STEPS: an empirical treatment of forecast uncertainty

A.W. Seed
BMRC, Melbourne, Australia

Abstract

Forecasts of rainfall in the very short term, say less than 2 hours, are usually based on advecting a radar rainfall field forwards into time. The Bureau is developing a flash flood warning service that will require rainfall forecasts at high spatial and temporal resolution. It will be useful therefore to understand the magnitude and sources of forecast errors in advection forecasts and to consider how this uncertainty can be conveyed to the users.

Intoduction

The Short Term Ensemble Prediction System (STEPS) has been developed as a joint project between the Bureau and the Met Office (UK) and has been configured so as to enable one to generate ensembles that are based on a single radar only for lead times out to 120 minutes, or a blend of radar mosaic and NWP rainfall forecasts for lead times out to 6 hours. The latter option is currently undergoing operational trials in the UK where it is being used to merge the radar-based nowcasts with the Met Office high-resolution rainfall forecasts with a 15-minute update cycle over a 1000 km x 1000 km domain. The Bureau is about to trial the use of nowcast ensembles in the context of a flash flood warning system, but will not be able to test the merging aspect of STEPS until work quantitative rainfall estimates from a radar network mosaic has been completed.

Sources of uncertainty

Uncertainty in extrapolation rainfall forecasts arises from errors in the initial radar rainfall analysis, errors in the estimation of the motion of the field, and temporal development of the rainfall field during the forecast period.

Radar estimation error

Data from the Kurnell Radar and a 200-gauge network of rain gauges that are operated by Sydney Water Corporation were used to evaluate the relative importance of the initial errors in the radar rainfall estimates. Significant rainfall was observed in Sydney during the period 12-17 May 2003. Ordinary block Kriging was used to estimate the 15-minute rainfall accumulations at 1 km resolution in the area covered by the rain gauge network, and the same gauge data were used by Rainfields to produce maps of 5-minute radar rainfall. STEPS was used to generate forecasts of 15-minute rainfall accumulations out to 90 minutes. . Figure 1 shows the mean squared error between the gauge and forecast fields over the area covered by the gauge network and Figure 2 shows the a time series of the mean standard error for a 24-hour period during the storm. It is clear from Figure 1 that errors in the initial radar rainfall estimation are significant over the 60-minute forecast period, accounting for about half the error variance, and Fig. 2 shows that these errors can vary significantly within the storm.

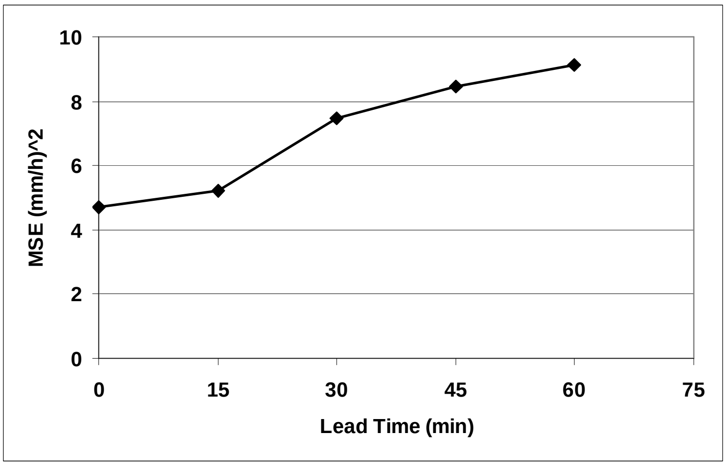

Figure 1. Mean square error $(mm/h)^2$ for 1 km, 15 minute rainfall accumulation forecasts as a function of lead time.

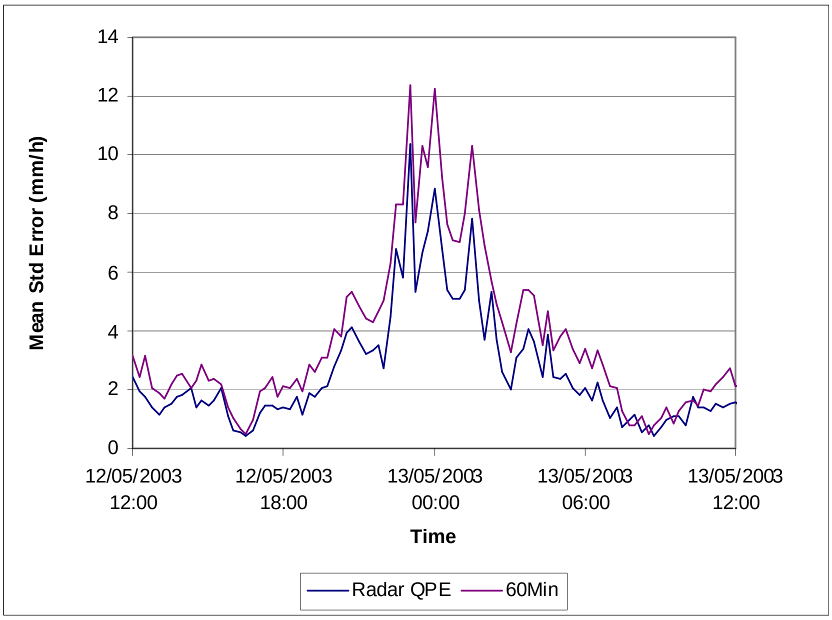

Figure 2. Mean standard error (mm/h) for the radar rainfall estimation and the 60-minute rainfall forecast for the period 12:00 12 May to 12:00 13 May 2003.

