Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 201 (2022) 255–264

Procedia Computer Science

www.elsevier.com/locate/procedia

The 13th International Conference on Ambient Systems, Networks and Technologies (ANT)
March 22 - 25, 2022, Porto, Portugal

Year-ahead Ambient Temperature Forecasting in Pharmaceutical Transport Lanes Thermal Conditions

Annabel P.B. van den Berg[a,*], Lotte R. Bootsma[a], Thomas F.A Bovenber[a], A. Rosa Moerbeek[a], Eelco de Jong[b], Sergey Khalil[c], Thomas Koch[a,c], Dr. Elenna R. Dugundji[a,c]

[a] *Vrije Universiteit Amsterdam, Department of Mathematics, Faculty of Sciences, De Boelelaan 1111, 1081 HV Amsterdam, Netherlands*
[b] *Validaide BV, Johan Huizingalaan 400, 1066 JS Amsterdam, Netherlands*
[c] *Centrum Wiskunde & Informatica, Stochastics Research Group, Amsterdam Science Park 123, 1098 XG Amsterdam, Netherlands*

Abstract

This research aims to predict year-ahead ambient temperature aggregated at monthly level per hour on airport locations using historical temperature data. In this research, an Extreme gradient boosting regression model, LSTM model and the benchmark model, a persistence model, are compared in accuracy. Models are fine-tuned on the cities with the highest variability in temperature and grid searches are implemented only for these cities (one per data source). Overall, we have seen that the LSTM model with output size 12 months x 24 hours predicts for the next year the best. The Persistence model is closely followed by the Extreme gradient boosting model, with a small deviation in the quantiles. The point predictions for each of the other models are a bit further off (with more than 3 degrees Celsius) and LSTM 365 days x 24 hours is the worst in this case. These models can be used to give an indication for the ambient temperature on lane level.

Keywords: Ambient temperature; Forecasting; Pharmaceutical products; Packaging; Air cargo; Tarmac; LSTM; Persistence model; Extreme gradient boosting

1. Introduction

Pharmaceutical products are time and temperature sensitive products, which means that they require a strictly controlled transport process to keep the effectiveness of the product optimal. It often happens that a product must be distributed by various means of transportation each entailing potential fluctuations in ambient temperature. The temperature of these products can be maintained by multiple solutions throughout the logistical lane of the transport process, such as the use of thermal blankets, liquid nitrogen and temperature-controlled containers.

* Corresponding author.
E-mail address: annabelvdberg@hotmail.nl

10.1016/j.procs.2022.03.035

In light of the COVID-19 crisis, it can be found that the required temperature of vaccines is very important. The first authorised vaccine was the Pfizer-BioNTech, and was, after the Ebola vaccine, the second human vaccine which required distribution at temperatures between -80°C and -60°C [8]. Such extreme conditions are not standard; nevertheless: they certainly need to be considered. Other pharmaceutical products with less extreme conditions are easier to transport, but as James (2021) observes, there are no licensed vaccines yet that meet the condition of being stored and transported at ambient temperatures [8]. This renders the storage and distribution chain of pharmaceuticals very vulnerable, complicated and costly [22]. The deeper one examines the pharmaceutical supply chain, the fewer checks are performed and the less money is invested in controlled storage, handling and distribution. Yet, tracking this chain is not only important to preserve medicines as well as possible in order to maintain their efficiency, but also to limit costs and energy.

By predicting the ambient temperature throughout the lane, packaging and routing can be adapted to these outcomes. Nevertheless, while most studies have focused on the importance of temperature on medicines; little attention has been devoted to solutions to prevent or limit this impact when temperatures differ throughout a distribution lane.

The aim of this paper is to evaluate and predict the ambient temperature at the tarmac of 18 airports in 11 different transport lanes for every month. This forecast provides insight into the most fragile point in the lane assessment, doing this provides enough information to develop an advice on the heaviness of the packaging and most cost-efficient solution.

1.1. Problem description

In this research, we are aiming to use advanced data analytics and AI for the monthly analysis of thermal conditions on several global transport lanes. We are specifically focused on the tarmac period, which is the most fragile period in the lane. The tarmac period is the time frame in the transport process when freight is waiting on the airfield to be loaded onto a plane or to be transported further to a warehouse. At the tarmac there are no standard facilities to protect the freight from the ambient temperature as there would be in most warehouses or planes. Therefore, the ambient temperature can influence the quality of pharmaceutical freight during the tarmac period and the usability of the pharmaceutical freight after transportation. The only way to protect the freight against the influence of the temperature is using the right packaging.

