Machine Learning and Artificial Intelligence
J.-L. Kim (Ed.)

doi:10.3233/FAIA220433

Frost Forecasting System with Multiple Models of Machine Learning

Shugo YOSHIDA[a,1], Liya DING[a], Kosuke NOBORIO[b] and Kazuki SHIBUYA[b]

[a] *School of Science and Technology, Meiji University, 1-1-1 Higashimita, Tama-ku, Kawasaki, 214-8571, Japan*
[b] *School of Agriculture, Meiji University, 1-1-1 Higashimita, Tama-ku, Kawasaki, 214-8571, Japan*

Abstract. Frost causes damage to crops. Predicting the frost occurrence in advance is highly valuable for practitioners taking possible frost-prevention measures. In our previous study, we have applied machine learning to forecast frost occurrence using two methods. To better support user with a potential trend of frost occurrence in a future period, we propose an integrated system of frost forecast taking the advantages of the two methods.

Keywords. Frost forecast, Predictive models, Time series forecasting, Machine learning

1. Introduction

Frost causes significant damage to crops [1]. Predicting the frost occurrence in advance is highly valuable for practitioners taking possible frost-prevention measures. Previous studies [2-5] show that there is potential for using machine learning to predict the frost occurrence. Since these studies were based on daily time intervals (e.g., the next day), our goal is to enable the prediction of frost occurrence at finer time intervals (e.g., several hours). In our previous study [6-10], we have applied machine learning to forecast frost occurrence using two methods. One is the predictive model based on environment factors predicted through time-series forecast, and the other is prediction by taking into consideration of "cause-effect with delay" relation between early movements of environment factor and late occurrence of frost.

In this study, we first compare the results of the two methods and grasp the advantages of each of them. Then, we propose a frost forecasting system that takes advantages of these two methods with a model adoption strategy. Having the two methods compensating each other, the system can show users how likely a frost event will occur in the next few hours to one day.

The rest of the article is arranged as follows. Section 2 explains the predictive models used in the system and compares their performances. The system overview is provided in Section 3, with an execution example illustrated in the same section. Section 4 gives the conclusions.

[1] Corresponding Author, Shugo YOSHIDA, School of Science and Technology, Meiji University, 1-1-1 Higashimita, Tama-ku, Kawasaki, 214-8571, Japan; E-mail: ce211070@meiji.ac.jp; shugoy45@gmail.com

2. Models involved in the system

2.1. Datasets

We use three datasets *data1*, *data2*, and *data3*, which were observed at a vineyard in Ikeda-cho, Hokkaido, Japan in every five minutes by an actual instrument installed there. The *data1* is collected during the period from April 1, 2020, 00:00 to May 31, 2020, 23:55 (index: 1~17281, frost was observed 21 times), and the *data2* is from April 1, 2021, 00:00 to May 31, 2021, 23:55 (index: 1~17281, frost was observed 14 times). *data3* is from January 1, 2021, 00:00 to March 31, 2021, 23:55 (index: 1~25298, no label of Frost). The *data3* is used for time-series forecasting.

As input factors, Air Temperature (AT), and Vapor Pressure (VP) are used. As the output of prediction Frost takes a binary value of yes/no. The data statistics are summarized in Table 1.

Table 1. The data statistics.

	data1 (2020/4/1-5/31)		data2 (2021/4/1-5/31)		data3 (2020/1/1-3/31)	
	AT	VP	AT	VP	AT	VP
count (every 5 min)	17281	17281	17281	17281	25297	25297
mean	7.218	0.826	7.903	0.885	-5.986	0.3574
std	6.426	0.284	5.796	0.338	8.234	0.1957

2.2. Model A: causal modeling with time delay

To forecast future frost occurrence, we take into consideration of "cause-effect with time delay" relation between early movements of environmental factors and subsequent frost occurrence [6-8]. In this model, the "cause-effect" concept is expressed by using time delay between the input and output. Therefore, we use output labels "Frost after t hour" (*FrostAFt, t=1,2,3…)*.