Tracking errors

The apparent velocity of each pixel in the field must be estimated, and these estimates are a potential source of error in the forecasts. The impact of tracking errors on the mean square error in a forecast was estimated by Bowler et al. (2003) by comparing forecasts generated using the advection that was valid at the start of the forecast period with forecasts that were generated using the advection that was observed during the forecast period. Bowler et. al. (2003) found that advection errors only accounted for about 10% of the variance of the forecast error, but these errors started to increase after about 3 hours into the forecast due to temporal development of the velocity field.

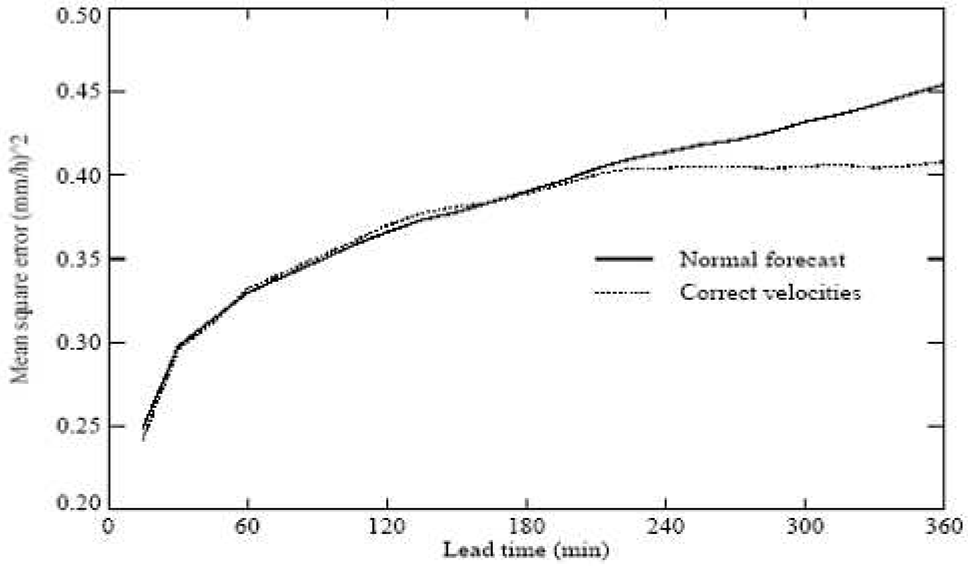

Figure 3. Mean square error of advection forecasts using the analysed advection fields (Correct velocities) and the advection field at the start of the forecast (Normal forecast). After Bowler et al 2003.

Temporal development

Continued development of the rain field during the forecast period is a major source of error in advection forecasting. The temporal development of a rain field is often conditioned on large-scale processes, for example the diurnal cycle, initiation of new convective storms in areas of significant topography and the like. Conceptual models have also been constructed that model the life cycle of particular physical phenomena, for example super-cell thunderstorms. Attempts have been made to include this knowledge into the forecast though identifying the object in the rainfall field and then imposing an appropriate life cycle to the object. The lifetime of an object is strongly linked to the scale of the object though a power law scaling relationship, and while this is a statistical description of the field it is a key feature that controls much of the observed scaling behaviour of rainfall.

This scaling relationship was studied by Turner et al. (2004) so as to quantify the predictability limits to precipitation forecasting as a function of the scale of a rainfall feature. Figure 4 shows the relationship between scale and lifetime for a number of cases. It is evident from this figure that the dynamic scaling relationship is dependent on the meteorological situation and therefore varies significantly between cases. The gain from using the meteorological intelligence based on typical life cycles or climatology to predict future initiation and decay of objects is quite small if the conceptual model used is inappropriate for the current situation.