2. Literature Review

This research consists of a forecasting problem for time-series. Therefore, a literature research was conducted to examine various papers related to time series and models to forecast the ambient temperature. First of all, several studies have examined the prediction of ambient temperature. [7] and [10] have shown in various machine-learning projects that extreme gradient boosting can be an excellent way to predict the temperature. It has often been an effective prediction for both classification and regression tasks, since the bias-variance trade-off can easily be controlled [23]. Decision trees are the base learner of the gradient boosting model. Decision trees are regression methods that use the partitioning of the input parameters into a distinct non-overlapping region with the use of if-then rules [24]. Advantages of the decision trees is that they can handle categorical and numerical data simultaneously. The variable selection is automatically performed due to the hierarchical structure of the decision tree [24]. However, this method has several limitations. The decision trees can be prone to overfitting, since they can feature high depth and may not find the generalisation of the relationship between the response and explanatory variables. Nevertheless, the gradient boosting machine can be used to increase the performance of decision trees [5].

In addition, besides the use of traditional methods, such as (S)ARIMA, other methods such as ANN and RNN have shown promising forecasting results [16]. Although RNNs are often used as reliable methods to perform time series predictions, it does not have the capability to predict long-term dependencies because the gradients in these models tend to vanish or explode over these longer periods [9]. To overcome this problem, the Long Short-Term Memory (LSTM) was presented by [6] for the first time in 1997 and is defined as a special kind of RNN. [19] states that LSTM is designed to model temporal sequences and has long-range dependencies that make the model more accurate than conventional RNNs.

To compare all of the models, research has shown that RMSE and MAE are extensively used measures in other temperature and time series related articles. These measures are therefore practical to use in this research, so they can be compared to other studies. However, [25] stated that the RMSE is an inappropriate and misinterpreted measure of the average error. Since the RSME tends to become increasingly larger than the MAE, but not in monotonic fashion. They claim that this is of particular interest because it is one of the most widely reported and misinterpreted error measures in the climatic and environmental literature. Nonetheless, the RMSE is still used for comparison reasons.

In addition, types of packaging for temperature controlled products were investigated, since there are some negative impacts that are commonly noticed when the temperature is out of range, such as loss of assay, increase of impurity, separation of layers of liquid products, change in dissolution pattern of solid dosage and discolouration of products [14]. Hence, it is of great importance that the package ensures that for every environmental condition on the transport lane the pharmaceuticals stay within their temperature range.

The type of packaging used is one of the most important variables that affects the temperature of temperature controlled products [20]. In this research 3 sub-classes of packaging are looked into. The first sub-class is active packaging. Active packaging is a mechanical or electric system which needs to be powered by an energy source at all times.These systems are thermostatic controlled to manage the products' temperature. Active systems can offer high thermal stability and can provide the exact temperature required. They also offer great security, since the units are locked and do not have to be opened during transportation.

However, active packages come at a high capital cost, and are almost exclusively leased. [18] stated that the costs of these systems can become prohibitive with the irregular lanes, variable load sizes and long distances as these systems also need to be collected and returned after use. The second subclass is passive packaging. Passive thermal systems commonly use phase change materials or insulation material to manage the temperature and these systems cannot be thermostatic controlled. The efficiency of most passive packaging systems have a maximum of 48 hours according to [1]. Passive packaging systems can be held in stock to provide volume flexibility [18] and the flexibility of passive packaging is strengthened due to the flexibility in volume size of the shipment and because of the single-use of these systems: there is no specialised infrastructure needed and no need to return the systems after use. This flexibility causes less logistic management and therefore passive packaging offers more route options and has considerably lower costs in comparison with active packaging systems.

Nonetheless, because these passive packaging systems are often not reusable, or only reusable a few times, it causes a lot of waste which is also not biodegradable in some cases. Passive packaging systems are also not thermostatic controlled and thus can fluctuate in temperature, which can cause the freight to evade their temperature range.

The last sub-class that is looked into is hybrid packaging. Hybrid packaging systems use a combination of phase change materials and thermostatic control to maintain the temperature of the freight. The hybrid packaging systems have the same benefits of active packaging systems and have some of the flexibility of passive packaging systems where they do not always need to be powered by an energy source. However, these systems do require the logistic management which is also required for active packaging.

