NAPSTER: NONLINEAR ANALYSIS AND PREDICTION STATISTICS FROM TIME-SERIES AND ENSEMBLE-FORECAST REALIZATIONS

A NERC FUNDED TECHNOLOGY TRANSFER PROJECT BY THE **CENTRE FOR THE ANALYSIS OF TIME-SERIES** AT **THE LONDON SCHOOL OF ECONOMICS DEPARTMENT OF STATISTICS**

Centre for the Analysis of Time Series
London School of Economics
Columbia House
Dept of Statistics
Houghton Street
London
WC2A 2AE

Tel: +44 (0) 20 7955 6015
Fax: +44 (0) 20 7955 7416
E-mail: cats@lse.ac.uk
www.lse.ac.uk/collections/cats/

Natural Environment Research Council
Polaris House, North Star Avenue
Swindon, SN2 1EU
Tel: 01793 411500
Fax: 01793 411501
www.nerc.ac.uk

NAPSTER: NONLINEAR ANALYSIS AND PREDICTION STATISTICS FROM TIME-SERIES AND ENSEMBLE-FORECAST REALIZATIONS

A NERC FUNDED TECHNOLOGY TRANSFER PROJECT BY THE **CENTRE FOR THE ANALYSIS OF TIME-SERIES** AT **THE LONDON SCHOOL OF ECONOMICS DEPARTMENT OF STATISTICS**

CONTENTS

INTRODUCTION

THE DAY-TO-DAY OPERATIONS OF MOST UK BUSINESSES AND PUBLIC SERVICES ARE AFFECTED BY THE WEATHER

Domestic gas usage, for example, is largely dependent on the ambient temperature, as are sales of foodstuffs such as cold drinks and ice-cream. The weather can demand that action be taken. Local authorities are required to grit their roads when there is sufficient risk of freezing conditions. The weather can also provide opportunities, as retailers of umbrellas, BBQ equipment and Wellington boots will no doubt testify.

Weather risk is the generic term given to weather specific impacts and opportunities that arise in the operations of business and public sector organisations. There are many different forms of weather risk, ranging from the immediate risk to human life arising from an extreme storm to changes in demand for basic products driven by relatively small differences in atmospheric conditions. Exposure to weather risk requires decisions to be made, sometimes in the form of a yes/no decision, sometimes in the form of a how much question: whether or not to take action, how much stock to order, what position to take. The quality of that decision depends not only on the decision maker but also on the information provided to support that decision.

In recent years meteorologists have developed and implemented new forecasting techniques. These state-of-the-art forecasts are probabilistic in nature, not merely guesses of what might happen but providing information on what is likely to happen. These forecasts are largely under-utilised, in part because many organisations are unable to process probabilistic information in its present form. Probabilistic forecast information needs to be better presented if its potential value is to be realised.

A framework for providing weather forecast information, along with the uncertainty in that information, is presented here for operational use. The following five snapshots give examples of the weather risk faced by a wide range of public and private business sectors. The use of weather forecasts for decision support is then discussed and examples of how probabilistic forecast information might be usefully communicated are presented.

SUPERMARKET GROCERY SALES

PERSONAL EXPERIENCE TELLS US THAT WEATHER CONDITIONS AFFECT OUR OWN GROCERY SHOPPING DECISIONS

Sales of many products are inherently weather dependent and store managers currently use weather forecast information qualitatively as part of their day-to-day decision making.

The relationship between recent local weather and sales in general is complicated. Total sales are dependent on many different types of customer and many different types of product, only some of which may be significantly exposed to the weather. Furthermore, demand is influenced by a multitude of factors other than the weather such as the day of the week, advertising campaigns, in-store promotions and even the weather forecast itself.

Deliveries are common on lead-times of 3 days and store managers are interested in not only maximising sales but also in reducing costs. There is a complex trade-off between shelf life, delivery and expected demand for which weather is one, potentially important, but generally unquantified component. Relevant, robust weather forecasts may well be able to assist store managers in making better quantitative decisions regarding stock levels of weather dependent items. Before this can be investigated, weather dependency must first be established. Are grocery sales significantly affected by extreme weather events such as runs of excessively hot days, or are these effects insignificant compared to core business activity? If weather dependency can be established, how accurate would weather forecasts have to be in order to be useful?