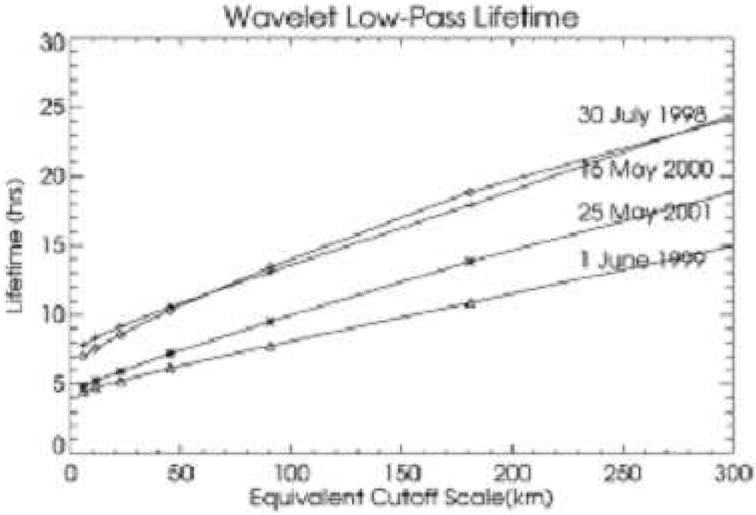

Figure 4. Lifetime as a function of scale. Reproduced from Turner et al. 2004.

Managing forecast uncertainty

It is useful to consider what can be achieved with imperfect forecasts since the perfect forecast is not a possibility. In general, one can look to reducing the uncertainty, improving understanding of the nature of the uncertainty, and informing users on the extent of the impact of uncertainty on their application.

Forecast uncertainty can be reduced though the development of better observation and forecast techniques. Radar observation errors can be reduced through better quality control of the radar reflectivity measurements and by using rain gauge data to reduce the mean bias. Extrapolation forecasts are based on a more-or-less instantaneous rain field rather than say a 30-minute accumulation so it is difficult to reduce the estimation errors in fields that are used for nowcasting. There is growing consensus that dual-polarimetric radar observations are most likely to provide the best data for quantitative rainfall nowcasting applications since they are unbiased and less sensitive to fluctuations in the Z-R relationship.

Understanding the statistical nature of observation and forecast uncertainties is a key step in moving towards being able to inform users on the uncertainty in the current observation and forecast. There is no ground truth in rainfall measurement since rain gauges are point observations of a spatial field and radar is an indirect measurement, and this complicates the analysis of observation and forecast errors.

Forecast uncertainty can be evaluated though the use of ensemble forecasts. Ensembles can be generated by perturbing the initial state of an NWP model or by collecting model forecasts from several sources. Ensembles that have been generated using a single model have lower resolution than single model runs, so it is not straightforward to convert these forecasts into something that can be used in hydrological applications given the sensitivity of run-off forecasts to high rain rates.

STEPS

The Short-Term Ensemble Prediction System (STEPS, see Bowler et al. 2003) has been developed jointly by the Met Office (Joint Centre for Hydro-Meteorological Research) and the Bureau of Meteorology (Bureau of Meteorology Research Centre) to provide an operational, stochastic precipitation nowcast capability. The project was initiated in response to a user requirement to quantify uncertainty in short-range precipitation forecasts. The resultant model produces forecasts in ensemble form. This allows probability distribution functions of fluvial flows to be estimated using existing deterministic, operational rainfall-run-off models, and can also facilitate objective decision making in pluvial and fluvial flood warning through the use of cost-loss models.

STEPS exploits a cascade representation of precipitation fields. This provides a suitable framework for modelling their fractal and dynamic scaling properties (see Venugopal et al.,1999). A field of instantaneous rain rate (estimated from radar or generated by a NWP model), with dimensions of L by L pixels, can be decomposed into a hierarchy of component fields representing variability on a discrete set of horizontal scales. Expressed in units of rain rate (e.g. mm h-1), the cascade is multiplicative and takes the form:

$$R_{i,j}(t) = \prod_{k=1}^{n} X_{k,i,j}(t) \quad \text{for =1,..,L, j=1,..,L, L=2n} \qquad (1)$$

where n is the number of component levels in the cascade, and the kth field in the cascade, Xk(t), represents the variability in the original field with frequencies, ωk, in the range $\frac{2^{k-1}}{L} < \omega_k < \frac{2^k}{L} pixel^{-1}$ at time, t. However, it is more convenient from a modelling perspective to work in units of 10+10log10 R because the log transformation reduces the skewness of the probability distribution making it closer to a Normal distribution and the addition of 10 reduces the truncation effect for very light rain rates ? Expressed in units of decibels of rain rate, dBR (10 log10 R), the cascade becomes additive:

$$dBR_{i,j}(t) = \sum_{k=1}^{n} X_{k,i,j}(t) \quad \text{for =1,..,L, j=1,..,L, L=2n} \qquad (2)$$

where n=8 in STEPS.