3. Data Analysis

Different kinds of data from different sources are used. Besides the data provided by the host company, data is obtained from two additional sources for meteorological data. These sources contain information about the weather of certain locations. The company's data consisted of detailed information of 11 transportation lanes which are distributed on 16 airport locations.

Both data sources contain weather information with several variables. These variables differ between the two data sources. For each of the data sources, the measurements run from January 1, 1991 to December 31, 2020. Both data sets are checked for missing data, and duplicate rows are deleted. Columns with variables that have no correlation with temperature are deleted, as well as the columns with values that are causal. The remaining columns are renamed.

[21] suggest that when predicting temperature for several climates, a suitable approach is to train your models on the city with the highest variability in temperature and then fine-tuning the model for the other cities.

For each of the data sources, the city with the highest variability in temperature is determined to use this city for hyper parameter selection and fine-tuning of the models. By looking at the 5% quantile and 95% quantile, we can establish a 90% confidence interval of the temperatures. The city where this range is the highest, is used for hyper parameter selection and fine-tuning of the models. For the first data source, Zurich is the city with the highest variability and for the other data source, Chicago has the highest range.

3.1. Seasonality

To check for seasonality in the time series data, an Error-Trend-Seasonality decomposition (ETS) decomposition is used. An ETS decomposition for the last two years of data is used and this is compared to the remaining 30 years to check the trend. This is due to time limits. From the ETS decomposition, it is found that there is a double seasonality, which is both yearly and daily (24 hours).

4. Methodology

In this section, the persistence model, extreme gradient boosting (XGBoost) model and Long Short-Term Memory (LSTM) model are described. These models are used to forecast the ambient temperature at the given locations. Also, the forecasting errors Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) that are examined in this report are explained.

4.1. Persistence model

The persistence model is chosen as the benchmark model. This will be used as a lower bound on the other models discussed in this research paper. This model is a supervised machine learning model and is obtained by:

$$y(n+1) = y(n) \tag{1}$$

[12] indicates that the prediction at time n is then equal to the measurement at time $n-1$. So, it uses the value at the previous step to predict the expected outcome at the next time step. In this research, this means that the temperature on a given date and time is equal to the temperature of that exact date and time in the previous year.

4.2. Extreme gradient boosting

XGBoost is an advanced gradient boosting framework which was introduced by [3]. In comparison with gradient boosting, XGBoost adds an additional regularising term to smooth the final weights. This additional term prevents overfitting [3]. Furthermore, [29] indicates that XGBoost also uses row and column sampling to solve overfitting. The formula can be expressed as:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i) \tag{2}$$

Where f_k depicts an independent tree structure with leaf scores [15]. The regularised objective which needs to be optimised is given by

$$Obj(\Theta) = \sum_{i}^{n} l(y_i, \hat{y}_i) + \sum_{k}^{K} \Omega(f_k) \tag{3}$$

Where l is the differentiable loss function and Ω the regularisation term, which can be written as:

$$\Omega(f) = \gamma \frac{1}{2} \lambda \sum_{j-1}^{T} \omega_j^2 \tag{4}$$

Where T equals the number of leaves and ω describes the score on each leaf. γ and λ represent constants which can control the degree of regularisation [15]. XGBoost has multiple hyper-parameters to be determined. These are decided by using grid-search.

4.3. Long Short-Term Memory

A LSTM is an extension of a RNN model. The difference between a RNN model and the LSTM model is the structure of the hidden layers, which is more complex in the latter. Moreover, LSTM is a solution for both short and long term temporal dependencies in time-series data due to its handling of vanishing gradient and exploding gradient because of the structure of LSTM which includes hidden layers [27]. Also, the memory cells within the hidden layers of the LSTM model ensures that important features are ”remembered” and unimportant features are ”forgotten”, this helps the model to learn long-term dependencies and gives superior outcomes [17].

In this research, two LSTM models are tested:

- Single-Layer LSTM: 1 LSTM hidden layer; Output size: 365 days x 24 hours
- Single-Layer LSTM: 1 LSTM hidden layer; Output size: 12 months x 24 hours

The technical LSTM architecture descriptions are described in Table 1. In the first approach with output size 365 days x 24 hours, the hourly temperature measurements are predicted per day for the next year and used to calculate the hourly mean per month.In the second approach with output size 12 months x 24 hours, the hourly mean per month is predicted directly. The latter approach reduces the amount of output data.