SVC (Support Vector Classification) is a frontier that best segregates the two classes (hyper-plane/ line) by finding the hyper-plane that differentiates the two classes very well [11]. Therefore, we constructed the classifier using SVC, with inputs (AT, VP) and outputs (*FrostAFt*), having data1 for training and data2 for test.

2.2.1. Result of Model A

Figure 1 shows predicting frost and actual *FrostAF1*.

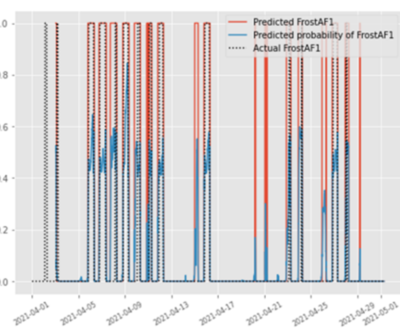


Figure 1. predicting frost and *FrostAF1*.

Causal modeling was used to create 72 different models in 5-minute increments from 5 minutes to 6 hours. The performance comparison of t (5 min ~ 6 hours) is given in Figure 2. The results show that causal modeling performance decreases monotonically as time passes.

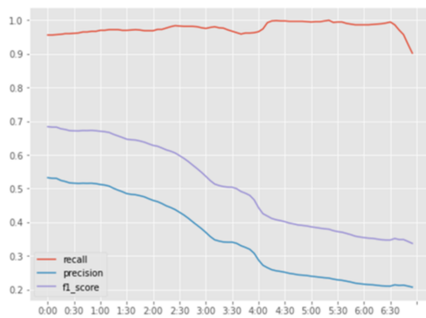


Figure 2. Performance summary of *FrostAFt.*

2.3. Model B: classification based on time-series forecasting

In this method, to forecast future frost, we first need to predict future environment factors. The predicted environmental factors in a future time are then used to forecast the frost at that time [6]. AT and VP have a periodicity, rising and falling within a day. SARIMA is adopted for forecasting future environment factors, which taking the seasonal element into account in ARIMA. SVC is used for frost forecast as a prediction based on the forecast of future environment factors.

First, we use SVC and *data1* to create a classifier that captures an associative relation between inputs AT and VP, and output Frost. Then, *data3* is used as the fitting period for SARIMA to forecast AT and VP for the period in *data2*. Then, the forecasted values of AT and VP are fed to the classifier to predict the frost for the corresponding period. SARIMA forecasts environmental factors for the next 24 hours every day. Parameters are automatically adjusted using auto-ARIMA.

2.3.1. Result of Model B

Figure 3 shows Air Temperature and Vapor Pressure predicted by SARIMA for each 12-hour period. Figure 4 shows the results of the frost prediction using these data.

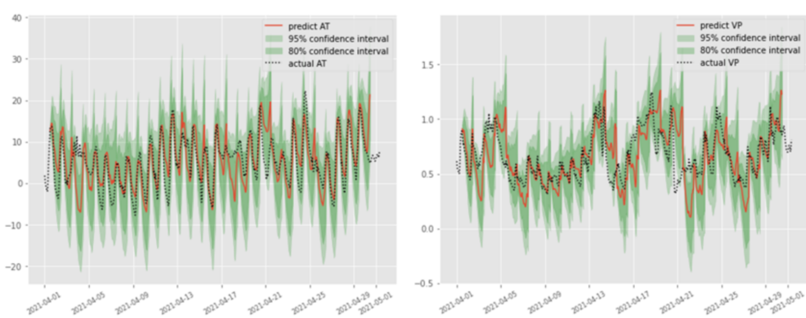


Figure 3. Time-series forecast of AT and VP.

Figure 5 gives the performance summary of Model B when it is used to forecast the occurrence of frost from 1 hour to 24 hours. This result shows that Model B has a slow decline in performance as time passes.

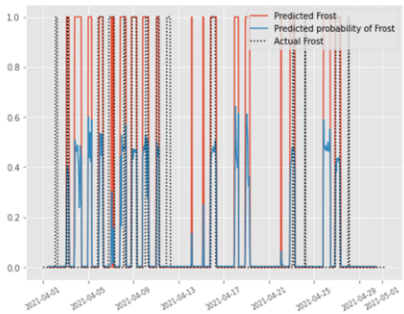


Figure 4. Forecasting Frost by Model B.

