The Nowcasting of Precipitation during Sydney 2000: An Appraisal of the QPF Algorithms


*Met Office, Joint Centre for Hydro-Meteorological Research, Wallingford, Oxfordshire, United Kingdom
+Bureau of Meteorology Research Centre, Melbourne, Victoria, Australia
#Telford Institute of Environmental Systems, University of Salford, Manchester, United Kingdom
@Department of Atmospheric Science, University of Missouri—Columbia, Columbia, Missouri
&Cloud Physics Research Division, Meteorological Service of Canada, Downsview, Ontario, Canada
**National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 8 August 2002, in final form 8 April 2003)

ABSTRACT

Statistical and case study–oriented comparisons of the quantitative precipitation nowcasting (QPN) schemes demonstrated during the first World Weather Research Programme (WWRP) Forecast Demonstration Project (FDP), held in Sydney, Australia, during 2000, served to confirm many of the earlier reported findings regarding QPN algorithm design and performance. With a few notable exceptions, nowcasting algorithms based upon the linear extrapolation of observed precipitation motion (Lagrangian persistence) were generally superior to more sophisticated, nonlinear nowcasting methods. Centroid trackers [Thunderstorm Identification, Tracking, Analysis and Nowcasting System (TITAN)] and pattern matching extrapolators using multiple vectors (Auto-nowcaster and Nimrod) were most reliable in convective scenarios. During widespread, stratiform rain events, the pattern matching extrapolators were superior to centroid trackers and wind advection techniques (Gandolf, Nimrod).

There is some limited case study and statistical evidence from the FDP to support the use of more sophisticated, nonlinear QPN algorithms. In a companion paper in this issue, Wilson et al. demonstrate the advantages of combining linear extrapolation with algorithms designed to predict convective initiation, growth, and decay in the Auto-nowcaster. Ebert et al. show that the application of a nonlinear scheme [Spectral Prognosis (S-PROG)] designed to smooth precipitation features at a rate consistent with their observed temporal persistence tends to produce a nowcast that is superior to Lagrangian persistence in terms of rms error. However, the value of this approach in severe weather forecasting is called into question due to the rapid smoothing of high-intensity precipitation features.

1. Introduction

As previously described in the literature (Fox et al. 2001), and elsewhere in this special issue, the first World Weather Research Programme (WWRP) Forecast Demonstration Project (FDP), titled Sydney 2000, was held in Sydney, Australia, between 4 September and 21 November 2000. Although Sydney 2000 placed an emphasis on the demonstration and appraisal of predictive algorithms for the nowcasting of high-impact, convective weather (Wilson et al. 2004, in this issue), as recommended in the WWRP mission statement (WMO 1998), the majority of the participating nowcast systems possessed short-range precipitation forecasting capability, and indeed, some were developed specifically for this purpose. In this paper, the aim is to examine the design and performance of the weather radar-based quantitative precipitation nowcast (QPN) algorithms run by these systems.

Of the six algorithm participants in Sydney 2000, four generated precipitation nowcast products routinely. These were the Auto-nowcaster [ANC; National Center for Atmospheric Research (NCAR), Boulder, Colorado], the Spectral Prognosis (S-PROG) model [Bureau of Meteorology Research Centre (BMRC), Melbourne, Australia], Nowcasting and Initialisation of Modelling Using Regional Observation Data System (Nimrod; Bracknell, Berkshire, United Kingdom), and the Generating Advanced Nowcasts for Deployment in Operational Land-surface Flood Forecasts System (Gandolf; an experimental version of the Met Office’s convective precipitation nowcasting system, developed jointly by the Met Office and the University of Salford, Salford, United Kingdom). In addition to these contributions, the regional forecast office of the Bureau of Meteorology (BoM) in Sydney ran a version of the Thunderstorm
Identification, Tracking, Analysis and Nowcasting (TITAN) scheme, components of which are also implemented in the ANC. The two remaining participating systems possessed limited capability to produce rainfall-related products [the Canadian Radar Decision Support system (CARDS), Meteorological Service of Canada, Downsview, Ontario, Canada; and the Warning Decision Support System (WDSS), National Severe Storms Laboratory, Norman, Oklahoma] but were not designed specifically for precipitation forecasting, and are not therefore considered here.

2. The observation network

The radar coverage afforded by the preexisting weather radar network (prior to the FDP) was sufficient to allow almost continuous radar observation of precipitation over a 705 km by 750 km domain, broadly centered on Sydney (the Nimrod operational domain). The radars in the network are a mix of conventional and Doppler C and S bands, with beamwidths between 1.5° and 3° and data ranges of 150 km (for radars operating in volumetric mode) or 512 km (for radars operating in long-range surveillance mode with fewer tilts). The BoM routinely generates a 2-km-resolution reflectivity composite (constant-altitude plan position indicator, CAPPI). Radar corrections applied prior to dissemination include the removal of ground clutter by a Doppler clutter filter, and corrections for range (1/r²), beam occultation, and attenuation. Since the radar data are used primarily for the early detection of severe convective storms, and their associated severe weather signatures, the BoM has yet to implement any automated, operational procedures for estimating surface rain rate by vertical profile correction or gauge adjustment.

