How to Optimize the Decision Making Using Ensemble Probabilistic Forecasts

Hui-Ling Chang[1], Shu-Chih Yang[2], Huiling Yuan[3,4], Pay-Liam Lin[2], and Yu-Chieng Liou [2]

1. Meteorological Satellite Center, Central Weather Bureau, Taipei, Taiwan;
2. Department of Atmospheric Sciences, National Central University, Jhong-Li, Taiwan
3. School of Atmospheric Sciences, and Key Laboratory of Mesoscale Severe Weather / Ministry of Education, Nanjing University, Nanjing, Jiangsu, China;
4. Jiangsu Collaborative Innovation Center for Climate Change, China

Introduction

Different from the deterministic forecasts (DFs), the ensemble probabilistic forecasts (EPFs) consider uncertainties during the forecast process, and convey the uncertainty information to the users using probability. For a skillful ensemble prediction system (EPS), the divergence of ensemble forecasts causes small forecast probabilities (P_f), which indicate that the possibility of the occurrence of a specific weather event is low and the forecast uncertainty is large. However, compared with the DFs with indication of “Yes” or “No” only, can such probabilities or uncertainties information really benefit the users or confuse them in decision making? For example, the information provided by EPFs is “the chance of heavy rainfall tomorrow is 70%”. Users usually have difficulty in making decisions based on such information because they’re not sure whether a probability of 70% indicates the evet will happen or not. Therefore, users most frequently wonder whether an optimal probability threshold (P_t) can be provided, so that when the P_f exceed the optimal P_t, they are able to take preventive actions, such as closing roads, harvesting crops in advance, and shutting offices and financial markets.

Besides, users also wonder how to best use the EPFs for risked-based decision making. For example, a flight controller must decide whether or not to change the runway of incoming flight based on the forecast of severe weather. Changing the runway before severe weather can avoid a plane crash, but disrupt airline schedules. Is there some guidance for the decision-making? Another example is some crops’ harvest season falls during the typhoon season, such as pomelos and Chinese dates. Pomelos are vulnerable to strong winds, and may be knocked down by strong typhoon winds. In contrast, Chinese dates are sensitive to the heavy rainfall resulting from typhoons. Heavy rainfall will result in date cracking and reduce the quality of dates. Therefore, the farmers must decide whether or not to harvest in advance to minimize their losses. Such kinds of problems are associated with users’ cost and loss, which are the so-called analysis of economic value (EV).

Analysis of economic value

The EV of a forecast system (Richardson 2000) is defined as:

$$EV = \frac{E_{climate} - E_{forecast}}{E_{climate} - E_{perfect}} \quad (1)$$

where $E_{climate}$, $E_{forecast}$ and $E_{perfect}$ are the expected expenses of a user who takes preventive action based on the climatological information, a forecast system, and a perfect deterministic forecast system, respectively. A perfect forecast system means that it always provides accurate predictions for the occurrence and non-occurrence of a weather event. According to the above definition, the EV can be interpreted as the relative performance taking the climatological information as a baseline.

In the EV analysis, we assume that a user takes preventive action totally depending on forecast information (i.e., this user takes action only when the event is predicted). Therefore, based on the past long-term forecast performance, we can evaluate the EV of a forecast system using a 2×2 contingency table (Table 1). Table 1 lists the relative frequencies and expected expense of a user for four possible conditions, where C is the cost of preventive action, L ($L = L_p + L_u$) is the total loss caused by weather events, including the protectable loss (L_p) and unprotectable loss (L_u) after taking preventive action. The cost-loss ratio (expressed as $r = C/L_p$) is unique for each user since the corresponding C and L_p are different. Given that users take preventive action only when $L_p > C$, the value of r is between 0 and 1.

Using the definition in (1), Zhu et al. (2002) showed that EV can be expressed as

$$EV = \frac{\min[\bar{o}, r] - (h + f)r - m}{\min[\bar{o}, r] - \bar{o}r}. \qquad (2)$$

This equation shows that EV is related not only to the forecast performance (i.e., forecast parameters h, f, and m) but also to the climatological frequency ($\bar{o}$) of a weather event and the cost-loss ratio (r) of a user.