A field of surface rain rate expressed in dBR is transformed into the frequency domain by means of a Fast Fourier Transform (FFT). Each level, Xk(t), of the cascade is calculated using a band-pass filter based upon a Gaussian window to pass the appropriate frequencies. An inverse transform is performed to return each Fourier component of the original field back into the spatial domain. Finally, the cascade levels are normalised using

$$Y_{k,i,j}(t) = \frac{X_{k,i,j}(t) - \mu_k(t)}{\sigma_k(t)} \qquad (3)$$

where $\mu_k(t)$ and $\sigma_k(t)$ are the mean and standard deviation of the kth level respectively.

The $Y_{k,i,j}(t)$ are referred to as the component levels of the cascade.

In keeping with the nowcasting methodology employed in Nimrod (Golding, 1998) and Gandolf, STEPS blends an extrapolation forecast with a deterministic NWP forecast of precipitation. In Nimrod and Gandolf this blending is achieved in physical space and the weight given to the extrapolation component takes the form of a fixed exponential decay with time. In STEPS, the merging of the extrapolation and NWP component forecasts is performed in a scale-dependent way using each level of the cascade. This allows the scheme to capture the scale dependent loss of predictive skill in the extrapolation forecast with advancing lead time (Venugopal et al., 1999; Germann and Zawadzki, 2002) and the scale-dependence of the skill of the NWP forecasts. The weights assigned to the extrapolation and NWP components are computed on-line for each STEPS run. This ensures that optimum use is made of both extrapolation and NWP models as the meteorology of a precipitation event evolves and the skill of the models changes.

An ensemble of nowcasts is generated by blending the extrapolation and NWP component forecasts with noise whose spatio-temporal statistical properties are derived from those of recent radar based analyses of instantaneous rain rate. Uncertainties in the advection and Lagrangian evolution of the extrapolation nowcast are modelled. Also, some accountant is taken of errors in the NWP component, although, at present this is limited to: a bias correction applied to the NWP model precipitation fields; the merging of the advection and NWP diagnosed velocity fields makes some adjustment for displacement/timing errors in the NWP forecast ; the weight assigned to the NWP component will vary according its skill as measured against the latest rainfall analysis.

A single forecast realisation is a blend of three cascades: extrapolation, NWP and noise. A set of weights are calculated for each of these forecast components, for each cascade level. The weights applied to the extrapolation component are derived from a hierarchy of second order Auto-Regressive (AR-2) models; those ascribed to the NWP forecast are based upon its skill as measured by cross correlating the latest rain analysis with a time synchronous NWP forecast. The weights are formulated such that, for each level, they preserve the variance of the level rather than sum to one since the mean of each level has been set to zero .

This model formulation ensures that, in each ensemble member, extrapolated features are progressively replaced by noise from the smallest scales upwards. This process proceeds at a rate that is consistent with the diagnosed temporal persistent of features on each level in the cascade, but is arrested by the predictive skill of the NWP model.

STEPS can be used to generate an ensemble of nowcasts without blending into the NWP forecasts, and this option will be used to generate ensembles for 0-90 minute nowcasts over single radars for the major Australian cities during the 2005-2006 summer as a trial. The Met Office have commenced using STEPS to provide a single blended rainfall forecast out to 6 hours as an operational trial, and verification studies will be continued once sufficient data have been collected.

Conclusions

While the accuracy of NWP forecasts of rainfall have improved steadily, and are likely to continue to improve in the future, the forecasts contain major uncertainty, which needs to be taken into account when presenting forecasts to users. Ensemble forecasts are able to provide a measure of forecast uncertainty, but much more work is required to develop techniques to present forecast uncertainty to users. Techniques use uncertainty information in applications also requires considerable development and this should be done in parallel with developing ensemble forecast techniques.

References

Bowler, N.E., C. Pierce, and A.W. Seed, 2003. STEPS: A probabilistic forecast scheme which merges an extrapolation forecast with downscales NWP. *Forecasting Technical Report 433*, Met Office

Germann, U., and Zawadzki, I., 2002. Scale dependency of the predictability of precipitation from continental radar images. Part 1: description of the methodology. *Monthly Weather Review*, 130(12), 2859-2873.

Golding, B., 1998. Nimrod: a system for generating automatic very short range forecasts. *Meteorol. Appl*., 5, 1-16.

Marsan, D., Schertzer, D. & Lovejoy, S. (1996) Causal space–time multifractal processes: predictability and forecasting of rain fields. *J. Geophys. Res.* 101, 26333–26346.

Turner, B., I. Zawadzki, and U. Germann, 2004. Predictability of precipitation from continental radar images. Part 3: Operational nowcasting implementation (MAPLE). *J. Applied Meteorology*, 43, 231-247.

Venugopal, V. E., Foufoula-Georgiou and Sapozhnikov, V., 1999. Evidence of dynamic scaling in space-time rainfall. *J. Geophys. Res.*, 104 (D24), 31599-31610.

Zawadzki, I., Morneau, J., and Laprise, R., 1994. Predictability of precipitation patterns: an operational approach. *Journal of Applied Meteorology*, 33, 1562-1571.