Several enhancements were made to the observational network in the vicinity of the Sydney basin for the duration of the FDP. The installation of a C-band polarimetric (C-Pol) radar at Badgery’s Creek (approximately 40 km west of Sydney) provided improved “clear air” radar coverage for detection of convergence zones, and polarimetric monitoring for precipitation typing. Volumetric reflectivity and Doppler velocity data were available from both the C-Pol and Kurnell (20 km south of Sydney) radars. May et al. (2004, in this issue) discuss the technical aspects of the C-Pol and Kurnell radar operations during Sydney 2000. The challenges faced included the identification of sites to optimize radar coverage and data quality, and the operational provision of high-resolution reflectivity and Doppler velocity data from the same radar. Kurnell proved more prone to the effects of ground clutter than C-Pol, and only some of this could be removed effectively by Doppler clutter filtering. Clutter was particularly prominent in the lowest plan position indicators (PPIs) over the sea.

3. A description of the algorithms

Objective approaches to QPN have been reviewed on numerous occasions, for example, in Browning and Collier (1989), Joss and Waldvogel (1990), Collier (1994), and Wilson et al. (1998). The aim here is to present a concise description of each of the algorithms demonstrated during Sydney 2000, and in so doing, place their formulation within the wider context of historical developments in quantitative precipitation forecasting (QPF) research.

Since the earliest work on the derivation of precipitation motion from temporal sequences of radar echoes (Hilst and Russo 1960; Wilson 1966), three broad approaches to precipitation nowcasting by radar echo extrapolation have evolved.

1) Pattern matching approaches: These attempt to establish the motion of an entire field or its component rain areas by some form of cross correlation (Zawadzki 1973; Austin and Bellon 1974; Rinehart and Garvey 1978; Rinehart 1981; Bellon and Austin 1984; Li et al. 1995).

2) Centroid tracking approaches: These use a variety of objective methods for feature identification, for example, contouring (Duda and Blackmer 1972), clustering (Haralick and Kelly 1969; Barclay and Wilk 1970), thresholding (Witt and Johnson 1993), or spectral methods (Zittel 1976). A temporal sequence of observed centroid positions is then used to derive a motion vector for each feature (Zittel 1976; Bjerkas and Forsyth 1979; Barclay and Wilk 1970; Austin and Bellon 1982; Rosenfeld 1987; Witt and Johnson 1993). Some of the resultant extrapolation algorithms have been designed to handle the merging and splitting of precipitating showers and thunderstorms (Dixon and Wiener 1993).

3) Wind advection approaches: These exploit independent observations or forecasts of wind velocity (Parsons and Hobbs 1983; Hand 1996; Golding 1998).

In the 1970s, Elvander (1976) concluded that linear extrapolation by cross correlation was the most reliable means of obtaining motion vectors for a wide variety of precipitation features seen in radar imagery. However, in cases of convective precipitation, when individual precipitation features tended to be relatively small and well defined, he noted that extrapolation by a linear least squares fit to past centroid positions was effective. Several decades later, these conclusions explain the basic differences in QPF algorithm design between some of the North American (TITAN, WDSS) and U.K. (Nimrod, Gandolf) nowcasting systems.

In the United States and Australia, nowcasting has been concerned primarily with the prediction of convection and its associated severe weather (damaging winds, hail, lightning). Consequently, efficient centroid tracking algorithms have evolved to extrapolate thunderstorms using volumetric radar reflectivity data. In
these countries, the prediction of surface precipitation is secondary to that of severe weather, and thus relatively little research and development effort has been expended in this area. By contrast, the focus of nowcasting research in the United Kingdom has been on improving predictions of rain rate and rain accumulation, since the most important meteorological hazard there is widespread, persistent rain, resulting in fluvial flooding. Here, the optimum methods for nowcasting such rain have proved to be extrapolation, either by cross-correlation vectors or NWP model winds (Ryall and Conway 1994; Ryall 1995).

Table 1 summarizes some of the essential attributes of the nowcasting systems demonstrated during Sydney 2000, including their QPN schemes. Each of these schemes is described in more detail in the following subsections.

### a. TITAN, Bureau of Meteorology

TITAN utilizes the centroid tracking algorithm developed by Dixon and Wiener (1993). Such algorithms are well suited to the extrapolation of small, well-defined radar echoes associated with updraft convection, but, in principle, can be applied to more extensive precipitation echoes. In common with similar algorithms proposed by Rosenfeld (1987) and Witt and Johnson (1993), this scheme defines “storms” as reflectivity volumes delimited by a predefined reflectivity threshold. A selection of storm properties is computed from multiple-elevation, Cartesian-gridded PPI scans. These include reflectivity-weighted storm centroid, storm volume, and the size and shape of the storm’s 2D footprint. The latter is described in terms of a set of ellipsoid attributes: ellipsoid centroid, ellipsoid major and minor axis radii, and ellipsoid major axis orientation.

TITAN storm tracks are computed by matching storms in successive radar scans using a combinatorial optimization technique. This involves identifying the most probable track for a given storm from a set of possible tracks, generated by pairing the storm with each of its neighbors. The “true” storm track is found by minimizing a cost function that expresses the quality of the match in terms of storm separation distance and volume. The TITAN tracking algorithm is designed to handle scenarios involving the merging and splitting of storms. Thus, when two storms merge, the scheme will extend only one of the two original tracks. In the case of a splitting storm, only one track can be extended, and a new track is created for the unmatched storm (Dixon and Wiener 1993).

The advection algorithm in TITAN assumes that storms move in straight lines. When storms are first identified, a Lagrangian persistence extrapolation is applied. In subsequent runs, forecasts may be generated using linear or nonlinear trend models in which the weight given to past storms tracks decays exponentially with time. During Sydney 2000, this feature of TITAN was not used.