Table 1: Technical LSTM Architecture Descriptions

Model	**Layer (type)**	**Output Shape**	**Nr. of Parameters**
Single-Layer LSTM	LSTM	(None, 11)	572
Output 365x24	Dropout	(None, 11)	0
	Dense	(None, 8760)	105120
Single-Layer LSTM	LSTM	(None, 31)	4092
Output 12x24	Dropout	(None, 31)	0
	Dense	(None, 288)	9216

4.4. Forecasting errors

For forecasting the temperature, different models are built. For each model, different loss functions are calculated to be able to compare each model. The loss functions that will be considered are: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE).

The RMSE is one of the most common measures to calculate the difference of a value predicted by a model and the observed values, also known as the error. The RMSE is the square root of the variance of the residuals and describes the root of the second sample moment and it is calculated by [2]: $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}$

where n is the number of observations, $\hat{y}_i$ are the predicted values and y_i the observations. Whenever the RMSE equals 0, it means that the prediction is a perfect fit to the observed value. The RMSE is proportional to the error. This means that large errors are given a relatively high weight. However, this loss function can only be used to compare the error between different models as it is scale-dependent.

The MAE is another commonly used loss function. This loss function takes average of the absolute difference of all the predicted values and observed values, thus: $MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|$

where n is the number of observations, $\hat{y}_i$ represents the predicted values and y_i the observations.

The MAE is a more robust loss function than the RMSE, which causes the function to be less sensitive for outliers. Thus, when there is a higher chance of outliers being corrupted, the MAE may be a more appropriate function to use. As the RMSE, this is a scale-dependent function and thus only can be used to compare different models. As the RMSE, this is a scale-dependent function and thus only can be used to compare different models.

5. Results

In this section, the results of Extreme gradient boosting regression and LSTM are shown and compared to each other and to the benchmark model. Table 2 displays the quantile scores, and the overall scores of the models. The Persistence model is divided into two parts, since the results are computed for the two data sources separately.

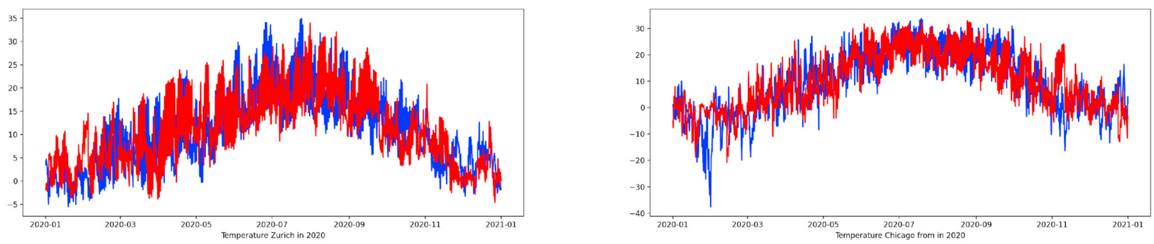
Table 2: Quantiles and Overall Scores for all Models with Output 365 days x 24 hours

Model	**RMSE 5%**	**MAE 5%**	**RMSE 95%**	**MAE 95%**	**RMSE**	**MAE**
Persistence 1	1.554	1.198	1.867	1.398	3.865	2.844
Persistence 2	2.471	1.657	2.594	1.967	4.826	3.579
XGBoost	2.486	1.842	2.558	2.045	3.979	2.964
LSTM (365x24)	6.997	5.934	6.114	4.563	6.617	5.081

5.1. Benchmark model: Persistence model

The persistence model is used to predict the ambient temperature. For both fine-tuning cities, the predictions (blue) are plotted against the measured values (red) in 2020 as can be seen in Figure 1. For modelling scores, the RMSE and MAE are used. The overall scores of the model can be found in Table 2.

For the predictions on the lanes, the temperature is given per month per location. For an indication of the expected climate, a 90% confidence interval for the temperature in a certain month is used. The overall scores for the monthly quantiles can be found in Table 2.



