FROST FORECASTING MODEL USING GRAPH NEURAL NETWORKS WITH SPATIO-TEMPORAL ATTENTION

Hernan Lira, Luis Martí & Nayat Sanchez-Pi
Inria Chile Research Center
Av. Apoquindo 2827, Las Condes, Santiago de Chile, Chile.
{hernan.lira, luis.marti, nayat.sanchez-pi}@inria.cl

ABSTRACT

Frost forecast is an important issue in climate research because of its economic impact in several industries. In this study, a graph neural network (GNN) with spatio-temporal architecture is proposed to predict minimum temperatures in an experimental site. The model considers spatial and temporal relations and processes multiple time series simultaneously. Performing predictions of 6, 12, and 24 hrs this model outperforms statistical and non-graph deep learning models.

1 INTRODUCTION

Weather prediction is an important issue across several industries with plenty of studies aiming to face different environmental dangers. Frost events are one of these dangers, which causes product and economic losses in agricultural-related activities. During some periods of the year, the temperature can vary several degrees during the day, and at nightfall, it can become suddenly negative. This variation of the temperatures can cause the flowering to come early, or the flower can freeze, losing its value for production. (M. Á. Guillén-Navarro & Martínez-España, 2017).

Frost events are challenging to predict since they are a localized phenomenon, influenced by environmental factors, including air temperature, humidity, and others. Most of the studies and methods of frost forecasting are related to numerical weather prediction by simulating partial differential equations or traditional statistical models. This approach requires a recurrent theoretical upgrade to incorporate weather and atmospheric assumptions. On the contrary, machine learning algorithms do not make any assumptions about weather behavior. Instead, they use historical weather data as an input and train a model to predict future weather values (Diedrichs et al., 2018).

Since frost can be highly variable across a small area, the collection of temperature data usually is not available with sufficient resolution. In addition, the number of frost events during the year is small, then the available data is scarce to build an accurate prediction model. Moreover, other variables such as humidity, radiation, and wind have their contribution as well. From a practitioner's view, frost severity and action time are also key factors to consider in a prediction (Ding et al., 2020).

To build a model that considers the constraints mentioned above, especially data scarcity and small datasets, different deep learning models have been proposed. In particular, graph neural networks (GNN) and attention mechanisms are suitable for this problem since they incorporate spatial knowledge to the model field and environmental interactions (Cheng et al., 2020). In addition, the occurrence of frost is caused by a prior movement of environmental factors. GNNs can be naturally extended to model this type of temporal interaction. Therefore, this study is presented as a time series forecasting problem using GNNs and attention.

In this study, air temperature and humidity data are collected from an experimental site. In addition, we use data from 10 weather stations. We propose a GNN with spatio-temporal attention architecture for frost forecast. We map weather stations' locations to nodes on a graph and construct the edges based on geographical proximity. Furthermore, the adjacency matrix is optimized during the training phase, therefore other interactions can be learned. We utilize spatio-temporal attention to incorporate similar locations and time.

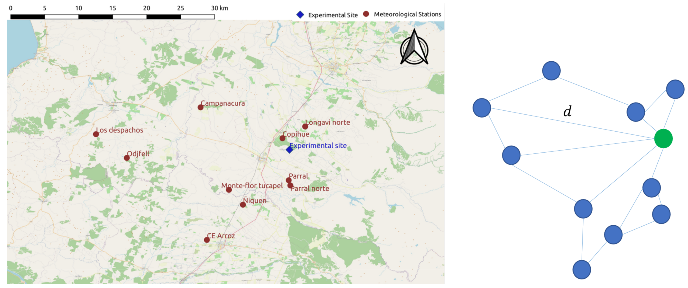

Figure 1: Map of the meteorological stations and experimental site.

2 Related Work

Initial efforts in facing the frost phenomenon are shown in (Mort & Chia, 1991; Verdes et al., 2000), studies in which authors present a temperature prediction system for agriculture using artificial neural networks. The basic goal is to predict the minimum temperature of the next day using historical data.