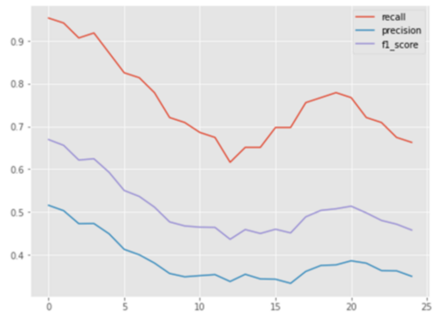


Figure 5. Performance summary of *FrostAFt* by Model B.

2.4. Performance comparison of models

Figure 6 shows a performance comparison of Model A and Model B. We can see that Model A performs better up to two hours, but Model B performs better after that time.

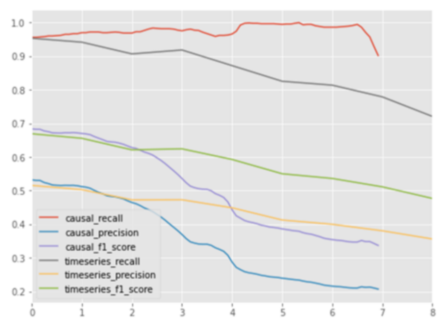


Figure 6. Performance comparison of Model A and Model B.

3. Forecasting system

3.1. System architecture

Using these two models, we propose a system that can announce how likely frost will occur after hours from now. The forecasting system overview is shown Figure 7, where the training of Model A and Model B [4,6] is omitted. In the ensemble of Model A, we combined the 72 different models of prediction after 5 minutes to 6 hours to obtain frost trend from the current time to 6 hours later. For example, if the current time is 9:00, a total of 72 models (a model to forecast 5min later, a model to forecast 10min later...a model to forecast 1 hour later...a model to forecast 6 hours later) are used to predict from 9:05 to 15:00 in 5-minute intervals.

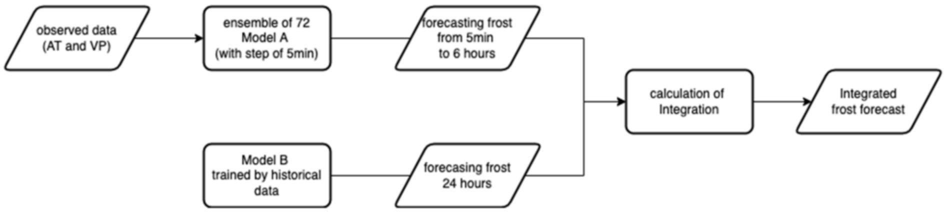


Figure 7. System overview.

Model B consists of two stages of processing. Future AT and VP are forecasted first by time series forecasting, and the forecasted results are fed to a SVC trained using historical data to forecast frost 24 hours ahead.

3.2. Model adoption strategy

From Section 2.4, It turns out that two hours is the borderline. Therefore, an integrated forecasting is calculated as follows.

$$Integrated\ forecast = \begin{cases} Model\ A, & t \leq 2 \\ (Model\ A + Model\ B)/2, & 2 < t \leq 6 \\ Model\ B, & 6 < t \end{cases} \qquad \ldots(1)$$

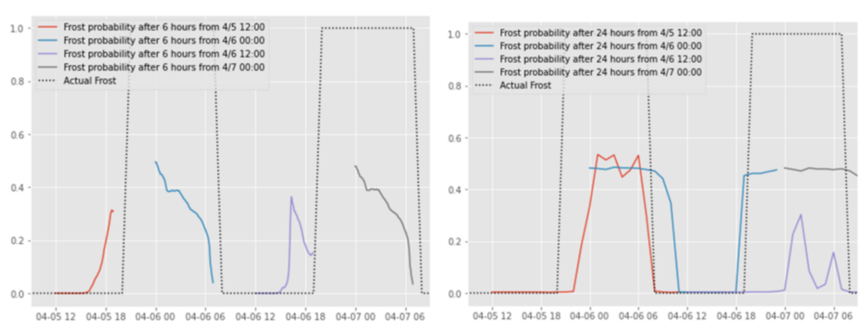
In addition to the above, in order to smooth data connections, the integrated forecast takes a 3-hour simple moving average.

