, and S. Schlüter, *Short-term temperature forecasts using a convolutional neural network — An application to different weather stations in Germany.* Machine Learning with Applications, 2020. **2**: p. 100007.
16. Kuhn, M. and K. Johnson, *Applied predictive modeling*. 1 ed. 2018, New York, NY: Springer. 600.
17. Chen, P., et al., *Time Series Forecasting of Temperatures using SARIMA: An Example from Nanjing.* IOP Conference Series: Materials Science and Engineering, 2018. **394**: p. 052024.
18. Ray, S., et al., *Time Series SARIMA Modelling and Forecasting of Monthly Rainfall and Temperature in the South Asian Countries.* Earth Systems and Environment, 2021. **5**(3): p. 531-546.
19. Dwivedi, D.K., J.H. Kelaiya, and G.R. Sharma, *Forecasting monthly rainfall using autoregressive integrated moving average model (ARIMA) and artificial neural network (ANN) model: A case study of Junagadh, Gujarat, India.* Journal of Applied and Natural Science, 2019. **11**(1): p. 35-41.
20. Zhang, G.P., *Time series forecasting using a hybrid ARIMA and neural network model.* Neurocomputing, 2003. **50**: p. 159-175.
21. Tealab, A., *Time series forecasting using artificial neural networks methodologies: A systematic review.* Future Computing and Informatics Journal, 2018. **3**(2): p. 334-340.
22. Hajirahimi, Z. and M. Khashei, *Hybrid structures in time series modeling and forecasting: A review.* Engineering Applications of Artificial Intelligence, 2019. **86**: p. 83-106.
23. Saba, T., A. Rehman, and J.S. AlGhamdi, *Weather forecasting based on hybrid neural model.* Applied Water Science, 2017. **7**(7): p. 3869-3874.
24. Deng, Y., H. Fan, and S. Wu, *A hybrid ARIMA-LSTM model optimized by BP in the forecast of outpatient visits.* Journal of Ambient Intelligence and Humanized Computing, 2020.
25. MIECZKOWSKI, Z., *THE TOURISM CLIMATIC INDEX: A METHOD OF EVALUATING WORLD CLIMATES FOR TOURISM.* 1985. **29**(3): p. 220-233.
26. Rutty, M., et al., *An Inter-Comparison of the Holiday Climate Index (HCI:Beach) and the Tourism Climate Index (TCI) to Explain Canadian Tourism Arrivals to the Caribbean.* 2020. **11**(4): p. 412.
27. Hyndman, R.J. and Y. Khandakar, *Automatic Time Series Forecasting: The forecast Package for R.* Journal of Statistical Software, 2008. **27**(3): p. 1 - 22.
28. Hyndman, R.J., et al., *forecast: Forecasting functions for time series and linear models*. 2021.
29. Petnehazi, G., *Recurrent Neural Networks for Time Series Forecasting*. 2018.
30. Lewis, N.D.C., *Neural networks for time series forecasting with R : an intuitive step by step blueprint for beginners*. 2017.
31. Elsaraiti, M. and A. Merabet, *A Comparative Analysis of the ARIMA and LSTM Predictive Models and Their Effectiveness for Predicting Wind Speed.* Energies, 2021. **14**(20).
32. Ruder, S., *An overview of gradient descent optimization algorithms.* CoRR, 2016. **abs/1609.04747**.
33. Kingma, D.P. and J. Ba, *Adam: A Method for Stochastic Optimization*. 2015.

 doi:10.20944/preprints202205.0147.v1

34. Dave, E., et al., *Forecasting Indonesia Exports using a Hybrid Model ARIMA-LSTM.* Procedia Computer Science, 2021. **179**: p. 480-487.