In recent years, IoT systems have come to help with the prediction of weather (Y Muck & J Homam, 2018). Levin Varghese (2019) present a weather forecast system based on a Raspberry Pi 3 Model B+, sensors, and a weather forecast algorithm. The presented system monitors air temperature, humidity, pressure, and altitude, of locations. The weather forecast algorithm is based on a Linear Regression Model. Diedrichs et al. (2018) uses IoT devices to extract data from selected locations and applied machine learning techniques. In addition, Ding et al. (2020) use weather stations and locally collected weather data to create causal-effect machine learning models of frost prediction. In summary, they show that future improvements are related to minimizing false positive predictions.

Regarding the models of weather forecasting, recent studies have focused on the adaptation and application of deep learning architectures. In (Shi et al., 2015) authors proposed a LSTM architecture combined with convolutions for precipitation prediction. The study (Mehrkanoon, 2019) introduced models based on 1D-CNNs to process tensor 3D data and extract spatio-temporal relations. In these approaches there are no spatial relations between the entities (e.g. different cities), the interaction is determined by the entity order in the tensor.

Graph neural networks have shown recent progress regarding time series forecasting and spatio-temporal relations. Zhang et al. (2019) include spatial relations between weather stations by creating a graph. Different studies (Wilson et al., 2018; Khodayar & Wang, 2018) have created applications of spatio-temporal graphs for predicting weather variables. In addition, in (Shang et al., 2021) proposed a model for multivariate time series forecasting using GNN in which they consider pairwise interactions between features in a node representation. Finally, Cheng et al. (2020) and Gao et al. (2021) propose GNN models with spatio-temporal attention mechanisms to extract features relevant for spatial and temporal relations. After attention is applied they use convolutions and GRU respectively.

3 Approach

3.1 Data sources

For this study, data from air temperature and air humidity sensors installed at a local orchard is used. The devices deployed consist of 12 low-power wireless motes, divided as 8 sensor data nodes and 4 repeaters, connected to a gateway through a SmartMesh IP manager. The wireless sensor network is exposed directly to sun, dust, rain, and snow. With the sensor equipment, we have collected data every 10 seconds at four different heights (one, two, three, and four meters above the ground).

In addition, temperature and humidity data from 10 meteorological stations are used. These stations are located near the orchard, and along with the orchard, data are structured as a graph. This data is collected every hour. Figure 1 shows their geographical location and their representation as a graph.

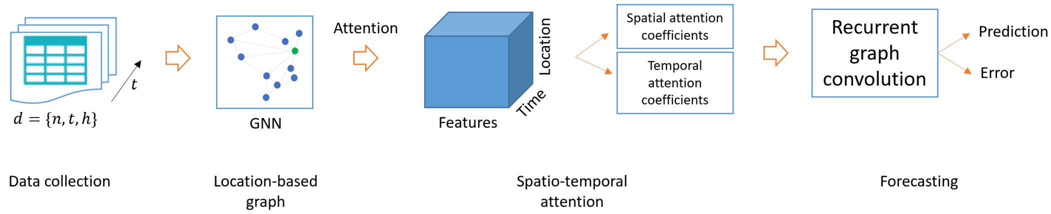

Figure 2: Overview of the GNN spatio-temporal attention model for frost forecast

The period included in the analysis is from September 4th, 2020 to April 5th, 2021. Data from September to February is used to train the model and the rest is used for testing. It is assumed that environmental factors can have an accumulated impact on frost, then the prediction is done using time series. Therefore, we do not randomly split training and test data, so the model can learn time-related patterns.

3.2 Forecasting model

In this study, we develop the GSTA-RGC model to predict the minimum temperature for a fixed number of hours into the future in the local experimental field using the information from sensors and weather stations. For the stations shown in the left of Figure 1, at a single time step the graph is the one shown at the right of the figure. For multiple time steps, we expand this graph into a spatial-temporal graph where feature values for a given node are related to its previous and future values, and its spatial neighbors.

Let's set the notation. n denotes the number of nodes. X is the training data with three dimensions: feature $\{temp, hum\}$, time t, and n. The model uses a window of T steps to forecast the following $t*$ steps. Figure 2 shows the main components of the model: 1) Input data processing, 2) GNN that models geographical interactions, 3) spatio-temporal attention that captures relevant spatial and temporal features, and 4) Recurrent graph convolution to perform the prediction.