OUTDOOR LEISURE ACTIVITIES

A NUMBER OF OUTDOOR LEISURE ACTIVITIES, FROM HORSE RACING TO THEME PARK ATTENDANCE, ARE SIGNIFICANTLY IMPACTED BY THE WEATHER

Some of these impacts are similar to those faced by supermarkets, i.e. weather related demand for consumables such as food and drink. There are, however, other mitigable weather risks that are specific to each industry.

The horse racing industry, for example, is concerned with the going - the quality of the turf. Ideally, the going should be good; not too hard, not too soft. Overnight freezing conditions result in hard going and when the going is too hard, racehorse owners are likely to withdraw their horses for fear of injury. In the worst cases the entire meeting has to be abandoned resulting in significant loss of revenue. The effects of frost can be mitigated by putting down overnight frost covers, but this action is costly. During summer months, a lack of rainfall can also result in hard going. This can be addressed by watering but short-term irrigation needs to take into account any potential large rainfall events. Summer storms on an already saturated course can result in soft going which is also undesirable.

Racecourses are well aware of the particular weather "events" that are of interest to them and have extensive records of on-course atmospheric conditions. This site-specific data, along with commercial weather forecasts, is used by managers to support decisions regarding racetrack maintenance. Site relevant data can be used to enhance commercially available weather forecasts to give more relevant and robust warnings of events such as overnight frosts and thus better inform intervention decisions.

This is done by purchasing gas in the forward market. The amount of gas purchased depends on the weather forecast. A change in the weather forecast results in a change in expected demand and gas traders frequently adjust their position accordingly. For long lead-times, say 8-10 days, the forecast can change a good deal in the time to actual delivery of the gas.

Fluctuations in the forecast result in trades that not only cost money in transaction costs, but also due to price changes. An increase in forecasted demand may require that gas be purchased when the market price is high. Similarly, a decrease in forecasted demand may require excess gas to be sold when the price is low. These trading decisions are made every day until the gas is delivered. Excessive trading costs increase operating costs which are ultimately passed on to the consumer.

Gas traders are numerate users of weather forecast information. The weather forecasts they use often provide no intrinsic information regarding the uncertainty in that forecast. Given additional robust information on the likely errors in a forecast, gas traders can better position themselves in the market and reduce unnecessary trades.

DOMESTIC GAS DEMAND IS HEAVILY INFLUENCED BY THE WEATHER AND SUPPLIERS MUST ENSURE THAT THEY HAVE SUFFICIENT SUPPLIES IN PLACE TO MEET DEMAND

ENERGY

FOOD RETAIL

RETAIL FOOD OUTLETS SPECIALISING IN TAKE-AWAY BREAKFASTS AND LUNCH ARE EXPOSED TO WEATHER DEPENDENT DEMAND

Purchasing decisions are largely carried out on the day and influenced by the weather conditions at the point of sale.

Deliveries are made every day based on orders received up to 3 days in advance and the shelf life of fresh ingredients is very short. Store managers are concerned not only with increasing income, but also in reducing waste. Many stores have sought to reduce their exposure to weather risk by offering a range of products. Coffee shops now sell iced coffee drinks and sandwich shops sell soup and hot sandwiches to counter the effects of changing weather.

Extreme weather events can provide opportunities for food outlets in particular locations. The prolonged period of fog at Heathrow during the Christmas period of 2006 created an unexpected large demand for food and drink. Given forewarning of these kinds of events stores could better prepare themselves. How difficult is it to provide forecasts for such rare events that are both accurate and reliable enough in order for businesses to take action? Given a better understanding of how sales are related to fluctuations in the day-to-day weather we can investigate whether or not there is useful information in existing weather forecasts to help store managers better plan their operations.

WINTER ROAD MAINTENANCE

LOCAL AUTHORITIES THROUGHOUT THE UK ARE RESPONSIBLE FOR ENSURING THAT ROADS ARE SAFE

Local authorities are concerned with a particular event, namely "freezing conditions", where the roads are covered, or partially covered, in ice. When there is a perceived threat of freezing conditions, gritting trucks are dispatched to spread salt on the roads.

