RADAR NOWCASTING OF HIGH INTENSITY RAIN EVENTS IN AUCKLAND AND WELLINGTON

L. Sutherland-Stacey & J. Nicol (Weather Radar New Zealand Ltd.), N. Brown & H. Foreman (Auckland Council Healthy Waters Department), B. Fountain (Wellington Water), G. Austin (University of Auckland) & T. Joseph (Mott MacDonald Ltd.)

ABSTRACT

It can be a challenge to manage the complex flooding risks in urban areas. With New Zealand's topography and climate, many densely developed and vulnerable communities are often located in small, steep flood prone catchments. Those attempting to minimise damage to property and prevent injury or loss of life are increasingly looking for alternative solutions to expensive infrastructure.

Recent short-duration / high-intensity rainfall events in both Auckland and Wellington have led to notable flooding events resulting in inundation of habitable floors and disruption to the community. Weather forecasts based on global numerical models provide valuable advanced warning but have lacked the resolution in time and location to facilitate targeted operational responses to flooding. On the other hand, the telemetered rain gauge network that provides real time intelligence does not provide sufficient warning in small catchments with short time-to-concentration.

Rainfall radar is the best available technology for measurement of the spatial distribution and evolution of the short-duration / high-intensity rainfall events which have been a major contributor to flooding damages. In order to make better use of the available national rain radar observations from the NZ MetService network, the authors established a collaboration which has already generated a shared real-time GIS platform for regional radar-derived Quantitative Precipitation Estimation (QPE).

To extend the operational usefulness of the radar data and provide enhanced warnings in catchments with short response times, a Quantitative Precipitation Forecast (QPF) based on a radar echo extrapolation nowcasting method has also been implemented. The radar nowcast QPF is generated by the Short Term Ensemble Precipitation System or "STEPS" and provides ensemble estimates of possible rainfall distributions up to 2 hours into the future, every 7.5 minutes.

The existing Auckland Council / Wellington Water radar QPE product, which includes clutter suppression, attenuation corrections, advection-interpolation and gauge scaling provides the input data for STEPS. The millions of data points which comprise the radar QPE and radar nowcast QPF are combined on-the-fly in a web-based GIS portal, allowing for treatment of antecedent condition and intensity/duration accumulation-based alarming in target catchments. The system provides a platform to enable location specific operational response to the event as it occurs and ultimately alerting to at risk properties.

In order to characterise the operational performance of the catchment alarming system, the QPE analysis and radar nowcast QPF have been run in hindcast mode covering the last 2 years. The radar nowcast QPF is validated against the radar QPE product at stormwater-catchment scale, achieving the highest skill at shorter lead times. This is the expected result given the chaotic nature and limits of prediction of the evolution of convective systems. Case studies of the (hindcast) performance of STEPS for the recent flooding events are also presented and implications for operational hydrology and hydraulic modelling are discussed.

KEYWORDS

RAIN-RADAR, FLOODING, FORECASTING, NWP, NOWCASTING

PRESENTER PROFILE

Dr. Luke Sutherland-Stacey is the Director of Weather Radar NZ Limited and an expert in real-time, high space/time remote and in-situ sensing.

1 INTRODUCTION

It can be a challenge to manage the complex flooding risks in urban areas. With New Zealand's topography and climate, many densely developed and vulnerable communities are often located in small, steep flood-prone catchments. Those attempting to minimise damage to property and prevent injury or loss of life are increasingly looking for alternative solutions to expensive infrastructure.

Recent short-duration / high-intensity rainfall events in both Auckland and Wellington have led to notable flooding events resulting in inundation of habitable floors and disruption to the community. Accurate estimation of the spatial and temporal variation of rainfall across urban catchments is essential for accurate sewer and stormwater modelling, and operational activities. Sampling the true areal rainfall with rain gauges is difficult, because rainfall varies on spatial scales much smaller than the typical separation between gauges (Morrissey et al. 1995, Steiner 1996, Nystuen 1998, Villarini et al. 2008). Forecasting rainfall for stormwater catchments is equally challenging, as the effective spatial resolution of Numerical Weather Prediction (NWP) models is typically five times the grid spacing (Harris et al. 2001), meaning the 4 or 1.5 km spacings available in New Zealand is still too coarse for detailed stormwater catchment modelling, aside from the lack of predictability at these scales.

The implications of spatial undersampling of rainfall measurements for stormwater and sewer modelling activities are well documented (e.g. Berne et al. 2004, Cooper and Fernando 2009). Too sparse gauge spacing may lead to significant under and over estimation of rainfall over short time periods and therefore guidelines about minimum gauge spacing have been developed (e.g. ARC 1999, WaPUG 2002). However, even if it were possible to meet deployment and running cost of the hundreds of gauges which would be required to adequately instrument the Auckland region, adherence to minimum gauge density requirements can be difficult in urban settings because of the limited availability of sites compliant with World Meteorological Society guidelines for rain gauge deployment (WMO 2008). Deployment of rain gauges too close to buildings may cause shadowing, while deployment above ground level (e.g. on rooftops) can result in significant low biases due to wind flow.