In this study, we take the Central Weather Bureau’ operational 0-6 h probabilistic quantitative precipitation forecasts (PQPFs) as an example to illustrate how to optimize the decision-making through the EV analysis. The PQPFs are generated from the LAPS EPS, which has 12 members (Chang et al. 2012). Eight typhoon cases in 2008 and 2009 were used to evaluate the EV obtained by users.

With 12 members, the LAPS EPS provides 12 P_t values (i.e., 1/12 to 12/12). Fig. 1a shows a set of EV curves at the 10 mm (6 h)$^{-1}$ threshold using 12 P_t generated from the LAPS 0-6 h calibrated PQPFs. The choice of P_t has a decisive influence on the EV obtained by users. For example, users with r = 0.1 will obtain 36% of EV based on P_t = 2/12 (Fig. 1b), but only 16% of EV can be obtained if they choose a more strict P_t (say, 4/12). But it they do not take preventive action until the P_f exceeds 5/12, they will gain nothing at all from the LAPS forecasts. Murphy (1977) showed that if perfectly reliable (i.e., unbiased) forecasts are adopted, the optimal P_t for maximizing EV is equal to the r value of users. Different users should choose the optimal P_t based on their r so that their EV can be maximized.

Application of EV analysis to decision-making

In Zhu et al. (2002), an example for using the EV analysis was demonstrated by explicitly knowing the r value of users. Unfortunately information of users' r is sometimes implicitly known. In this situation, can users still optimize their decision-making to obtain the maximum EV? For example, before the arrival of typhoon, the date farmers must decide whether or not to harvest in advance to minimize their losses. Compared with normal harvesting, the action of harvesting in advance does not seem to pay any cost; however, it may lead to a hidden loss since unripe dates are sold at a lower price. Therefore, two conditions should be considered for price drop: premature harvest and being affected by heavy rainfall. The ratios between the reduced and original prices for these two conditions are denoted as R_1 and R_2.

Assume two weeks before dates ripen, the LAPS forecast indicates a 50% probability of heavy rainfall (P_f) in the coming six hours. Since the dates are nearly ripe, R_1 is equal to 80%. In addition, R_2 is 40% and the total price of ripe dates is A. We can use Table 2a to analyze the expected expense of the farmers in four possible situations. In situation 1, harvesting in advance and heavy rainfall not occurring. In this situation, the expected expense of the farmers is equal to the cost of preventive action (C), but how to get it? The key point is that the premature harvest will lower the price of dates and thus reduce the total income of the farmers. The reduction of income should be regarded as the cost of preventive action; therefore, the expected expense of the farmers is C = (1-80%)A. Following the same concept, the expected expense is calculated by considering the reduction of the farmers' total income in the remaining three situations. With this contingency table, we can get the r value of the farmers is about 0.33; therefore, if P_f =50%, P_f is greater than the optimal P_t and the farmers should harvest in advance. In contrast, if P_f is only 10%, P_f is smaller than the optimal P_t and the farmers do not need to take action.

By contrast, if the typhoon may hit four weeks before dates ripen and the premature price is only 60% of the mature price, should the farmers still harvest in advance? The r value of the farmers can be calculated in the same way and it's about 0.67. Therefore, premature harvest is required only when P_f >=67%. Furthermore, if the farmers should harvest in advance based on the forecast information, they may wonder what percentage of dates should be harvested to minimize their losses. The answer is "full harvest". Please refer to Chang et al. (2014) for a detailed description of this interesting question.

Future works

From the experience we gain, it's feasible to extend the application of decision-making from the LAPS to other EPS, or from the rainfall field to other meteorological fields. For example, fishermen may regard wind speed and wave height as vital indicators in terms of their economic benefits, and the wine industry is concerned about sunshine hours, temperature, and humidity.

In addition, we are working on a dynamic cost-loss ratio model for the EV analysis.

Although the static cost-loss ratio model in this study is very simple and useful, it cannot be applied to all real-world scenarios. For example, in the case of date harvesting some factors are not considered. This includes wages of labor for harvesting, which might become more and more expensive as the lead time gets shorter and shorter. Therefore, it's necessary to extend the static model to a dynamic one to take into account multiple occasions with increased complexity for real-world scenarios.