(a) Zurich (b) Chicago

Fig. 1: Plots RMSE Persistence Chicago and Zurich

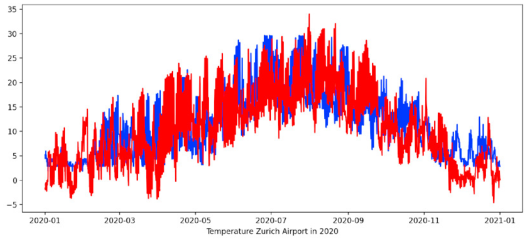
5.2. Extreme gradient boosting regression model

A grid-search using the tool Optuna was implemented to determine the hyper-parameters for the gradient boosting model. The final hyper parameters can be found in Table 4.

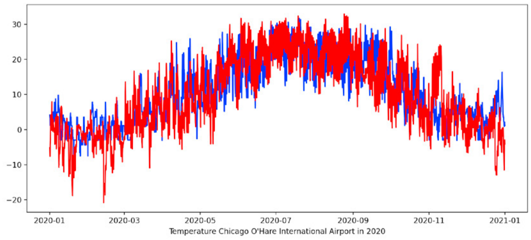
Using these parameters, the model results in the overall model scores and in the overall quantile scores as can be found in Table 2. In Figure 2, the predictions (blue) of the extreme gradient boosting regression model are plotted against the measured values (red) in 2020.

5.3. Long Short-Term Memory

For the LSTM hyper parameters, a grid-search is used to evaluate all possible model combinations for each of the data sources. This resulted in the hyper parameters as depicted in Table 3. These parameters are used for training and



(a) Zurich



(b) Chicago

Fig. 2: RMSE XGBoost Chicago and Zurich

testing on all other airport locations. After creating the predictions using the LSTM model, the RMSE and MAE are calculated for the point predictions and the quantiles for all locations. The modelling scores can be found in Table 2.

Data pruning was necessary for the single-layer LSTM model to increase the search space. The parameters which resulted in the best model can be found in Table 3. What was noticeable is that models with a high number of neurons were failing to complete.

Table 3: Hyper-parameters for Grid-search LSTM

Variable	**Zurich (365x24)**	**Chicago (365x24)**	**Chicago (12x24)**
Neurons	20	50	31
Activation	Softplus	Softplus	Softplus
Kernel initializer	Uniform	Glorot norm.	Glorot uniform
Dropout rate	0.9	0.9	0.3
Activation dense	Softsign	Linear	Linear
Optimizer	Nadam	Nadam	Nadam
Epochs	400	200	50
Batch size	3	1	1
MSE score	0.02665	0.03593	0.004195

Table 4: Hyper-parameters for Extreme Gradient Boosting

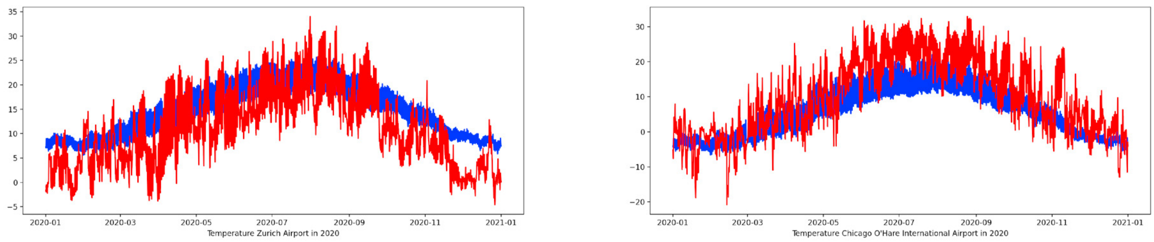
Variable	**Value**
booster	dart
lambda	0.0005
alpha	0.0041
subsample	0.4242
colsample bytres	0.7610
max depth	3
min child weight	4
eta	0.2882
gamma	0.0347
grow policy	depth wise
sample type	uniform
normalize type	forest
rate drop	2.2689e-06
skip drop	0.00060

In Figure 3, the predictions (blue) of the LSTM model with output size 365 days x 24 hours are plotted against the measured values (red) in 2020.

6. Discussion

Overall, only the extreme gradient boosting regression model has shown to be competitive with the Persistence model and can provide reliable ambient temperature forecasts. The LSTM model with output size 12x24 significantly outperformed all models. Nevertheless, some aspects of the research were unknown, or have had an influence on the results which decreased the accuracy.