3.3. Illustration of system execution

The following four points of time were used to start the prediction.

- from 4/5, 2021, 12:00
- from 4/6, 2021, 00:00
- from 4/6, 2021, 12:00
- from 4/7, 2021, 00:00

The results of this prediction by ensemble are shown in Figure 8 (a). By using Model B, Figure 8 (b) shows the frost forecasting of 24 hours from the same four points of time when Model A is activated. The integrated result is shown in Figure 9.



(a) Model A. (b) Model B.

Figure 8. Frost forecasting by two models at different time points.

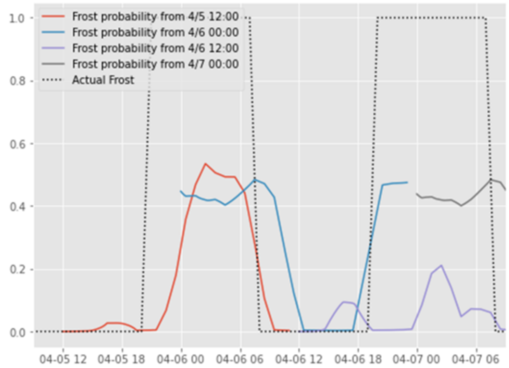


Figure 9. Integrated trend of frost by multiple models.

4. Conclusion

We proposed a frost forecasting system consisting of two predictive models developed by different modelling methods: Model-A (Causal modeling with time delay), and Model-B (classification based on time-series forecasting). We also further confirmed that Model-A is superior in short period while Model-B is advantageous in a longer period. Having Model-A and Model-B compensating each other, the proposed forecasting system can offer user a better forecast accuracy for overall trend of frost occurrence in next 24 hours.

References

[1] Nationnal Geographic Society. Frost. https://www.nationalgeographic.org/encyclopedia/frost.
[2] Verdes PF, Granitto PM, Navone HD, Ceccatto HA. Frost prediction with machine learning techniques. VI Congreso Argentino de Ciencias de la Computación, 2000.
[3] Fuentes M, Campos1 C, García-Loyola S. Application of artificial neural networks to frost detection in central Chile using the next day minimum air temperature forecast. Chilean J Agric Res. July-Sep 2018;78(3):327-338.
[4] Lee H, Chun JA, Han H-H, Kim S. Prediction of frost occurrences using statistical modeling approaches. Adv Meteorol. 2016; 2075186. (http://dx.doi.org/10.1155/2016/2075186), 2016.
[5] Diedrichs AL, Bromberg F, Dujovne D, Brun-Laguna K, Watteyne T. Prediction Of Frost Events Using Machine Learning and IoT Sensing Device. IEEE Internet Things J. Dec. 2018;5(6): 4589-4597.
[6] Shugo Y, Yosuke T, Kenta O, Liya D, Kosuke N, Kazuki S. A study of causal modeling with time delay for frost forecast using machine learning from data. AICI'2022, Vietnam, Jan. 14, 2022.
[7] Liya D, Yosuke T, Kosuke N, Kazuki S. Frost forecast: a practice of machine learning from data. Int. J. Reasoning-based Intelligent Systems, Vol. 13, No. 4, pp.191-203, Sep. 2021.
[8] Liya D, Yosuke T, Shugo Y, Kenta O, Tatsuya T, Yuto M, Kosuke N, Kazuki S. Ensemble causal modelling for frost forecast in vineyard. Procedia Comp Sci. 2021;192: 3194-3203.
[9] Yosuke T, Shugo Y, Kenta O, Liya D, Kosuke N, Kazuki S. Method of Data Labelling Using Machine learning for Frost Forecasting. ICICIC'2021, Japan, Sep. 15, 2021.
[10] Kenta O, Yosuke T, Shugo Y, Liya D, Kosuke N, Kazuki S. Prediction of Frost Occurrences Using Time Series Forecast and Machine Learning. ICICIC'2021, Japan, Sep. 16, 2021.
[11] Sunil Ray. Understanding Support Vector Machine(SVM) algorithm from examples (along with code). https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/#h2_4, 2017.