Reference

Chang, H. L., H. Yuan, P. L. Lin, 2012: Short-Range (0-12h) PQPFs from Time-Lagged Multimodel Ensembles Using LAPS. *Mon. Wea. Rev.*, **140**, 1496–1516.

—, S.-C. Yang, H. Yuan, P. L. Lin and Y. C. Liou, 2015: Analysis of relative operating characteristic and economic value using the LAPS ensemble prediction system in Taiwan area. *Mon. Wea. Rev.*, **143**, 1833–1848.

Murphy, A.H., 1977: The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation. *Mon. Wea. Rev.*, **105**, 803–816.

Richardson, D. S., 2000: Skill and relative economic value of the ECMWF ensemble prediction system. *Quart. J. Royal Meteor. Soc.*, **126**, 649–667.

Zhu, Y., Z. Toth, R. Wobus, D. S., Richardson, and K. Mylne, 2002：The economic value of ensemble-based weather forecasts. *Bull. Amer. Meteor. Soc.*, **83**, 73–83.

TABLE 1. Contingency table for forecasts and observations of a binary event.

		Forecast / action	
		Yes	No
Observation	Yes	Hit (*h*) Mitigated loss (C+L_u)	Miss (*m*) Loss (L_p+L_u)
	No	False alarm (*f*) Cost (C)	Correct rejection(*c*) No cost (N)

TABLE 2. Contingency tables in case of (a) two weeks (R_1=80%) and (b) four weeks (R_1=60%) before dates ripen. Assume that R_2=40% and the total price of ripe dates with normal harvest is A.

(a)

		Forecast / action	
		Yes	No
Observation	Yes	Hit (*h*) (2) C+L_u=(1-80%)A	Miss (*m*) (3) L_p+L_u=(1-40%)A
	No	False alarm (*f*) (1) C=(1-80%)A	Correct rejection (*c*) (4) N=0

(b)

		Forecast / action	
		Yes	No
Observation	Yes	Hit (*h*) *(2)* C+L_u=(1-60%)A	Miss (*m*) *(3)* L_p+L_u=(1-40%)A
	No	False alarm (*f*) *(1)* C=(1-60%)A	Correct rejection (*c*) *(4)* N=0

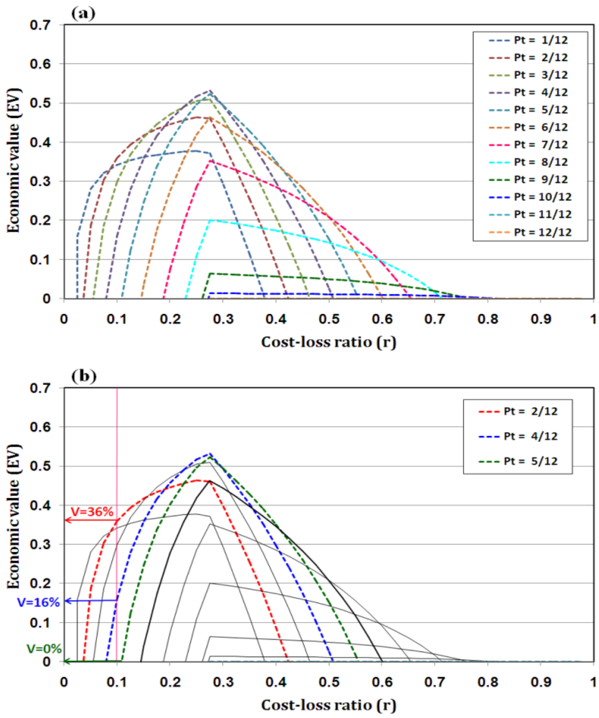

FIG. 1. For the LAPS 0-6 h PQPFs at the 10 mm (6 h)$^{-1}$ threshold, (a) economic value (EV) against the cost-loss ratio (r) at different probability thresholds (P_t=1/12 to 12/12), (b) illustration of economic values obtained by users with r = 0.1 when adopting different P_t values (P_t =2/12, 4/12 and 5/12).