The scheme was applied to PPI scans from the Kurnell and C-Pol radars. An agreed upon, single $Z-R$ relationship was used to estimate instantaneous rain rate from TITAN-thresholded reflectivity. In common with the other extrapolation-based nowcast schemes, a vertical profile correction was not applied to estimate surface rain rate.
b. Auto-nowcaster, National Center for Atmospheric Research

The Auto-nowcaster (ANC) developed by NCAR is an expert system for the nowcasting (0–60 min) of thunderstorm evolution. It incorporates a data fusion system designed to assimilate a variety of datasets including those from radar, wind profiler, satellite, mesonet, and upper-air soundings. Conceptual models of thunderstorm initiation, growth, and decay are employed to produce nowcasts. These are embodied within a series of mathematical functions (predictors) whose application depends upon a fuzzy logic approach. Although these functions are physically based, performance statistics are employed to tune them.

Some of the more important predictors of convection relate to boundary layer convergence. While automated techniques have been developed to detect convergence lines within the ANC, these were judged to be inadequate for the FDP. Consequently, human-identified boundary layer convergence lines were entered manually. A single-Doppler radar wind retrieval technique (Sun and Crook 1994, 2001) was used to diagnose the vertical motion associated with convergence lines. A detailed discussion of this and other aspects of the ANC can be found in Mueller et al. (2003).

The ANC extrapolates radar reflectivity using a modified pattern matching technique based upon the Tracking Radar Echoes by Correlation (TREC; Tuttle and Foote 1990) algorithm. The extrapolated reflectivity field is then modified by the ANC to account for predicted initiation, growth, and decay. Here, information on extrapolated thunderstorm characteristics obtained from TITAN (Dixon and Wiener 1993) is employed. Estimates of rain rate are derived from the predicted reflectivity using a single Z–R relationship. In common with the other schemes on trial, no corrections were made to estimate rain rate or accumulation at the earth’s surface. Examples of a precipitation analysis and nowcast from the ANC are shown in Fig. 1 (top left and bottom left, respectively).

c. Nimrod, Met Office

The Nimrod system developed by the Met Office during the early 1990s is a fully automated expert system designed to produce near-surface nowcasts (0–6 h) of a variety of meteorological variables, including precipitation (Fig. 1, bottom right). The system design is similar in philosophy to that of the ANC, the aim being to exploit the benefits of a range of observational data and nowcasting algorithms. Ryall (1995) examined a wide range of extrapolation techniques for application to QPN in the United Kingdom. She concluded that the combination of a cross-correlation-based linear extrapolation technique similar to TREC and a numerical weather prediction (NWP) model wind field advection method provided the optimum means of generating radar-based short-range precipitation forecasts in the United Kingdom. In Nimrod, these two advection algorithms may be applied interchangeably to individual contiguous rain areas, referred to as rain objects. The choice of advection method for each object can vary from forecast to forecast depending on the relative performance of hindcasts generated by the two approaches.

One significant distinction between Nimrod and the other participating systems relates to the estimation of surface rain rate and rain accumulation. Unlike ANC, S-PROG, and Gandolf, Nimrod normally applies a sophisticated vertical profile correction to radar reflectivity to account for the effects of beam occultation, range (1/\(r^2\)), attenuation by precipitation, bright band, and droplet growth below the lowest radar beam (Kitchen et al. 1994). This correction is applied to Cartesian-gridded reflectivity data on a pixel-by-pixel basis, using a 1D physical parameterization of the vertical reflectivity profile. Following this, a single Z–R relationship is utilized to convert the derived, near-surface reflectivity estimates to an equivalent instantaneous rain rate. Finally, a mean field bias adjustment is made to the rain-rate field by comparing time integrations of the derived surface rain-rate fields with rain gauge accumulations (Gibson 2000).

Due to time limitations, the Nimrod vertical profile and gauge adjustment schemes were not implemented in Sydney. Furthermore, data from the BoM’s weather radar network were subjected to only limited quality control (May et al. 2004). As a result, the quality of the rain-rate analyses and forecasts generated by Nimrod were not on a par with those normally produced in the United Kingdom.

The version of Nimrod run in Sydney included a rain accumulation scheme capable of generating 1–6-h rainfall totals. This scheme utilizes an algorithm that computes surface accumulations by considering the dwell time of rain objects over the ground.

d. Gandolf, Met Office and University of Salford

In the mid-1990s, the Met Office undertook research and development aimed at improving quantitative nowcasts of convective precipitation. They explored the conceptual modeling of convection using object-oriented programming techniques. Hand and Conway (1995) proposed a conceptual life cycle model of a convective cell, capable of ingesting high-resolution radar and satellite data, and forecasts from a mesoscale NWP model. In the scheme, five developmental stages in the life cycle of a typical, midlatitude, convective cell are recognized and distinguished on the bases of their vertical reflectivity profiles, size, and depth. A forecast algorithm allows each observed cell and its associated attributes to proceed through an evolutionary cycle determined by the conceptual life cycle model, while the cell is advected by a representative NWP model wind vector. A key component of the scheme is an algorithm designed to simulate severe storm propagation by daughter cell
FIG. 1. Selected QPF products for 7 Sep 2000. From left to right and top to bottom: an ANC rain-rate analysis valid at 0128 UTC; S-PROG and ANC $T + 30$ min rain-rate forecasts valid at 0135 and 0133 UTC, respectively; and a Nimrod $T + 90$ min rain-rate forecast valid at 0130 UTC.