Rain radar is a well-established technology for addressing the spatial sampling problem. Rain radar has been used in a variety of stormwater (Löwe et al. 2014), runoff (Shaw et al. 2010) and sewer system modelling (Sempere-Torres et al. 1999, Heinonen et al. 2013) applications internationally. Auckland Council Healthy Waters and Watercare Services Limited have reported on their experience with the use of rain radar in sewer modelling (Joseph et al. 2014) and stormwater (Sutherland-Stacey et al. 2016) settings and recently implemented operational use of weather radar (Sutherland-Stacey et. al. 2017).

New Zealand has been covered by the network of weather radars run by the Meteorological Service of New Zealand Limited (MetService) for many years (for a review, see Crouch 2003). However, until recently there have been only limited attempts to make use of radar data in stormwater and wastewater engineering applications. Limited use of rain radar measurements in the engineering modelling community in New Zealand may be attributed to the technical barriers which exist in making use of complex radar data compared to simpler rain gauge measurements (for a discussion, see Milsom 2007).

In order to remove these barriers and foster more widespread use of radar data, automated data quality control for the MetService C-band radars, and operationalised real-time calibration of the radar precipitation estimates using the regional rain gauge networks has been implemented. The high quality radar derived accumulations are prepared at spatial and time resolutions suitable for urban hydrology (1-minute time step, 500x500m pixel resolution rasters).

Extrapolation Nowcasting is an effective method to extend the radar observations into a short term forecast. To generate forecasts, the radar quantitative precipitation analysis is used as input data for the Short Term Ensemble Precipitation System (STEPS) (Bowler et al 2006). STEPS is a probabilistic nowcasting system which deals with both the evolution and advection of the observed precipitation in order to generate estimates of possible future rainfall in the 0-2 hour range.

2 METHODOLOGY

Extrapolation nowcasting relies on both a nowcasting model and accurate, spatially resolved rainfall estimates.

2.1 RADAR RAINFALL ESTIMATES

The Auckland and Wellington regions are served by an extensive rainfall observation network (Figure 1) comprising of telemetered tipping bucket rain gauges run by regional government and a single-polarisation, Doppler weather radar run by NZ MetService Ltd.

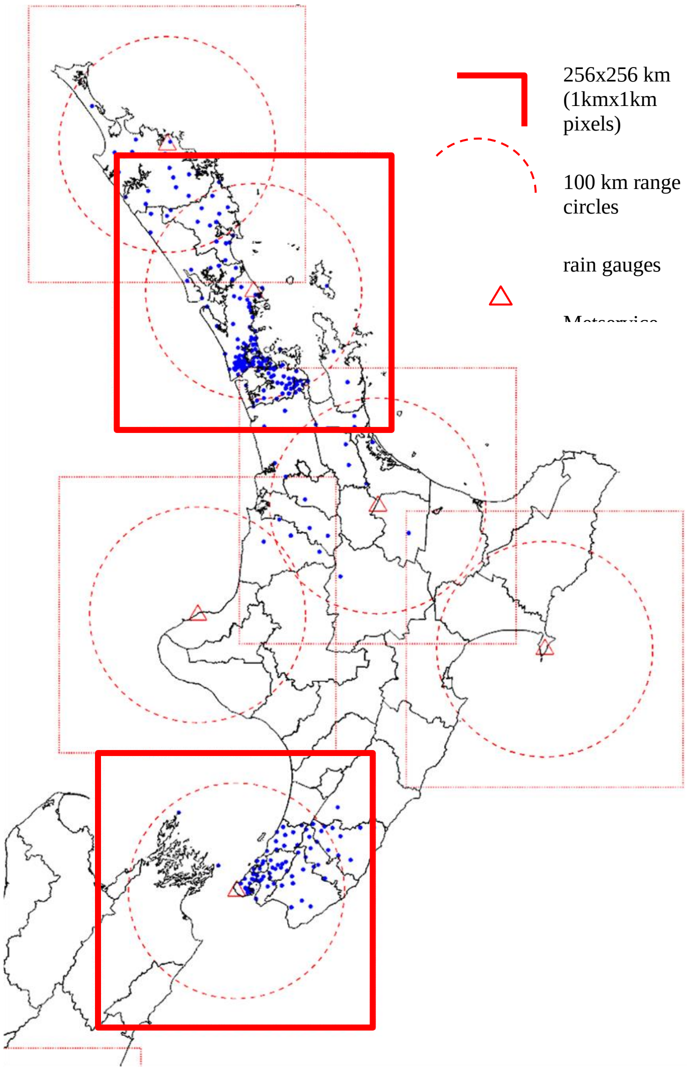

Figure 1: Map of the Auckland and Wellington radar nowcasting analysis areas.