Graph construction. There are two types of input data. First, a 3D tensor that includes the node, features at each node (temperature, humidity), and the specific time. In addition, a 2D matrix includes the geographical information of the nodes (latitude, longitude). Using this input data, the geographical proximity is calculated which is used to construct the edges. Nodes are simply a subset of the tensor in a time t. We utilize an aggregation function to reduce the edge updates to a single element. Therefore for a single node, we summarize the interactions with other nodes.

Spatio-temporal attention. We use the GAT model (Veličković et al., 2017) to extract Spatio-temporal similarity features. The idea is to update the embedding information of each node using the aggregate data from its neighbors. Therefore, weather stations and the experimental site receive the temperature and humidity data that previously occurred in near areas. In this way, we are considering that for a specific prediction in a certain area, it is more likely that near areas along with recent measures have more impact on the result than far away nodes and old measures. In particular, the temporal attention layer is a fully-connected layer with ReLU activation and parameters W_t and $\lambda_1 \in [0, 1]$ which represent a temporal penalty factor. The spatial attention layer is applied on top of the temporal layer with parameters W_s and $\lambda_2 \in [0, 1]$ which represents a spatial penalty factor.

Graph neural network forecasting. In the last part of the forecasting model, we applied a Recurrent Graph Convolution. In particular, we use a sequence-to-sequence model (seq2seq) (Sutskever et al., 2014) to map the series used in training with series used for forecasting. The reason for using seq2seq is that in a graph structure, we can perform recurrent graph convolution to handle all series simultaneously, different to usual mechanisms in which each series is treated separately (Shang et al., 2021). Specifically, for each time t, the seq2seq model takes X_t for all series and updates the internal hidden state from H_{t-1} to H_t. The encoder recurrently updates the training data to be included, producing H_{t+T} as a summary. The decoder takes that input and continues the recurrence to include all the testing data for the forecasting phase. The recurrence that accepts input and updates hidden states for all series uses a graph convolution to replace the usual multiplication with a weight matrix. Finally, by the back-propagation process, the loss and prediction values are calculated.

Table 1: Average MAEs and MSEs of time series forecasting models applied to frost forecast.

	MAE			RMSE		
Model	**6hr**	**12hr**	**24hr**	**6hr**	**12hr**	**24hr**
ARIMA	4.73	5.21	5.70	8.75	10.70	13.72
FNN	4.28	4.76	5.74	7.88	8.60	10.75
LSTM	4.03	4.76	5.46	6.93	7.35	9.11
AC-att	3.41	3.79	4.26	5.89	6.64	7.44
STA-C	3.77	4.02	4.83	6.55	8.02	8.71
STA-GRU	3.52	3.73	4.17	6.14	7.20	7.39
GSTA-RGC	**3.08**	**3.25**	**3.86**	**5.42**	**5.79**	**6.19**

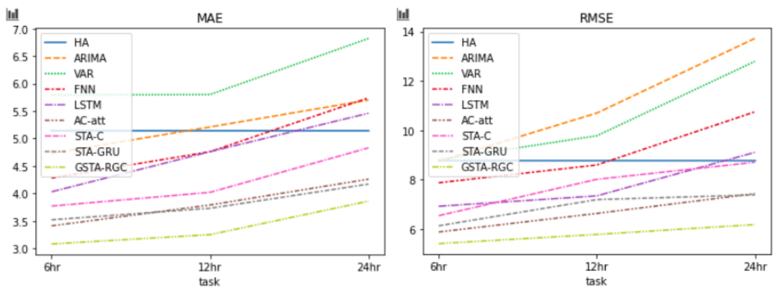

Figure 3: MAE and RSME for forecasting tasks.

4 Results

Baselines. The model proposed in this study is compared with the following forecasting methods: 1) Non-deep learning methods ARIMA; 2) Deep learning methods that produce forecasting for each series separately such as feed-forward neural network and LSTM; 3) Autoencoder with attention mechanism; 4) Variants of this architecture using convolutions (Cheng et al., 2020) and GRU (Gao et al., 2021).

All methods are evaluated with two metrics: mean absolute error (MAE) and root mean square error (RMSE).