In collaboration with the University of Salford, the Met Office demonstrated an experimental version of the Gandolf system during Sydney 2000. This incorporated conceptual life cycle models of single-cell, multicell (daughter cell propagation), and supercell thunderstorms. The bases for this three-way cell classification were various convection diagnostics computed from mesoscale NWP forecasts (Pierce and Cooper 2000). These included convective available potential energy (CAPE), convective inhibition energy (CIN), and vertical shear in the horizontal wind. Each cell, identified from multibeam radar reflectivity and a Geostationary Meteorological Satellite-5 (GMS-5) infrared cloud-top temperature field, was assigned various attributes. These included an estimate of cell mean surface rain rate, calculated from the lowest available Kurnell radar PPI...
scan. Echoes observed only in this scan and not in those at higher elevations were assumed to be ground clutter, and were ignored. The average, cell-oriented evolutionary cycle of surface rain rate, estimated from observed cells at different developmental stages, was utilized to forecast the growth and decay of rain rates associated with each cell. In common with Nimrod’s approach to the estimation of surface rain accumulation, OOM rain accumulation forecasts were produced by considering cell advection velocity and the consequent dwell time of each cell over the ground in Gandolf’s operational domain.

e. S-PROG, Bureau of Meteorology Research Centre

Implicit in the design of S-PROG, developed by the BMRC, is the assumption that radar-derived rainfall fields exhibit multifractal properties and can be modeled using a bounded multiplicative cascade (Seed 2003). This approach assumes that a rainfall field is the product of a hierarchy of correlated random fields for which the correlation length of the field at a given level in the hierarchy is equivalent to the scale represented at that level. The variance of the rainfall field decreases as a power law with increasing scale. Instantaneous rain rate and radar reflectivity are related through a power law. Thus, a multiplicative cascade of instantaneous rain rate can be transformed into an additive cascade of reflectivity.

In S-PROG the cascade of random fields is calculated from a radar reflectivity field using a notch filter in the frequency domain (Seed 2003; see also Bellon and Zawadzki 1994). Nowcasts are generated by modeling the temporal development of reflectivity at each level in the cascade using a second-order autoregressive process [AR(2)]. The Lagrangian autocorrelations at lags of one and two radar scan cycles (5 and 10 min, respectively during Sydney 2000) are computed for every model run, by comparing the current radar reflectivity with that observed in the previous two scans. The parameters used to describe the hierarchy of AR(2) models are estimated from these autocorrelations using the Yule–Walker equations (Seed 2003). The evolution of the forecast reflectivity field is constrained such that the resultant, derived rain-rate fields exhibit stationarity and the mean instantaneous rain rate and rain area do not change over time. A single Z–R relationship is used to convert the forecast reflectivity fields into instantaneous rain rate. Figure 1 (top right) shows an example of an S-PROG forecast from 7 September 2000.

4. Performance evaluation of the QPF algorithms

a. Overview

Between 4 September and 21 November 2000, there were 20 days when some precipitation fell in Sydney and its environs. On 12 of these days the precipitation was predominantly convective with observed radar reflectivities in excess of 40 dBZ. On the remaining 8 days, the rain was generally stratiform in character. The following appraisal of the QPF algorithms draws upon selected model runs and performance statistics generated for two of these events: severe convective precipitation associated with supercell development [3 November 2000; see Wilson et al. (2004, in this issue) and Sills et al. (2001)], and orographically forced precipitation in an easterly flow regime (12 November 2000). The selected statistics used to demonstrate comparative algorithm performance are described by Ebert et al. (2004, in this issue) elsewhere in this special issue.

b. Validation methodology

For the duration of the FDP, QPF products generated by each of the nowcasting systems were archived routinely. These included TITAN storm tracks (BoM), and rain analyses and forecasts generated by the ANC, Gandolf, Nimrod, and S-PROG. Since these systems did not share a common domain, projection, resolution, or data feed (see Table 1), it was proposed that forecast products be evaluated against their associated rain analyses. This would avoid penalizing system performance for reasons other than those directly attributable to algorithm design. However, it was agreed that all systems should be assessed over a common domain, namely, the Severe Weather Warning Area recognized by the BoM’s regional forecast office in Sydney (see, e.g., Fig. 1).

The verification statistics produced fall into two broad categories: those computed using entire forecast fields (continuous, categorical, and event-oriented performance measures) and those that are based upon contiguous rain areas (CRAs; Ebert and McBride 2000). In the following issue-based examination of algorithm performance, reference is made to a small selection of these statistics, full descriptions of which can be found in Ebert et al. (2004) and Ebert and McBride (1998). Since a direct statistical comparison of algorithm performance is not recommended for the reasons given above, Eulerian and Lagrangian persistence forecasts generated from the rain analyses of each nowcast system provide common performance benchmarks against which algorithm performance can be compared.

Associated with the performance statistics for each QPF algorithm in Table 2 is an equivalent persistence forecast based upon the relevant rain-rate analysis, and an extrapolation (Lagrangian persistence) forecast derived using a single, field-wide, cross-correlation vector.

c. Issue-oriented case studies

Here the aim is to highlight some of the key performance issues brought to light during the FDP, while avoiding, except where justified, an explicit, statistical comparison of algorithm performance. The issues addressed in the following subsections include a review...
of the impacts of data quality on algorithm performance (issue 1); the relative merits of pattern matching, centroid tracking, and NWP model wind advection approaches to extrapolation forecasting (issue 2); and the impact of algorithm complexity on performance (issue 3). Where relevant, and to avoid duplication, cross references are made to additional material in other papers in this special issue.