During winter months the dispatching decision has to be made at least once every day. Activity is relatively frequent, with gritters being dispatched during winter months, on average, around one night in every three. Information on the current atmospheric conditions and short-term weather forecasts are used and the lead-time for decision making is relatively short, with trucks dispatched within hours.

A good deal of expertise has been developed around winter road maintenance. The potential loss of human life and damage to property associated with inaction are much larger than the costs of deployment. Consequently, road maintenance managers are extremely risk averse and any changes to operational processes would need to be justified with significant improvements in skill.

DECISION SUPPORT

DIFFERENT ORGANISATIONS ARE EXPOSED TO A WIDE RANGE OF WEATHER RISK

For some, weather is a principle concern in day-to-day operations and in general these businesses are already consumers of weather forecast information. For others, the weather is one component amongst many other important factors. In order for weather forecasts to be used more effectively the relationship between weather and operations needs to be well understood and quantified.

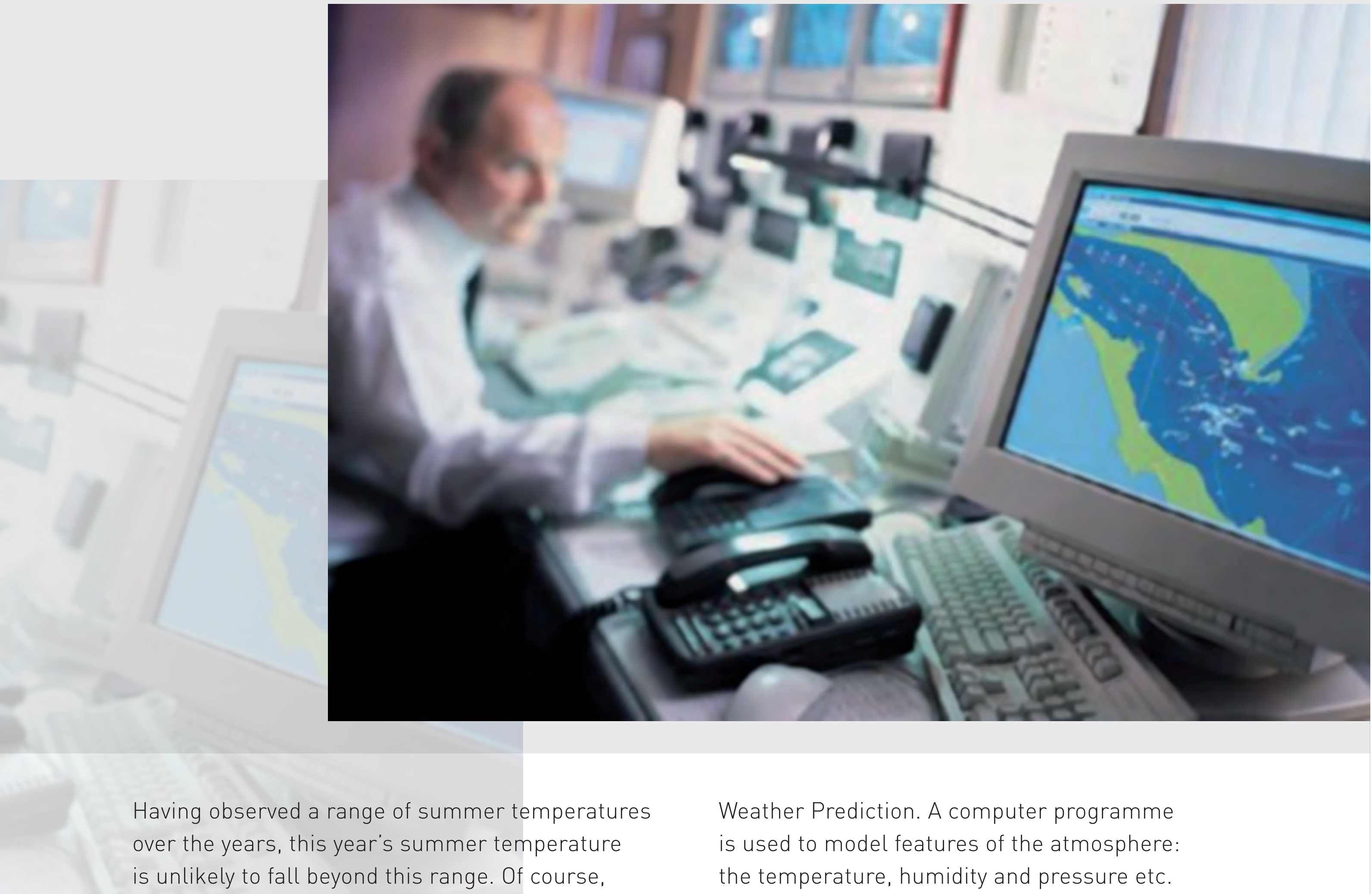
The response to weather risk is also varied. For many companies the choice is a binary decision: action is either taken or it is not. For others, such as gas traders and supermarkets, there is a further notion of quantity: how much gas to purchase or how much stock to order. Moreover, these decisions are rarely solely dependent on the weather. A comprehensive decision support tool encompassing all possible relevant decision variables is neither feasible nor necessary.