Rainfall accumulation estimates are prepared for the Auckland and Wellington regions by combining the weather radar and rain gauge observations (Figure 1). Extrapolation nowcasting is generated automatically from the radar analysis.

2.1.1 QUANTITATIVE PRECIPITATION ESTIMATION WITH RADAR

MetService operate a number of scanning rain radars around New Zealand (for description of the radar see Crouch 2003). The radars perform a scan cycle every 7.5 minutes, measuring radar reflectivity at increasing altitudes and at up to 250 km in range.

MetService operate a single-polarisation, C-band (5.4 cm wavelength) scanning rain radar located on Mount Tamahunga near Warkworth. The radar is well positioned to provide observations of the meteorological situation for both Auckland and the Northland regions. However, the most southern parts of the Auckland region are over 100 km away from the radar. As such, radar measurements in South Auckland are made between 1.5 and 3 km above the ground which somewhat decreases the representativity of the measurements.

The Wellington Radar is located immediately to the west of the city. The radar is very well positioned to provide meteorological observations for the Wellington urban area, however the significant topography around the region means many sectors of the radar scans are impacted by "beam blocking".

Photograph 1: The NZ MetService scanning radars at Mount Tamahunga (left, courtesy MetService Ltd.) and Wellington (right)

Radar is an active sensing technology which illuminates targets with electromagnetic energy and measures the properties of the reflected (or "back-scattered") radiation in order to elucidate some physical property of the targets. In the case of meteorological radars, repetitive pulses of electromagnetic energy are focused into the distance by a parabolic dish, by scanning the dish and recording the bearing and time taken for pulses of energy to return, a map of precipitation location and intensity can be constructed.

The principle radar measurement is reflectivity (Z, mm^6m^{-3}), which for meteorological applications is the scattering cross section of all the targets in the radar beam at a particular range bin:

$$Z = \int_0^\infty D^6 N_v(D) dD$$ (Equation 1)

where D is the drop diameter and N_v is the number of drops with that diameter. Reflectivity is usually expressed in decibel units, and values typically range from 20 dBZ for light rain to 55 dBZ for very heavy rain. Values over 55 dBZ are likely to indicate solid precipitation (hail).

The scattering cross section, and hence reflectivity, depends on the usually unknown raindrop size distribution, and must be converted to rainfall rate (R, mm hr^{-1}) to be useful. Other factors influencing the estimation of rainfall which must also be taken into account are attenuation, ground clutter, beam blocking, uncertainty in the vertical profile of reflectivity, spatial smoothing and time intermittency of the radar measurement.

2.1.2 RADAR DATA ARCHIVE FOR AUCKLAND AND WELLINGTON

For hydrological applications, detailed quality control and processing is required to generate useable rainfall estimates. For Auckland Council Healthy Waters Department and Wellington Water's requirements, precipitation estimates were required at sub-hourly frequency, sub-kilometre resolution and with minimum systematic bias and error. This level of detail and accuracy was not available from the 1-hour accumulation product generated by the C-band radar's bundled software, so raw radar data in polar format (range, bearing and reflectivity) were sourced directly from the C-band radar output files and ingested in the cloud based GIS system through a customised post processing system; see Sutherland-Stacey (2017, 2018) for a description.

Radar data is available back to about 2007 for both the Auckland and Wellington regions. While the weather radar network has been operated for significantly longer than a decade, the historic measurements have not been archived. The lack of a longer data set is unfortunate, given the requirement to understand the high-impact / low-frequency rainfall hazards associated with stormwater operations. Nonetheless, both Wellington Water and Auckland Council Health Waters Department have arranged access to all available radar data.

2.2 RADAR NOWCASTING

The physical basis of all radar nowcasting systems is the observation that real-time radar estimates of precipitation contain a large amount of useful information about the current intensity and distribution of rainfall. Determination of the motion of the current rain field allows estimation of its future position and hence construction of rainfall accumulation forecasts (Figure 2). While the skill of radar nowcasting falls away quickly with lead time, the inclusion of real-time observations means nowcasting outperforms other forms of precipitation forecasting for short lead times