First of all, due to limited time, limited computational power or memory, a high number of hyper-parameter combinations and large data, concessions had to be made with respect to the research. For both the grid searches of LSTM and SARIMA we have not been able to investigate all targeted combinations of hyper-parameters and have only considered the combinations that we were able to investigate for the results of our research. Thus, we have only considered



(a) Zurich

(b) Chicago

Fig. 3: RMSE LSTM (365 x 24) Chicago and Zurich

a small part of all possible combinations for the grid searches. Maybe the optimal solution happens to be among those combinations, otherwise we have only used the local solution from the part of the grid that we investigated.

In line with the aforementioned problems, we have also used the approach of [21] to fine-tune models on the city with the highest variability. This way, the choice of hyper-parameters depends on the climate of this city. Though [21] proved that this is a suitable approach, it might influence performance of the model on the other climates. This was experienced in the LSTM model, where the best model for the city with the highest variability did not yield results for all cities.

In this research, we have not considered the use of exogenous variables in the prediction of the ambient temperature, which would be possible in the case of for instance the LSTM model. Some of these variables can be correlated to the ambient temperature and help yield better predictions. Also, by including a prediction of the solar radiation on the tarmac in the dashboard a more complete indication can be given of expected tarmac circumstances.

Besides this, the outside temperature in the provided data was measured in the shade. This temperature does not directly say something about the effect that freight might experience by the ambient temperature in the sun. Our predictions are also based on this shade measured temperature and therefore do not give a full indication of circumstances on the tarmac. Hence, one can conclude that freight might be exposed to even higher temperatures than predicted.

Point predictions have been used to estimate the 90% confidence interval per hour. Thus, the temperature is predicted per hour for the next year and based on these predictions, the intervals and mean for each month is calculated. Another approach is to not predict the hourly measurements but to predict the mean and 90% confidence interval. In this research, we have not been able to investigate this, but this is an approach that will reduce the amount of data.

Moreover, it would be even more interesting to predict the temperatures of the freight itself. Methods such as the mean kinetic temperature (MKT) and Newton's Law of Cooling give insight into the influence of temperature fluctuations on the freight [13]. This was not done in this research due to a lack of time and missing information such as, the weight, the specific content of the package and the air within the package. Since pharmaceutical products differ in products and differ in primary and secondary packaging, the MKT will behave differently for each product. When taking these variables into account, a more accurate decision on packaging could be made.

Furthermore, detailed information on the performance of the packaging was lacking. Despite the information provided by a freight forwarder, not much was known. Therefore some assumptions were made based on the information on hand and numbers were suggested based on proportion. These assumptions are indications and could be improved with more detailed information. Also, with more exact information on the performance of the packaging at certain temperatures, the MKT would also increase in accuracy. Additionally, the packaging would need to be specified further as the thermal stability of packaging differs extremely between different brands and models.

7. Conclusion and Recommendations

Overall, we have seen that the LSTM model with output size 12 months x 24 hours predicts for the next year the best. The Persistence model is closely followed by the Extreme gradient boosting model, with a small deviation in the quantiles. The point predictions for each of the other models are a bit further off (with more than 3 degrees Celsius) and LSTM 365 days x 24 hours is the worst in this case.

Although this article ambitions completeness, there are some parts that could be investigated even further or could be improved. First, the possibilities of changing transport routes could be considered. This article focuses on different kinds of packaging to prevent freight from exceeding the critical temperature boundaries for fixed lanes and not on the best routing between locations. We have chosen not to emphasise on this because we have found that in practice there are often few possibilities to adjust routes and times. This was affirmed by a major freight forwarder. However, we recommend having a closer look at this before routing can be assumed to be optimal.

Furthermore, the packaging used for pharmaceutical products is chosen before a possible route is found. It is our recommendation to also look at the possibility of choosing these two variables at the same time. However, this might be hard as packaging is chosen by the shipper, while the route is chosen by the forwarder.

It is also recommended to look into the details of the performance of packaging more. Since there is still a lot to learn about performance of packaging and possibilities and difference between products and even different versions of the same product. In-depth research on packaging, such as costs and influence of ambient factors on the packaging could create more accurate advice on what kind of packaging can be used depending on the situation.

On the modelling side, it is recommended to investigate more models. Performance of the LSTM model might be improved by investigating all of the other hyper-parameter combinations that have not been included in this research. Stacked LSTM consists of multiple layers of LSTM combined to create a deep network, as mentioned in the articles of [28] and [11]. Bi-directional LSTM makes it possible to retain information from the past and the future by running the inputs forward and backward. By combining Bi-directional LSTM with stacked LSTM, all layers have this property. Also, the predictions are now based on models that have been fine-tuned on the cities with the highest variability. By fine-tuning on each city separately, better results might be achieved.