Organisations develop their own expertise regarding their day-to-day decision making. It is extremely unlikely that an automated decision making system could outperform the existing processes supported with better quality weather forecast information and a better understanding regarding the uncertainty in that weather forecast information.

Different organisations also have very different attitudes towards risk. Commercial organisations may well be able to use forecast information that gives insight into likely outcomes but is not 100% correct. Organisations that are extremely risk averse have demands on forecast information that may be beyond that which is currently available. Nevertheless, it is useful to explore alternative methods of presenting forecast information and to map out the limits of useful forecast information.

WEATHER FORECAST INFORMATION

THE HISTORICAL, OR CLIMATOLOGICAL, RECORD OF WEATHER INFORMATION PROVIDES A BASELINE FOR WHAT TO EXPECT



Having observed a range of summer temperatures over the years, this year's summer temperature is unlikely to fall beyond this range. Of course, extreme events do happen and records are broken. Hottest summers and driest winters appear to occur ever more frequently, but even in a slowly changing climate, historical records provide a useful guide for the possible weather.

Although very reliable, climatological forecasts are not particularly skilful; they are rarely misleading but may be too vague to provide useful decision support. In order to say more regarding what the weather is likely to be, methods based on a scientific understanding of the atmosphere are generally used. Modern weather forecasts are almost exclusively the product of Numerical Weather Prediction. A computer programme is used to model features of the atmosphere: the temperature, humidity and pressure etc. Changes in these quantities are then described by equations based on the laws of physics.

The model requires a starting point called an initial condition. This initial condition is an estimate of the Earth's current weather based on satellite images and observations from weather stations around the world. The model then evolves this initial condition into a prediction of the model's future weather using mathematical equations. The quality of the forecast depends on the ability of the model to describe real world atmospheric conditions and the suitability of the initial condition.

OVERCONFIDENCE & UNCERTAINTY

TRADITIONAL NUMERICAL WEATHER FORECAST INFORMATION IS OFTEN PACKAGED AS A "BEST GUESS FORECAST"

These forecasts undoubtedly have skill over climatological forecasts for lead-times on the order of days. Sometimes, however, the best guess weather forecast can be misleading. The great storm of 1987, which was dismissed on national television, is a good example. The best guess forecast in that instance was wrong, and the storm caught many people by surprise largely due to the misplaced confidence of the weather forecasters. Naively taking a best guess forecast at face value can be very costly.

In practice, users develop their own confidence levels of a best guess forecast system based on trial and error. This heuristic interpretation obscures potentially useful information. It has been recognised that our ability to predict the weather changes from day-to-day and as such our confidence in that forecast should vary accordingly. Modern weather forecasts attempt to convey this uncertainty information using the language of probability. Rather than giving a definitive statement on the temperature or rainfall in the coming days, forecasts refer to the chance of an event. Suitably presented, these forecasts contain potentially valuable information for decision makers.