1) ISSUE 1: THE IMPACT OF DATA AND PRODUCT RESOLUTION ON ALGORITHM PERFORMANCE

In recent years substantial theoretical and empirical evidence has been produced to suggest that precipitation fields exhibit spatiotemporal properties that are consistent with their representation by multifractal models (Foufoula-Georgiou 1996; Menabde et al. 1999). These models assume that similar features exist in precipitation fields on a continuum of scales from the large (10^4 m) to the very small (<10 m). Observations show that the temporal persistence of such features is broadly proportional to their size. Consequently, the temporal range over which the Lagrangian persistence assumption is valid for the purposes of deterministic precipitation forecasting will be a function of spatial resolution (Germann and Zawadzki 2002).

It follows that the measured statistical performance of a forecast in terms of rain–no-rain discrimination can be improved, simply by reducing its spatial resolution (i.e., by smoothing the forecast). However, such an improvement would be at the expense of losing potentially valuable, quantitative information on the finescale precipitation structure, and it is this latter content that is likely to be most important for short-range, severe weather prediction. Clearly then, the drawing of inferences regarding forecast performance from a limited objective, statistical verification may lead to an erroneous perception of product quality. This is illustrated below.

During the afternoon of 3 November 2000, several thunderstorms were observed to develop in the vicinity of Campbelltown to the southwest of Sydney (Figs. 2a–c). Their subsequent motion and evolution were highly nonlinear as a result of interactions between storm near-surface outflows and a sea-breeze front (Wilson et al. 2004; Sills et al. 2001). Given the convective nature of the event, it was believed that the centroid tracking extrapolators of TITAN would provide the best quantitative forecast guidance. Subsequently, this was largely confirmed by case study appraisals and consultation with human forecasters (Wilson et al. 2004).

Nonetheless, when considered alone, selected categorical performance statistics for the event may be construed as indicating otherwise. The statistics shown in Table 2 suggest that Nimrod’s 5-km-resolution rain-rate forecasts were superior, in terms of rain–no-rain discrimination, to the 2-km-resolution products of the ANC and TITAN. Thus, the event-average equitable threat score (ETS) for T + 30 min Nimrod forecasts was 0.307 compared with equivalent values of 0.158 for the ANC and 0.194 for TITAN. However, the temptation to conclude that the Nimrod product was more useful to the human forecaster is dispelled when outputs from each model are compared side by side. Figure 2a clearly highlights the loss of information on storm volume and intensity in the coarser-resolution Nimrod product. This feature of Nimrod performance is confirmed by the more comprehensive statistical evaluation undertaken by Ebert et al. (2004).

The above discussion emphasizes the important point made earlier regarding the relationship between data resolution and the useful range of extrapolation forecasts. Although a 2-km-resolution extrapolation forecast is likely to provide better quantitative guidance than an

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Categorical statistics</th>
<th>CRA statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ETS</td>
<td>HK</td>
</tr>
<tr>
<td>ANC</td>
<td>0.158</td>
<td>0.304</td>
</tr>
<tr>
<td>Persistence</td>
<td>0.061</td>
<td>0.116</td>
</tr>
<tr>
<td>Lagrangian persistence</td>
<td>0.149</td>
<td>0.261</td>
</tr>
<tr>
<td>Nimrod</td>
<td>0.307</td>
<td>0.515</td>
</tr>
<tr>
<td>Persistence</td>
<td>0.186</td>
<td>0.312</td>
</tr>
<tr>
<td>Lagrangian persistence</td>
<td>0.308</td>
<td>0.470</td>
</tr>
<tr>
<td>Gandolf</td>
<td>0.097</td>
<td>0.289</td>
</tr>
<tr>
<td>Persistence</td>
<td>0.083</td>
<td>0.153</td>
</tr>
<tr>
<td>Lagrangian persistence</td>
<td>0.194</td>
<td>0.327</td>
</tr>
<tr>
<td>S-PROG</td>
<td>0.189</td>
<td>0.309</td>
</tr>
<tr>
<td>Persistence</td>
<td>0.075</td>
<td>0.138</td>
</tr>
<tr>
<td>Lagrangian persistence</td>
<td>0.169</td>
<td>0.287</td>
</tr>
<tr>
<td>TITAN</td>
<td>0.194</td>
<td>0.419</td>
</tr>
<tr>
<td>Persistence</td>
<td>0.079</td>
<td>0.151</td>
</tr>
<tr>
<td>Lagrangian persistence</td>
<td>0.154</td>
<td>0.276</td>
</tr>
</tbody>
</table>
Fig. 2. (a) Selected precipitation analysis products for the afternoon of 3 Nov 2000. From left to right and top to bottom: a C-Pol precipitation-type analysis valid at 0200 UTC, a Kurnell PPI reflectivity image for 0210 UTC, an ANC rain-rate analysis for 0217 UTC, Nimrod and Gandolf rain-rate analyses for 0200 UTC, and an S-PROG rain-rate analysis for 0205 UTC. (b) Selected rain-rate analyses generated by Gandolf on 3 Nov 2000. The first two analyses in the sequence, valid at 0420 and 0430 UTC, show an intense supercell moving northeastward. The analysis on the right, valid at 0520 UTC, shows the existing supercell has tracked northward to the west of the city center, while a new severe storm has developed on its northeast flank just to the north of Sydney.
equivalent 5-km product in the very short range (up to 1 h ahead), it will tend to lose value more rapidly. Thus, if one is not to mislead the end user with spurious, and therefore unjustified, detail, a deterministic forecast must reflect the growth in uncertainty at the smaller spatial scales, by degrading the product resolution with lead time. This approach has been adopted in the formulation of S-PROG (Seed 2003) and Nimrod (Ryall 1995; Golding 1998).