Table 1 and Figure 3 show the evaluation of the proposed GSTA-RCN model compared with the baselines. The tasks are to forecast the minimum temperature of the experimental field 6, 12, and 24 hrs in advance.

The proposed model outperforms all the compared baselines for frost forecast in 6, 12, and 24 hrs tasks. It is noticeable that a non-graph model such as an autoencoder with an attention mechanism outperforms GNN with Spatio-temporal attention using convolution and GRU. To improve these results it is necessary to collect more weather data, more weather variables and use more weather stations for modeling geographical and temporal interactions.

5 Conclusions

We have presented a frost forecasting model that uses and optimizes a graph structure between multiple time series using a GNN architecture with a recurrent graph convolution mechanism to process the series simultaneously. The model is completed with spatio-temporal attention to consider spatial relations and extract temporal dynamics.

Acknowledgments

This work was funded by project CORFO 10CEII-9157 Inria Chile and Inria Challenge project OcéanIA (desc. num 14500).

REFERENCES

Dawei Cheng, Xiaoyang Wang, Ying Zhang, and Liqing Zhang. Graph neural network for fraud detection via spatial-temporal attention. *IEEE Transactions on Knowledge and Data Engineering*, 2020.

Ana Laura Diedrichs, Facundo Bromberg, Diego Dujovne, Keoma Brun-Laguna, and Thomas Watteyne. Prediction of frost events using machine learning and iot sensing devices. *IEEE Internet of Things Journal*, 5(6):4589–4597, 2018.

Liya Ding, Kosuke Noborio, and Kazuki Shibuya. Modelling and learning cause-effect—application in frost forecast. *Procedia Computer Science*, 176:2264–2273, 2020.

Junyi Gao, Rakshith Sharma, Cheng Qian, Lucas M Glass, Jeffrey Spaeder, Justin Romberg, Jimeng Sun, and Cao Xiao. Stan: spatio-temporal attention network for pandemic prediction using real-world evidence. *Journal of the American Medical Informatics Association*, 28(4):733–743, 2021.

Mahdi Khodayar and Jianhui Wang. Spatio-temporal graph deep neural network for short-term wind speed forecasting. *IEEE Transactions on Sustainable Energy*, 10(2):670–681, 2018.

snd Santhanavijayan A. Levin Varghese, Gerard Deepak. An iot analytics approach for weather forecasting using raspberry pi 3 model b+. *2019 Fifteenth International Conference on Information Processing (ICINPRO)*, 2019.

F. Pereñíguez-García M. Á. Guillén-Navarro and R. Martínez-España. Iot-based system to forecast crop frost. *International Conference on Intelligent Environments (IE)*, 2017.

Siamak Mehrkanoon. Deep shared representation learning for weather elements forecasting. *Knowledge-Based Systems*, 179:120–128, 2019.

N. Mort and C. L. Chia. Minimum temperature prediction in agricultural area using artificial neural networks. *IEE Colloquium on Neural Networks for Systems: Principles and Applications*, 1991.

Chao Shang, Jie Chen, and Jinbo Bi. Discrete graph structure learning for forecasting multiple time series. *arXiv preprint arXiv:2101.06861*, 2021.

Xingjian Shi, Zhourong Chen, Hao Wang, Dit Yan Yeung, Wai Kin Wong, and Wang Chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, 2015:802–810, 2015.

Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. *arXiv preprint arXiv:1409.3215*, 2014.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.

Pablo Fabián Verdes, Pablo Miguel Granitto, Hugo Daniel Navone, and Hermenegildo Alejandro Ceccatto. Frost prediction with machine learning techniques. In *VI Congreso Argentino de Ciencias de la Computación*, 2000.

Tyler Wilson, Pang-Ning Tan, and Lifeng Luo. A low rank weighted graph convolutional approach to weather prediction. In *2018 IEEE International Conference on Data Mining (ICDM)*, pp. 627–636. IEEE, 2018.

P Y Muck and M. J Homam. Iot based weather station using raspberry pi 3. *International Journal of Engineering & Technology*, 2018.

Si Zhang, Hanghang Tong, Jiejun Xu, and Ross Maciejewski. Graph convolutional networks: a comprehensive review. *Computational Social Networks*, 6(1):1–23, 2019.