Businesses are frequently exposed to just one-side of weather forecast error, with temperatures colder than forecast being more costly than vice versa for example. Given an estimate of the size and direction of the uncertainty in a forecast, a decision maker can better choose how to act on that information. If the chance of rain is too high, even though the most likely outcome is no rain, an alternative to outdoor activities can be investigated. Similarly, additional electricity generation may be put on stand-by if there is a small chance of a significant cold snap.

FORECAST ERROR

WHY DO FORECASTS GO AWRY? FORECAST ERRORS CAN ARISE DUE TO A NUMBER OF FACTORS

A principle source of error is uncertainty in the state used to initialise the model. The best initial condition is never known at the time of initialisation. Of course, if the best initial condition were known it would be used. The problem is not that the correct initial condition is not used, but rather that the correct initial condition is not known. There is unavoidable uncertainty in how best to initialise the model.

Another principle source of error is due to imperfections in the model itself. The equations and their representation of the atmosphere are approximations of reality and are unable to capture every feature of the weather at all length scales. For example, numerical weather models have difficulty describing precipitation. The physical processes that are important in rainfall: cloud formation, the formation of ice crystals etc, occur on length scales that the models do not resolve. Instead, weather models use simpler models to capture these small-scale processes. Effectively, this means that no initial condition exists for which a given model can provide a realistic looking forecast. While this model inadequacy can be fundamentally addressed by developing better models and sometimes even detected by looking at more than one of today's models, initial condition uncertainty can be addressed using existing technology.

Numerical weather predictions are not expected to be perfect, since the model is an approximation and there is uncertainty in how the model is initialised. To guard against over-confidence a forecast should be presented with an estimate of the uncertainty in that prediction.

ENSEMBLE FORECASTING

ONE APPROACH TO QUANTIFYING UNCERTAINTY IS A TECHNIQUE CALLED ENSEMBLE FORECASTING

Rather than make a single forecast using the best guess initial condition, a number of forecasts are made each consistent with the observations. The result is a distribution of distinct forecasts for the same day. When the spread is wide, i.e. the forecasts are very different, the ensemble indicates increased uncertainty and confidence is reduced.

Ensemble forecast information is essentially probabilistic even if it does not provide probabilities that are useful as such. The exact outcome is not known with certainty and to pretend otherwise would be misleading. Statements must be qualified by a degree of confidence: "there is a good chance it will rain", or "it is likely that there will be above average temperatures". These statements are sometimes further quantified using probabilistic statements such as: "there is a 50% chance that it will rain", or "there is a 1-in-4 chance of very strong winds". A forecast that says that there is an 80% chance of rain when the outcome is no rain is not necessarily a bad forecast. An 80% chance implies that one expects the event to not occur 1-in-5, or 20%, of the time that that forecast is issued. A forecast system that consistently says there is an 80% chance of rain when no rain occurs is misleading because it is overconfident. Similarly, a forecast system that consistently says there is a 20% chance of rain when rain occurs is misleading because it is under confident.

Given this probabilistic information, how should a decision maker react? Whether or not action is taken depends on a number of factors, not least the risk appetite of the decision maker, the costs of taking action and the potential rewards at hand. The level of skill required for a forecast to be considered useful depends on the application. Some forecasts may be able to provide that information and some cannot. Different users with different tasks and risk tolerances may want different things from their forecast information: confidence that an event will occur given a warning, or confidence that an event is unlikely to occur given no warning.