Also in our literature review, we have found Generalised Regression Neural Networks and TBATS models to be successful in prediction of weather-related variables like the ambient temperature. Also, a hybrid LSTM-SARIMA model could be considered for modelling. According to [26] the hybrid LSTM-SARIMA model seems to outperform the standalone models when predicting daily values and has gained popularity lately. Another example is in the research of [4], where exports are predicted, based on monthly historical values. However, no weather predictions are made using the hybrid model yet, thus further research is required.

By including more weather variables in prediction of the ambient temperature, the prediction might be strengthened but also give a more complete overview of the expected climate in certain locations. Even though solar radiation is a bit different from the temperature in the sense that there is little to no solar radiation in the night, most of the suggested models can also be used for the prediction of solar radiation. Therefore, it is recommended to consider the effect of exogenous variables and include other weather variable predictions for a more concise indication of an airport location's climate.

8. Acknowledgements

We would like to thank the Faculty of Science at the Vrije Universiteit Amsterdam for giving us the opportunity to write this paper as part of the Bachelor Business Analytics. In addition, we would also like to thank Wout Olde Hampsink and Robert Knoppert for their insights into the data and their cooperation during this research. Hereby, we would also like to acknowledge our team member Mitch Lobbes for his contribution to the research. Furthermore, we would like to thank Elisabeth Fokker for her insight into the models and her support.

References

[1] Abscisse, 2011. Project on air cargo related to pharmaceutical products in canada. Transport Canada .

[2] Chai, T., Draxler, R., 2014. Root mean square error (rmse) or mean absolute error (mae)? – arguments against avoiding rmse in the literature. Geoscientific Model Development 7, 1247–1250. URL: https://gmd.copernicus.org/articles/7/1247/2014/, doi:10.5194/gmd-7-1247-2014.

[3] Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system, Association for Computing Machinery, New York, NY, USA. p. 785–794. URL: https://doi.org/10.1145/2939672.2939785, doi:10.1145/2939672.2939785.

[4] Dave, E., Leonardo, A., Jeanice, M., Hanafiah, N., 2021. Forecasting indonesia exports using a hybrid model arima-lstm. Procedia Computer Science 179, 480–487. URL: https://www.sciencedirect.com/science/article/pii/S1877050921000363, doi:https://doi.org/10.1016/j.procs.2021.01.031. 5th International Conference on Computer Science and Computational Intelligence 2020.