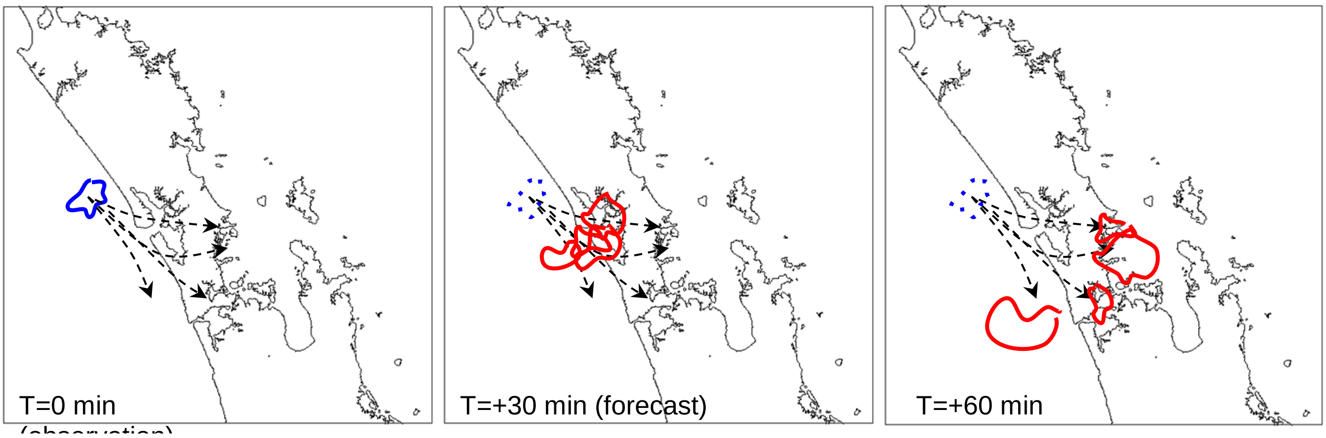

Figure 2: Conceptual representation of the increase in uncertainty in radar extrapolation nowcasts. At the observation time (T=0, left panel) possible paths of the current rainfall are diagnosed. The forecast error (ensemble spread) depends then on the spread of possible paths and evolution of the rainfall structure. At short lead times (middle plot), the possible future locations of the rainfall have not had very great opportunity to diverge, but at longer lead times uncertainty in the motion and development or decay of the rain field can lead to large uncertainty in future rainfall location (right plot).

The uncertainty introduced by divergent estimates in rainfall displacement and development leads to a "crossover" lead time where the NWP forecasts, which include physical details of the underlying weather systems and topography, are more skillful than extrapolation nowcasts (Golding 1998). The exact lead time of the crossover depends on the predictability of the rain field evolution, both in terms of advection velocity and internal details. The exact lead time at which NWP becomes more skillful than extrapolation nowcasting depends on the details of the local weather and observation range of the radar network, with varying values suggested in the literature. Mandapaka et al (2012), for example found extrapolation nowcasting to be more skillful up to about 2.5 hours, while Lin et. al. (2005) found a value of six hours. The longer lead time results are obtained using continental scale networks of radars. In the case of the Auckland Council nowcasting system, only the data from the Auckland radar is available, so the maximum lead time for which the Nowcast is useful is limited by the maximum observation range of the single radar.

2.2.1 RADAR NOWCASTING WITH "STEPS"

The Short Term Ensemble Prediction System (or "STEPS", Bowler et al 2006) is a Radar Nowcasting system which has been jointly developed by the UK Met Office (UKMO) and Australian Bureau of Meteorology (BOM) to provide short duration probabilistic precipitation risk forecasts.