In the S-PROG scheme a hierarchy of second-order AR(2) models are used to establish the temporal persistence of precipitation features on a range of spatial scales. This information is used to smooth the precipitation field accordingly. Since the smallest-scale pre-
cipitation features are generally the most ephemeral, this smoothing has the effect of averaging the advected precipitation field over progressively larger areas as lead time advances (i.e., it is smoothed toward the domain-mean rain rate). This is illustrated in Fig. 1 (top right). Statistical validation confirms the intuition that this approach minimizes the rms error in forecast rain rate (Ebert et al. 2004; Seed 2003).

The response of users to S-PROG during Sydney 2000 was mixed: on the one hand it provided useful guidance on the mean, observed longevity of small-scale precipitation features (i.e., thunderstorms); on the other, S-PROG forecasts were of little use for short-range, severe weather warning, where the most important forecast attributes are peak rain rate and rain accumulation. Potentially, this deficiency could be overcome by extending the S-PROG scheme to generate a set of ensembles that would capture the evolutionary uncertainty in the rain-rate distribution, while preserving the statistical properties of the observed precipitation field.

One alternative approach to that adopted by S-PROG involves blending a relatively high-resolution extrapolation forecast with a coarser-resolution NWP forecast. This is the basis of the Nimrod QPF scheme. Such a method has the advantage that the resultant forecast is both smoothed in space to reflect the loss of deterministic predictive skill at small scales, and, by virtue of its NWP component, is able to capture the synoptic-scale evolutionary trends in the precipitation field that are generally well handled by NWP models. Unfortunately, the performance of Nimrod during Sydney 2000 suggests that the benefits of this technique in Sydney were minimal. Ebert et al. (2004) show that Nimrod’s performance had declined to a minimum by between $T + 2$ and $T + 3$ h, when the influence of the NWP component tends to become important. This finding may be due to the predominance of convective rainfall during the FDP, and the fact that the Mesoscale Limited Area Prediction System (MesoLAPS; Puri et al. 1998) does not handle this well. Studies in the United Kingdom have supported this approach to QPN (Bowler and Pierce 2002).

2) Issue 2: The Choice of Advection Scheme and Its Impact on Performance

The majority of automated precipitation nowcasting tools rely upon the extrapolation of radar-observed precipitation. As summarized earlier, the motion of the precipitation can be identified by a variety of techniques, including pattern matching (cross correlation of precipitation fields or component rain objects: ANC, Nimrod), centroid tracking (thresholding of reflectivity fields or volumes to produce discrete objects whose attributes are matched in successive radar scans: TITAN), and NWP model wind advection (either discrete rain objects or individual pixels are assigned a velocity vector derived from a NWP model wind forecast: Gandolf, Nimrod).

Here, the relative merits and failings of these extrapolation techniques are illustrated by reference to two precipitation events with quite different meteorological characteristics. The pertinent findings of Ebert et al. (2004) are also summarized.

The severe thunderstorms of 3 November 2000 posed many challenges for the nowcasting systems participating in the FDP. As demonstrated by Wilson et al. (2004), the evolution and motion of these storms were complex, and markedly nonlinear. A critical factor in these respects was the interaction of the existing storms with a sea-breeze front. Among other things, this produced a dramatic change in storm motion. Figures 2b and 2c illustrate the impact of this sea-breeze front collision on the subsequent behavior of the storms, and the performances of Nimrod, Gandolf, and TITAN. Without exception, all the QPF algorithms failed to capture the directional shift in storm motion between 0430 and 0500 UTC (Wilson et al. 2004; Sills et al. 2001). This observation should come as no surprise given the reliance of the pattern matching and centroid tracking techniques on the linear extrapolation of past motion. The use of a simple NWP-based wind advection algorithm (Hand 1996) in Gandolf, although nonlinear, proved equally poor.

On 12 November 2000, the ascent of a warm, moist easterly flow over the Blue Mountains produced an almost stationary belt of precipitation over the high ground surrounding the Sydney basin (Fig. 3a). The lack of motion in the precipitation field as a whole was captured well by the pattern matching algorithms implemented in the ANC and Nimrod (Fig. 3b). Conversely, the centroid tracking [Storm Cell Identification and Tracking (SCIT) algorithm2] and NWP model wind advection (Gandolf) techniques performed poorly, exhibiting a tendency to move the precipitation away from the mountains (Fig. 3b). In the former case, the storm tracks identified by SCIT (Johnson et al. 1998) were almost entirely spurious.

These two case studies largely confirm the performance characteristics of the extrapolation algorithms reported elsewhere (i.e., Wilson et al. 1998). During Sydney 2000, centroid trackers (TITAN) tended to provide superior nowcasts of the motion of upright convection and its associated precipitation when compared with the other algorithms: over all precipitation events, the mean, $T + 60$ min forecast location error for TITAN was 15 km (see Ebert et al. 2004). Nonetheless, the performance of pattern matching extrapolators using multiple vectors (ANC) was on a par with that of centroid trackers during episodes of nonsevere convection (Wilson et al. 2004).