[5] Friedman, J., 2001. Greedy function approximation: A gradient boosting machine. The Annals of Statistics 29, 1189–1232. URL: http://www.jstor.org/stable/2699986.
[6] Hochreiter, S., Schmidhuber, J., 1997. Long Short-Term Memory. Neural Computation 9, 1735–1780. URL: https://doi.org/10.1162/neco.1997.9.8.1735, doi:10.1162/neco.1997.9.8.1735.
[7] Hutchinson, R., Liu, L., Dietterich, T., 2011. Incorporating boosted regression trees into ecological latent variable models.
[8] James, E., 2021. Disrupting vaccine logistics. International Health 13, 211–214. URL: https://doi.org/10.1093/inthealth/ihab010, doi:10.1093/inthealth/ihab010.
[9] Jaseena, K., Kovoor, B., 2020. Deterministic weather forecasting models based on intelligent predictors: A survey. Journal of King Saud University - Computer and Information Sciences URL: https://www.sciencedirect.com/science/article/pii/S1319157820304729, doi:https://doi.org/10.1016/j.jksuci.2020.09.009.
[10] Johnson, R., Zhang, T., 2014. Learning nonlinear functions using regularized greedy forest 36, 942–54. doi:10.1109/TPAMI.2013.159.
[11] Karevan, Z., Suykens, J.A., 2018. Spatio-temporal stacked lstm for temperature prediction in weather forecasting. arXiv preprint arXiv:1811.06341 .
[12] Kavasseri, R., Seetharaman, K., 2009. Day-ahead wind speed forecasting using f-arima models. Renewable Energy 34, 1388–1393. doi:https://doi.org/10.1016/j.renene.2008.09.006.
[13] Konovalenko, I., Ludwig, A., Leopold, H., 2021. Real-time temperature prediction in a cold supply chain based on newton's law of cooling. Decision Support Systems 141, 113451. URL: https://www.sciencedirect.com/science/article/pii/S0167923620302062, doi:https://doi.org/10.1016/j.dss.2020.113451.
[14] Kumar, N., Jha, A., 2017. Temperature excursion management: A novel approach of quality system in pharmaceutical industry. Saudi Pharmaceutical Journal 25, 176–183. URL: https://www.sciencedirect.com/science/article/pii/S131901641630069X, doi:https://doi.org/10.1016/j.jsps.2016.07.001.
[15] Mustapha, I., Saeed, F., 2016. Bioactive molecule prediction using extreme gradient boosting. Molecules 21. URL: https://www.mdpi.com/1420-3049/21/8/983, doi:10.3390/molecules21080983.
[16] Muzaffar, S., Afshari, A., 2019. Short-term load forecasts using lstm networks. Energy Procedia 158, 2922–2927. URL: https://www.sciencedirect.com/science/article/pii/S1876610219310008, doi:https://doi.org/10.1016/j.egypro.2019.01.952. innovative Solutions for Energy Transitions.
[17] Qing, X., Niu, Y., 2018. Hourly day-ahead solar irradiance prediction using weather forecasts by lstm. Energy 148, 461–468.
[18] Roskoss, A., 2011. Temperature-controlled packaging systems—active or passive? URL: http://www.iptonline.com/articles/public/Intelsius.pdf.
[19] Salman, A.G., Heryadi, Y., Abdurahman, E., Suparta, W., 2018. Single layer & multi-layer long short-term memory (lstm) model with intermediate variables for weather forecasting. Procedia Computer Science 135, 89–98. URL: https://www.sciencedirect.com/science/article/pii/S187705091831439X, doi:https://doi.org/10.1016/j.procs.2018.08.153. the 3rd International Conference on Computer Science and Computational Intelligence (ICCSCI 2018) : Empowering Smart Technology in Digital Era for a Better Life.
[20] Sharley, P., Williams, I., Hague, S., 2003. Blood transportation for medical retrieval services. Air Medical Journal 22, 24–27. URL: https://www.sciencedirect.com/science/article/pii/S1067991X03001251, doi:https://doi.org/10.1016/j.amj.2003.08.006.
[21] Srivastava, S., Lessmann, S., 2018. A comparative study of lstm neural networks in forecasting day-ahead global horizontal irradiance with satellite data. Solar Energy 162. doi:10.1016/j.solener.2018.01.005.
[22] Sykes, C., 2018. Time- and temperature-controlled transport: Supply chain challenges and solutions. P & T : a peer-reviewed journal for formulary management 43, 154–170.
[23] Taieb, S., Hyndman, R., 2014. A gradient boosting approach to the kaggle load forecasting competition. International Journal of Forecasting 30, 382–394. URL: https://www.sciencedirect.com/science/article/pii/S0169207013000812, doi:https://doi.org/10.1016/j.ijforecast.2013.07.005.
[24] Touzani, S., Granderson, J., Fernandes, S., 2017. Gradient boosting machine for modeling the energy consumption of commercial buildings. Energy and Buildings 158. doi:10.1016/j.enbuild.2017.11.039.
[25] Willmott, C., Matsuura, K., 2005. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. Climate research 30, 79–82.
[26] Wu, D., Ji, L., He, K., Tso, K., 2021. Forecasting tourist daily arrivals with a hybrid sarima–lstm approach. Journal of Hospitality & Tourism Research 45, 52–67. doi:10.1177/1096348020934046.
[27] Yu, Y., Cao, J., Zhu, J., 2019. An lstm short-term solar irradiance forecasting under complicated weather conditions. IEEE Access 7, 145651–145666. doi:10.1109/ACCESS.2019.2946057.
[28] Zhang, F., Feng, N., Liu, Y., Yang, C., Zhai, J., Zhang, S., He, B., Lin, J., Du, X., 2020. Pewlstm: Periodic lstm with weather-aware gating mechanism for parking behavior prediction. IJCAI , 4424–4430.
[29] Zhang, W., Wu, C., Zhong, H., Li, Y., Wang, L., 2021. Prediction of undrained shear strength using extreme gradient boosting and random forest based on bayesian optimization. Geoscience Frontiers 12, 469–477. URL: https://www.sciencedirect.com/science/article/pii/S1674987120300669, doi:https://doi.org/10.1016/j.gsf.2020.03.007.