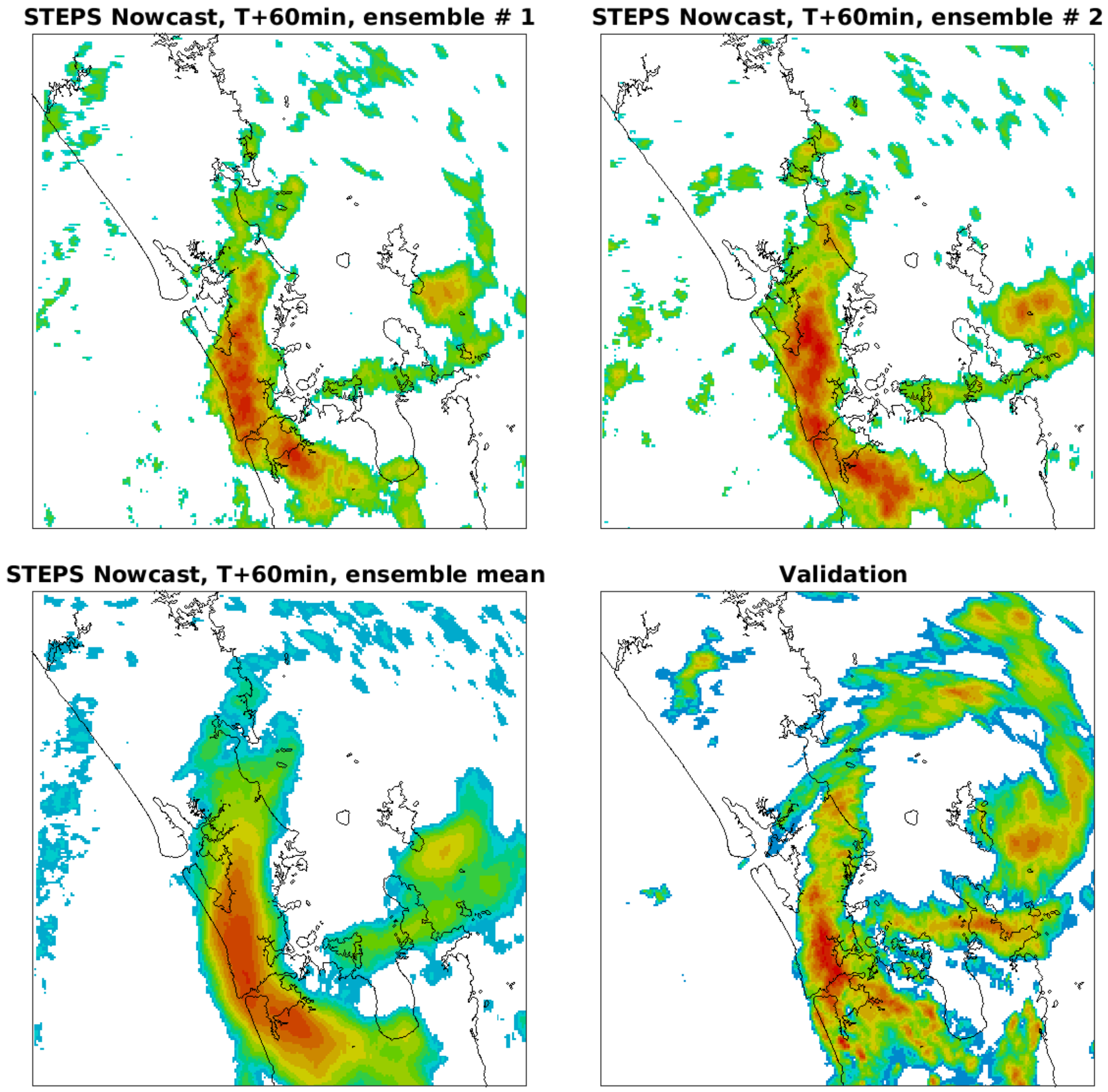

Figure 3: STEPS 60-minute lead time rainfall intensity forecast for Auckland valid at 2017/03/11 22:50 UTC. The first two ensemble forecast members are plotted (top panels). All 30 ensembles are averaged to construct the ensemble mean (bottom left). The validating rainfall observation is also provided (bottom right). In this example, STEPS properly forecasts the location of the rainfall hazard over West Auckland 60 minutes into the future.

STEPS partitions the uncertainty in the future rainfall into two parts:

1) The advection of the existing rain field in space
The advection of the rain field is estimated from the observed past motion of the rain field. Perturbations to the speed and direction of the advection are introduced to account for uncertainty in the future trajectory of the rainfall.

2) The evolution of the existing rainfall patterns in time
Modeling the evolution of the rain field is achieved with the Spectral Prognosis (S-PROG) model (Seed 2003). S-PROG decomposes the rain field into a multiplicative cascade at a range of spatial scales, which allows the persistence of the rain field at each spatial scale to be estimated and filtered. Rainfall details which are estimated to be short lived and unpredictable are gradually replaced with noise with equivalent spectral characteristics to the observed rainfall. In this way, ensembles of the possible evolution of the rain field are generated.

A detailed technical description of the STEPS system can be found in Bowler et. al. (2006). STEPS is implemented for Auckland Council Healthy Waters Department and Wellington Water as described in the paper, with the exception that NWP data is not yet included for Wellington Water.

STEPS makes use of gauge-scaled radar accumulation estimates from the rain radar analysis platform domain to construct forecasts. Input data is prepared from the real-time radar analysis as 256x256 pixel (pixel size 1x1 km), 5-minute accumulation rasters. For every 5 minute input radar accumulation estimate, STEPS generates 30 ensembles of possible accumulations for the next two hours at 5-minute intervals. Radar extrapolation nowcasts for both the Wellington and Auckland regions were prepared for the period September 2016 to September 2018. An example of STEPS output for a rain event which caused flooding in the West Auckland suburb of New Lynn in April 2017 is presented in Figure 3.

2.2.2 PRECIPITATION FORECASTING VALIDATION SKILL METRICS

The skill of a precipitation forecast may be evaluated in terms of a contingency table. For deterministic accumulation forecasts, a standard approach is to evaluate each forecast in terms of a success/hit, failure/miss, false alarms and correct negatives, given a particular accumulation intensity /duration threshold.

The threshold approach is appropriate for most rainfall hazards, as generally a rainfall related incident will occur after a certain accumulation threshold is breached, irrespective of by how much. A variety of skill score metrics can be constructed from the aggregated Hits, Misses and False Alarms and used to assess model performance.