\footnote{On this occasion the observed rain was below the threshold intensity used for cell identification in TITAN. As a result storm tracks produced by the SCIT algorithm in the Warning Decision Support System (WDSS2; National Severe Storms Laboratory) are shown instead. This comparison is justified because the TITAN and WDSS-2 use similar centroid tracking and extrapolation algorithms.}
Not surprisingly, centroid trackers were less reliable than either linear extrapolation or NWP model wind advection during widespread precipitation events. In these latter cases, the pattern matching extrapolators of the ANC and Nimrod proved most successful with mean locational errors for $T + 60$ min forecasts of around 18 km (Ebert et al. 2004).

The NWP model wind advection approach adopted in Gandolf proved least successful in statistical terms. This may be due, in part, to the observed tendency of the life cycle model to decay showers prematurely. However, observations of shower evolution in the Sydney basin during the FDP suggest that the concept of a steering-level wind for showers may not always be meaningful in regions of complex topography. Conversely, limited evidence from the United Kingdom indicates that the advection of showers with NWP model winds may be superior to cross-correlation-based linear extrapolation (Bowler and Pierce 2002). Either way, the evidence is inconclusive.

Over large domains where rotational and shear motions may be evident in fields of widespread precipitation, the use of multiple extrapolation vectors for identified contiguous rain areas (either from NWP model winds or cross correlation) is likely to be more appropriate than a single-field-wide extrapolation vector (Wilson et al. 2004; see also Germann and Zawadzki 2002).

3) ISSUE 3: THE IMPACT OF ALGORITHM COMPLEXITY AND AUTOMATION ON PERFORMANCE

The basis for the majority of short-range QPF algorithms is a linear extrapolation of recent radar echo motion. For all its simplicity, this method has proved to be the most reliable in published nowcasting studies,
and, to date, there has been little empirical evidence to support the trending or more sophisticated modeling of precipitation intensity in the very short range (<1 h ahead) and at high spatial resolutions (≤5 km). Wilson et al. (1998) concluded that such approaches have yet to demonstrate any consistent, quantitative advantages over linear extrapolation.

Mean statistics compiled from all precipitation events observed during Sydney 2000 largely confirm these earlier findings. Significantly, however, the ANC was able to demonstrate some modest improvement over linear extrapolation during certain, key, boundary layer–forced convective precipitation events. Wilson et al. (2004) attribute this improvement to the successful prediction of zones of convective initiation, growth, and decay. This achievement was not matched by any of the other more sophisticated QPF schemes evaluated during the FDP. Notably, the life cycle modeling approach used in Gandolf proved inferior to linear extrapolation on most occasions.

The object-oriented, convective life cycle models implemented within the Gandolf system utilize radar-observed vertical reflectivity profiles, and various mesoscale NWP forecast convection diagnostics to model the growth and decay of individual convective cells. The enhanced predictive capabilities (over linear extrapolation) reported of a similar scheme in the United Kingdom have been attributed to the successful diagnosis of severe convective environments and the conceptual modeling of the self-propagating characteristics of severe storms (Hand 1996; Pierce and Cooper 2000).

During convective precipitation events, including that of 3 November 2000, Gandolf exhibited a marked tendency to predict the rapid decay of observed convection (Wilson et al. 2004). On occasions, this appeared to be due to a static stability bias in the MesoLAPs model...
troposphere, and the consequent absence of sufficient buoyant instability to maintain vigorous convection. With respect to the severe storms of 3 November 2000, additional, contributing factors were inadequate tuning of the life cycle model parameters and the failure of the Mesolaps model to resolve the marked low-level wind shear that was central to the development and maintenance of the observed supercells.

The relatively poor performance of the Gandolf system during Sydney 2000 is confirmed by Ebert et al. (2004) in their statistical validation of algorithm performance. Gandolf’s mean $T + 30$ min frequency (rain area) bias of 0.5 (rain rate threshold $>1$ mm h$^{-1}$) compares with 0.9 for the ANC and S-PROG, 1.1 for Nimrod, and 1.3 for TITAN (reflectivity threshold of 35 dBZ). The CRA intensity error [expressed as a percentage of the mean squared error (mse)] in Table 2 confirms the observed tendency of Gandolf to decay the majority of the convection prematurely.

Of the remaining schemes, S-PROG exhibited a level of statistical (rmse) performance superior to that achieved using linear extrapolation (Ebert et al. 2004). However, this improvement results from a spatial smoothing of the predicted precipitation field. While this has the effect of minimizing the rms forecast error, the loss of critical information on peak rain intensity and accumulation renders the product far less valuable to the severe weather nowcaster.

5. Summary and conclusions

The Sydney 2000 FDP afforded a unique opportunity to demonstrate and compare the performance of a range of nowcasting tools, including those designed for short-range quantitative precipitation nowcasting (QPN). These QPN algorithms employed advection techniques variously based upon linear extrapolation (Nimrod, S-PROG, ANC), centroid tracking (TITAN), and NWP model winds (Nimrod, Gandolf). Some have the facility to nowcast trends in precipitation intensity, either by extrapolation of observed trends (TITAN—this component was not functional during the FDP), or by the application of more sophisticated statistical, time series techniques (S-PROG), conceptual models (Gandolf), or physically based nowcasting parameters (ANC).

The majority of the schemes generated nowcasts with a maximum range of 60 min (TITAN, S-PROG, ANC), but some of those exploiting information from NWP models routinely produced products with extended ranges of 120 min (Gandolf) or even 360 min (Nimrod). Two of the schemes were implemented within systems that allowed manual input or intervention (ANC, TITAN), while the remainder were fully automated (S-PROG, Nimrod, Gandolf).