FORECAST INTERPRETATION

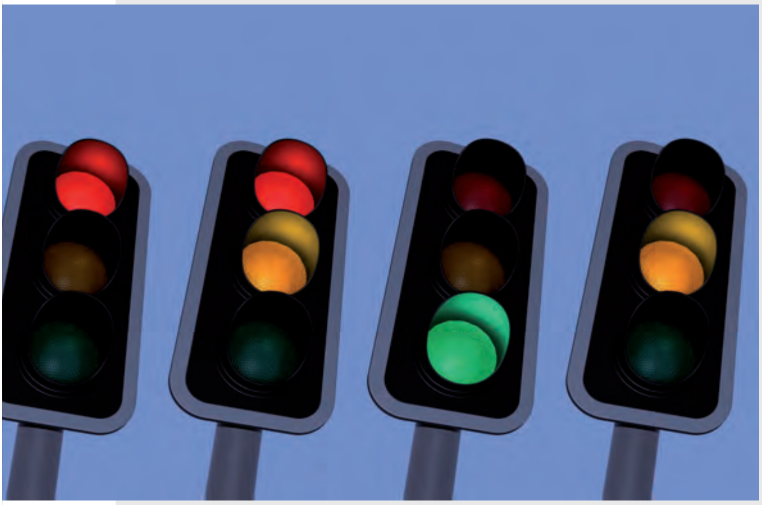
ENSEMBLE TECHNIQUES OFFER THE POTENTIAL TO MAKE FORECAST STATEMENTS THAT QUANTIFY THE CONFIDENCE WITH WHICH THEY ARE ISSUED

Consequently, ensemble forecasts are also able to convey when they do not have skill compared, say, to the climatology. This information is useful as it helps prevent action due to overconfidence in the forecast.

Model output is used as the basis for the issued forecast, but is rarely issued without additional interpretation. The UK Met Office, for example, employs many experienced meteorologists to process the model output for public consumption. Naive interpretations of an ensemble forecast have repeatedly proven to not be robust. For example, issuing the fraction of ensemble members that display precipitation as a probabilistic forecast for rain will rarely be robust due to fundamental errors in the underlying model. In practice we find that the observation falls outside the ensemble too frequently to be consistent with a probability forecast.

Systematic errors can be addressed, to some degree, by interpreting the forecast information statistically to compensate for repeatedly observed mistakes. A prerequisite for any statistical interpretation is a suitably large historical archive containing a number of examples of forecast error. Given enough data, the statistical interpretation can greatly enhance the utility of ensemble forecasts for decision support. These interpretations are less effective for rare events, since by definition there are few instances of rare events in the historical archive.

CONSISTENT WARNINGS



IN THE SIMPLEST CASE, THE USER IS INTERESTED IN A CRITICAL VALUE OF SOME VARIABLE AND WOULD LIKE TO KNOW WHEN THAT VALUE IS LIKELY TO BE EXCEEDED

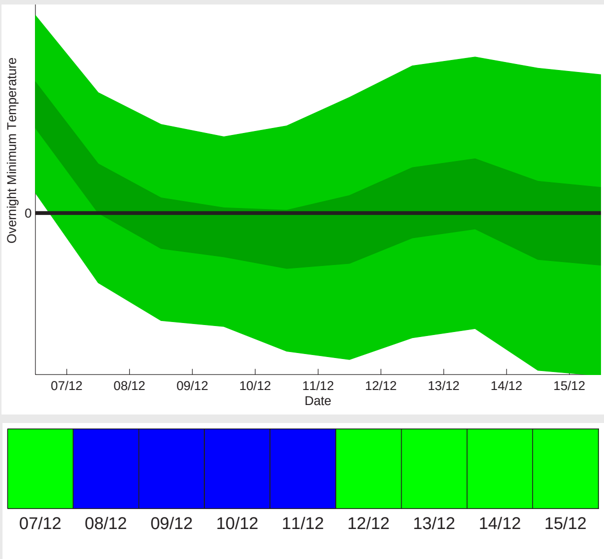
Best guess forecasts, taken at face value, give definitive statements about the expected weather. This confidence does not often withstand scrutiny. The alternative is to issue warnings along with an estimate of the confidence in that warning. We avoid providing a "probability" of an event occurring since we know that fundamental errors in the forecasting system prevent us from giving a reliable probability forecast. Instead warnings are issued with a targeted level of confidence, such as "we expect 3-out-of-4 warnings to be correct".

Issuing a warning every day would trivially capture every event but would not provide useful decision support. A skilful warning system will capture a good proportion of events and confidence will be justified by realised performance. Different warning systems can be developed to suit different needs. Warnings can be issued so that events are rarely missed, or warnings can be issued so that they coincide with an event more frequently. The level of confidence that can be achieved and the proportion of events that can be warned on will then depend on the skill of the ensemble forecast, the relative frequency of the event of interest and the size of the historical archive.