Table 1: Skill Score Metrics

		rain over threshold X mm observed?	
		yes	no
forecast over threshold M mm issued?	yes	Success ("Hit")	False Alarm
	no	Failure ("Miss")	Correct Negative

Meteorological forecasting validation makes use of a metric referred to as *Probability of Detection.* When testing for exceedance of a rainfall threshold POD is equivalent to *sensitivity.* Equally, POD is equivalent to *specificity* when testing for correct prediction of low rainfall conditions. Meteorological verification is often extended by consideration of *BIAS, False Alarm Ratio and Equitable Threat Scores* (Manson 2003). The extra metrics can be useful for understanding the overall skill of the meteorological forecast model.

Frequency BIAS gives the ratio of the frequency of forecast events to observed events and reflects if the forecast system tends to under-forecast (BIAS<1) or over-forecast (BIAS>1) events.

$$frequency\ BIAS = \frac{hits + false\ alarms}{hits + misses}$$

Probability of Detection (POD) reflects the fraction "yes" events which were correctly forecast. POD is sensitive to hits and ignores false alarms. A perfect forecasting system would have POD=1.

$$POD = \frac{hits}{hits + misses}$$

False Alarm Ratio (FAR) indicates the fraction of "yes" events which were incorrectly forecast (did not actually occur, so were "false alarms"). A perfect forecasting system would also issue no false alarms (FAR=0).

$$FAR = \frac{false\ alarms}{hits + false\ alarms}$$

Both POD and FAR are sensitive to the relative frequency of forecast events and should be considered together. For example, if the forecast system tends to over-predict events then POD will appear better but FAR worse. Therefore, it is useful to make use of the **Equitable Threat Score** which measures the fraction of observed and/or forecast events that were correctly predicted, adjusted for hits associated with random chance (for example, it is easier to correctly forecast rain occurrence in a wet climate than in a dry climate).

$$\boldsymbol{ETS} = \frac{hits - hits_{random}}{hits + misses + false\,alarms - hits_{random}}$$

$$hits_{random} = \frac{(hits + misses)(hits + false\,alarms)}{total}$$

The *ETS* is often used in the verification of rainfall in NWP models because its "equitability" allows scores to be compared more fairly across different regimes.

2.2.3 APPLYING PRECIPITATION SKILL SCORES TO PROBABILISTIC FORECASTING

Deterministic forecasts are relatively straightforward to verify with skill score metrics, in that for each forecast location there is a clear success or failure to predict rainfall over an observed threshold. Verification of probabilistic forecasts is more complex, as for each forecast location (for example, a catchment) the forecast system generates an ensemble of possible forecasts (an ensemble of possible accumulation hyetographs).

A simple approach to applying the deterministic skill score methodology to an ensemble forecast is to decide that a "yes" event has been forecast when *n* or more of the ensemble members breach the event threshold. Equivalently then, a forecast is "no" when less than *n* of the ensemble members breach the event threshold. If the precipitation forecasting system is unbiased at the target intensity/duration threshold compared to the rainfall climatology (**frequency BIAS**=1), then *n* should be set to *N*/2, where *N* is the total number of ensemble members.

First, Nowcast results were prepared for skill score assessment from each of the 30 ensemble members subsampled from the 5-minute update frequency to 30-minute frequency, in order to reduce the error correlation between sequential forecasts. Then, the 5-minute interval, 1x1 km pixels were accumulated to generate 30-, 60- and 120-minute per-pixel accumulations. STEPS generates 2-hour forecasts in the default radar extrapolation configuration (used here) therefore 30-minute accumulations can be validated at four independent lead times (0-30, 30-60, 60-90 and 90-120 minutes). Likewise, the 120- and 60-minute accumulation forecasts can be validated at one and two unique lead times respectively. 30-, 60- and 120-minute accumulations are representative of durations likely to be important for small impervious catchments.

For stormwater applications arguably the primary hazard is associated with high intensity, short duration localised rainfall events. In an operational context it may not be necessary to know the exact (per-pixel) location of such events, but rather the general area immediately at risk of deep accumulations. Therefore, following preparation of the per-1x1km pixel 30, 60 and 120 minute rainfall depths, the data is subsampled to identify the highest depth accumulations at each lead time and duration within 10x10km regions, across the analysis domain and for each ensemble member and the validation data. The validation metrics can then be applied to the sub-sampled data set to determine the forecast skill for localised exceedance of a critical depth over a 1x1 km area *anywhere within* a 10x10km sub-region (Figure 4) .

Due to the relative infrequency of large events in the one-year radar record processed so far, accumulation thresholds from 1 - 10 mm at each duration have been considered. By way of comparison, the HIRDSv4 (Carey-Smith et. al., 2018) 1.58-year ARI depth for 60-minute accumulations is 23.9 mm. A longer radar data series is clearly required to allow better estimation of the nowcasting system skill for rarer events.