Care is needed in the interpretation of the quantitative performance statistics produced by Ebert et al. (2004). For example, in comparing the statistical performance of Nimrod with the other schemes, the importance of data resolution must be recognized. All other things being equal, the use of radar data with a coarser resolution (5 km as opposed to 2 km) will have the effect of reducing the magnitude of the discrepancies between forecast and observed precipitation fields, and may be an advantage in terms of rain–no-rain evaluation when the size of individual, contiguous rain areas is small relative to that of a single, coarse-resolution data pixel.

Furthermore, while there may be scientific justification for employing techniques that smooth a forecast precipitation field with advancing lead time (as employed by S-PROG and Nimrod), the consequent improvement in quantitative statistical performance does not necessarily imply that the resultant product is more valuable to the weather forecaster or hydrologist. For the purposes of short-range flood warning, for example, preservation of the observed precipitation intensity distribution for estimating maximum accumulation is likely to be more important than smoothing the precipitation field in a way that is consistent with the growth of uncertainty in rain-rate estimates.

With these provisos in mind, some clear trends do nevertheless emerge concerning the relative merits and failings of the QPN schemes. When the performance of nowcasts with a range of 60 min is assessed over a variety of precipitation events, including cases of widespread rain, and showers, it is the linear extrapolation schemes that are generally superior. For small domains, a single, field-wide extrapolation vector may be adequate (S-PROG) to capture the observed movement, but for large domains, or for events involving marked, differential motion of rain areas (e.g., severe convection), multiple vectors are likely to produce better guidance (Wilson et al. 2004).

For convective QPN there is little to choose between centroid trackers and multiple-vector, pattern matching techniques (e.g., TREC), although the former are specifically designed to handle the merging and splitting of cells, and may be more reliable extrapolators of showers because they use radar reflectivity volumes rather than 2D reflectivity fields as input. However, the dependence of centroid trackers upon thresholding techniques to distinguish individual convective objects makes them inappropriate for the nowcasting of stratiform rain. This raises the related issues of algorithm robustness and automation. Algorithms that are designed for operation under specific meteorological conditions must be capable of recognizing those conditions if they are to be automated satisfactorily. This requirement complicates algorithm design and provides additional scope for error, as demonstrated by the poor performance of an experimental version of Gandolf during Sydney 2000.

Perhaps the greatest challenge in convective QPF is the deterministic prediction of the new convection in clear air. Currently, this is largely beyond the capabilities of operational nowcasting and NWP models. However, in cases of in situ, boundary layer forced convect-
tion some measure of success has been demonstrated by the ANC as shown on 3 November 2000 (Wilson et al. 2004; Fox et al. 2004, in this issue). The ANC uses high-resolution Doppler wind fields and various observation-based convection diagnostics to identify convective trigger and development zones (Mueller et al. 2003). Gandolf also attempts to forecast the development of showers in clear air by the postprocessing of mesoscale NWP fields. This technique proved largely unsuccessful in Sydney. In the United Kingdom, a more sophisticated, stochastic methodology has met with some success (Hand 2002).

One of the criticisms leveled at Lagrangian persistence approaches to QPF is the extrapolation of spurious, fine-resolution detail in the precipitation fields. Verification statistics from Sydney 2000 show that 2-km-resolution Lagrangian persistence forecasts are little better than those of Eulerian persistence beyond about $T + 60$ min. In recognition of the limited predictability of precipitation at fine spatial scales, both Nimrod and S-PROG smooth their forecast precipitation fields with advancing lead time. In S-PROG this smoothing is consistent with the observed temporal persistence of precipitation features at different spatial scales.

Of greater merit for longer-range nowcasting (up to 6 h ahead) is the method of smoothing adopted by Nimrod in which extrapolation forecasts are blended with those of a coarser-resolution mesoscale NWP model. Previously published statistics show that these “merged” forecasts are superior to those of their component forecasts in the approximate range $T + 90$ to $T + 240$ min (Bowler and Pierce 2002). This methodology is also appealing in that it provides a sound means of introducing significant, broader-scale trends in growth and decay that derive from the NWP-predicted atmospheric evolution.

Neither of the above methods provides an ideal means of accounting for the growth of error in forecast precipitation fields, while retaining the information content pertinent to severe weather and hydrological forecasters. A better approach currently being explored jointly by the BoM and the Met Office involves generating nowcast ensembles that can capture the evolutionary uncertainty in rain intensity while retaining the spatial detail required of the end user.

Although the Sydney 2000 FDP has served to confirm many of the earlier reported characteristics of short-range QPF schemes, a number of key aspects of QPF remain to be addressed within the FDP framework. Most importantly, the issues surrounding surface rain-rate estimation by vertical reflectivity profile correction and rain gauge adjustment were not explored because the emphasis was very much on severe weather forecasting. Despite these limitations, Sydney 2000 was successful both in demonstrating the operational viability of state-of-the-art nowcasting tools, and in establishing a framework for the sharing of research and development with other member states in the WMO.

Acknowledgments. The authors would like to thank the other participants in the FDP: Bill Conway, Jannelle Johnson, DeWayne Mitchell, and Arthur Witt of NSSL; and David Hudak, Paul Joe, David Patrick, Dave Sills, and Pierre Vaillencourt from the Meteorological Service of Canada. We are very grateful to our Australian hosts at the Bureau of Meteorology regional office in Sydney, and also to the organizers of the project: Tom Keenan, John Bally, Sandy Dance, Geoff Freeman, Peter May, and Phil Purdom of the BMRC.

REFERENCES