By comparing the expected performance with observed performance we can test the warning system. By making sure that no data used to develop the system is used for the testing, the results better reflect operational performance. We check whether or not performance is statistically consistent with the expected behaviour along with the proportion of events successfully warned on by the system.

OVERNIGHT FREEZING CONDITIONS

THE FOLLOWING GIVES AN EXAMPLE OF HOW ROBUST WARNINGS ON A PARTICULAR WEATHER EVENT MIGHT BE PRESENTED



Consider an organisation that is concerned with overnight freezing conditions and in particular the minimum overnight air temperature. Given regular air temperature observations, recorded every hour for example, we can produce a data set of overnight (between 8pm and 6am say) minimum temperatures.

The available ensemble forecasts correspond to surface ("2 meter model") temperatures at midnight. While this forecast information is relevant to overnight minimum temperatures there is no direct correspondence between the two quantities. In order to be useful the forecast must be interpreted using historical observations. By looking at the historical relationship between observed minimum overnight temperatures and midnight forecast temperatures we can build a statistical relationship between the two – translating forecast information into the quantity of interest.

The figure above shows a probabilistic minimum overnight temperature forecast going out 9 days. The light green range denotes the range of values that we expect roughly 9-in-10 observations to fall. The darker band denotes the range of values that we expect to see roughly 1-in-3 observations fall. As expected, the uncertainty, i.e. the range of likely outcomes, increases with the forecast lead-time.

This probabilistic forecast forms the basis for the warning system. Suppose we want to issue a warning when it is likely that the minimum overnight temperature will be less than 0 degrees Celsius. We can compute the chance of the event, according to the forecast, and then issue a warning when this chance exceeds some threshold. The thresholds determine how sensitive the warning system is and are determined from past data so that historical warnings are correct a certain proportion of the time.

Suppose, for example, that we want 3 warnings out of every 4 to be correct, on average. For each lead-time we find the smallest threshold for which 3 out of 4 warnings would be correct over the historical data and use this value for the warning system. The coloured squares underneath the plot provide information on likely overnight freezing conditions. The blue squares indicate likely freezing conditions, while the green squares indicate that the forecast has no information regarding freezing conditions. There are limits to the reliability that can be achieved, however. It may not be possible to identify thresholds that are able to provide warnings that are very confident, i.e. being correct 99 times out of 100 warnings. The level of confidence that can be achieved will vary from problem to problem, depending on the background frequency of the event, the amount of historical data available to determine the thresholds and the skill of the forecast system. It is also important that the thresholds are robust. Changing the data used to determine the thresholds should not change the threshold values unduly.