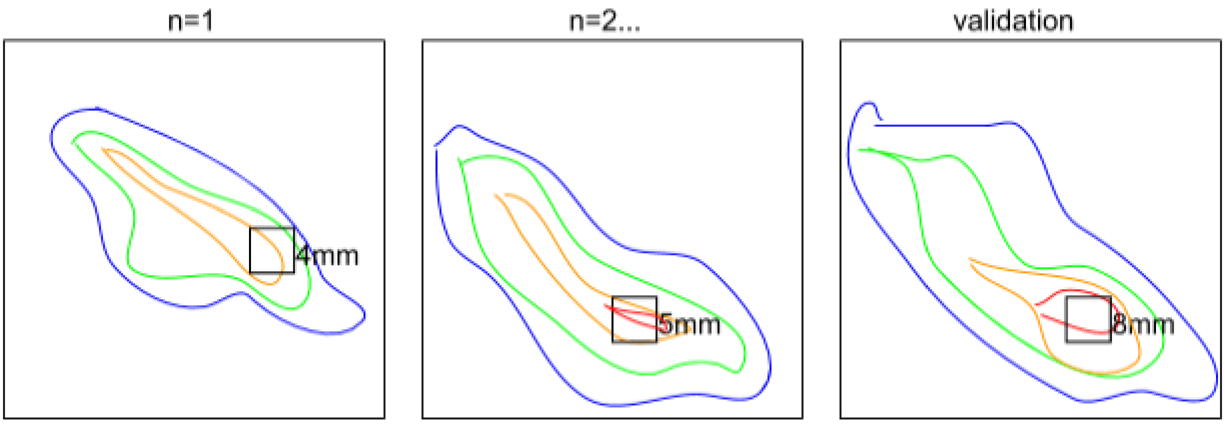

Figure 4: Schematic of the nowcasting validation metric. Within each 10x10km sub-region (large black bounding boxes) the local maximum depth (small black boxes) is identified in the underlying high resolution rainfall data (coloured contours) for each of the 30 ensemble members (n=1,2...) and validation data. Skill scores are then calculated for the entire 10x10km sub-region from those local maximum depth values. In this figure ensemble member 1 records a maximum depth of 4mm, ensemble member 2, 5mm and the validation data had a local maxima of 8 mm. Both ensemble members would have recorded a "HIT" at the 4mm threshold and likewise both would have "MISSED" at the 6mm threshold.

STEPS does not include a physical growth or decay model for rainfall, so it is to be expected that orographic effects will introduce systematic biases into the STEPS forecast. Therefore, the forecasts were analysed to identify and remove the long-term biases at each depth/duration/leadtime. The optimal alarming threshold ($n_{unbiased}$) was thereby determined for each 10x10 km alarming region to minimise the **frequency BIAS** in the forecasts at each intensity/duration threshold. By way of example, Figure 5 gives the spatially resolved **frequency BIAS** for *n*=10, 15 and 20 of the 30 total ensemble members agreeing on rainfall depth/durations of over 1 mm in 60 minutes. It is clear from Figure 5 that around higher elevation regions, the local 1x1km maximum accumulation depths derived from STEPS forecasts tend to occur too infrequently compared to the observed rainfall climatology at the 1-mm depth / 60-minute duration, so alarming on a *lower* ensemble agreement level than 50% (i.e. *n*<15) generally provides an optimal **frequency BIAS** result. The reverse appears to be true to lower elevation regions. The physical basis for the spatial **frequency BIAS** differences is discussed in more detail later.

After applying the optimal climatological **frequency BIAS** correction (n_{unbiased}) for each validation catchment, **POD, FAR** and **ETS** scores were then determined for each depth/duration/lead time combination. Selected spatially resolved maps of the skill statistics are provided in Figure 6 for the 60-minute, 5-mm forecast depth/durations at the 0-60 minute lead time. Regional average results (the average of all pixels over land in the Auckland or Wellington regions) are summarised in Table 2.

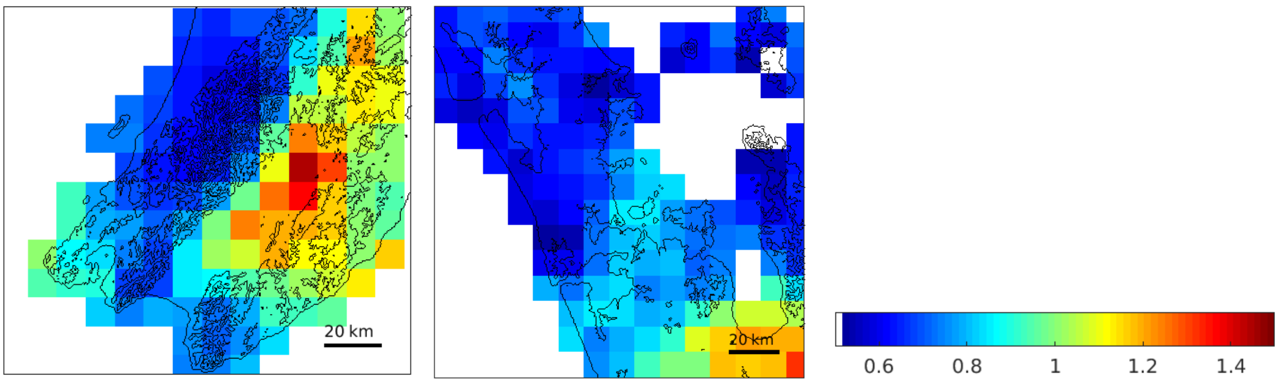

*Figure 5: STEPS **frequency BIAS** for 5 mm depth / 60-minute duration around the Wellington (left) and Auckland (right) regions.*