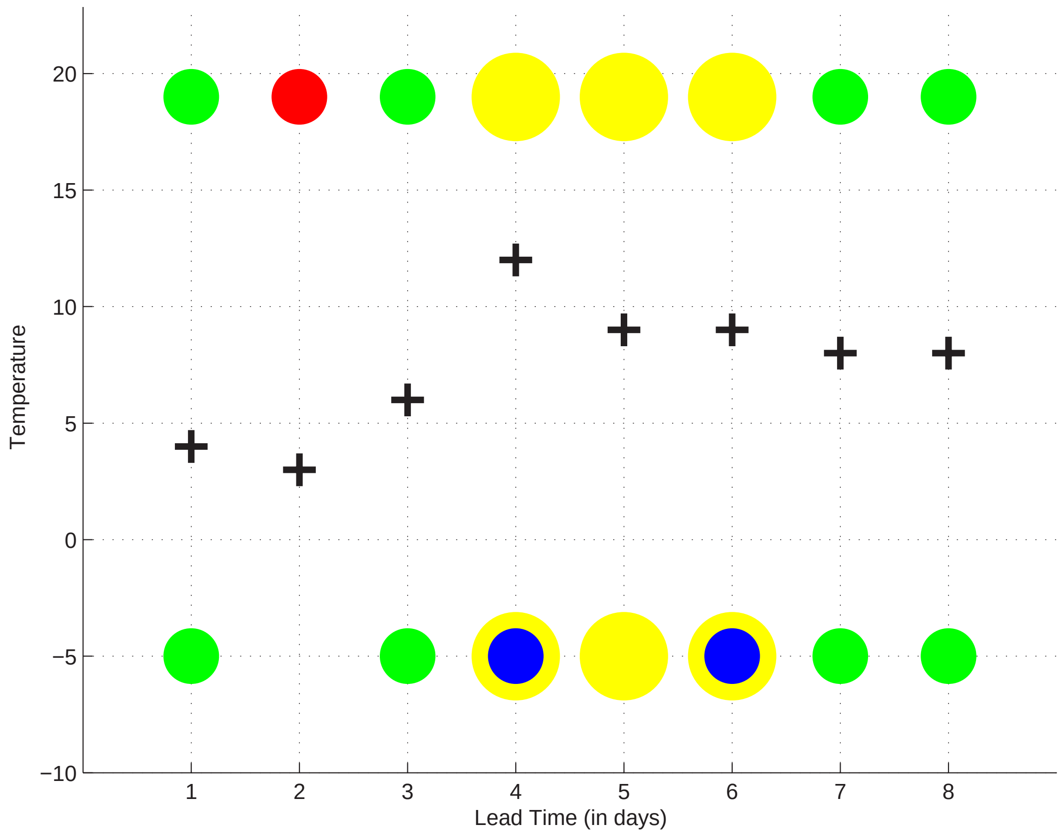
FORECASTING FORECAST ERROR

THE COST OF WEATHER RISK IS RARELY SYMMETRIC WITH ORGANISATIONS MORE OFTEN EXPOSED TO ONE SIDE OF FORECAST ERROR THAN ANOTHER

In these cases it is useful to know both the magnitude and sign of any likely error in a weather forecast. Best guess weather forecasts are rarely accompanied by an estimate of the likely error in that forecast. We can, however, use ensemble forecast information to inform on the sign and size of likely errors in a best guess forecast.

The forecast error warning system uses the ensemble forecast to estimate the chance that the error in the best guess forecast will exceed some pre-determined value. When the chance of an error is sufficiently large, a warning is issued. As before, the warning thresholds are determined using historical data and are tuned to achieve a targeted level of confidence. The level of confidence that can be achieved will depend on the frequency of the event, the quality of the best guess forecast, the quality of the ensemble forecast and the size of the historical archive.

The error information is presented with the best guess forecast (shown as black crosses), shown in the figure below. The dots at the top and bottom of the plot indicate likely forecast errors greater than 2 degrees Celsius, with the colour denoting the sign of the error. Blue dots indicate that the observed temperature will likely be 2 degrees colder than the forecast temperature, the red dots indicate that the observed temperature will likely be 2 degrees warmer than the forecast temperature. The yellow dots provide information on the days where there is increased uncertainty in the ensemble. These are days when the ensemble is wider than usual and the best guess forecast information can be used cautiously. The green dots indicate when there is no information regarding likely errors.



SUMMARY

The effective dissemination of weather forecasts turns on the clear communication of robust and relevant information. Traditional best guess forecasts, lacking accompanying estimates of the uncertainty in their predictions, appear too good to be true, and often prove to be just that. Useful estimates of forecast confidence can be provided by ensemble forecasting techniques. These probabilistic forecasts are, however, largely under-utilised in operational decision making. This, in part, is because many organisations are unable to process ensemble forecast information in its current form.

Presenting ensemble forecast information in the form of user-relevant warnings is a natural extension to traditional weather forecast information. Not every warning is expected to be correct however, and the uncertainty information is communicated by associating a level of confidence with the warnings. The resulting forecast information is made more robust, since, by design, warnings should verify as frequently as expected. By combining ensemble information with user-specific data, either site-specific meteorological observations or business relevant records such as road maintenance activity or retail sales, forecast information can be made more relevant.

The authors would like to thank all those people and organisations that have provided the data for this project and in particular Melvin Brown at the Smith Institute for Industrial Mathematics and System Engineering.

Centre for the Analysis of Time Series
London School of Economics
Columbia House
Dept of Statistics
Houghton Street
London
WC2A 2AE