*A **frequency BIAS** of less than 1.0 (cool colours) indicates that STEPS tends to under-estimate the frequency of occurrence of 5 mm depth / 60 minute, events, while a frequency BIAS above 1.0 (warm colours) indicates an overestimation. Because STEPS overall conserves rain volume during the 2 hour nowcast, frequency BIAS could be due to orographic or coastal enhancement effects or range-related biases in the radar data.*

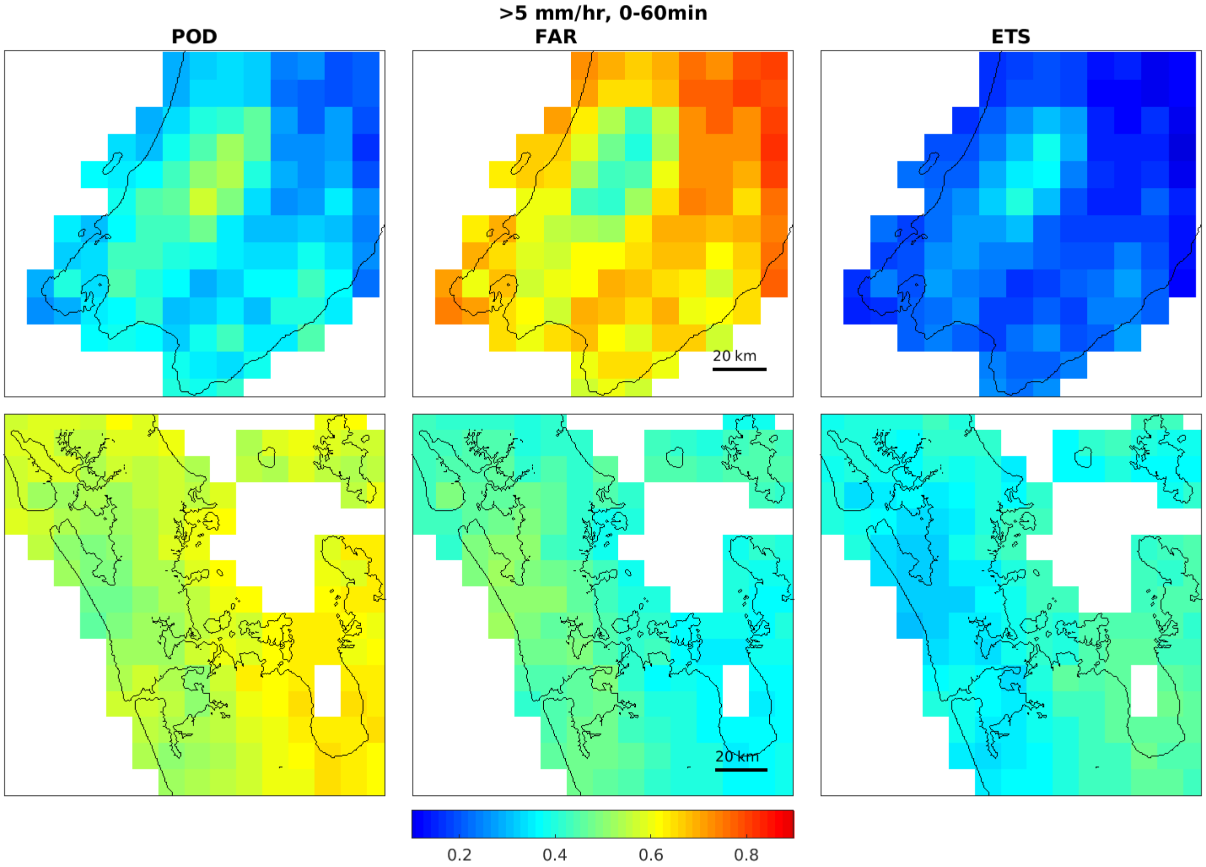

Figure 6: STEPS spatially resolved skill score results for 5 mm depth / 60-minute durations for Wellington (top and Auckland (bottom).
***POD** (left), Probability of Detection, fractional chance an event of 5mm depth over a 1x1km area within each 10x10km sub region will be detected. The long term statistics indicate about 60-70% of events are detected in Auckland, which is an operationally useful skill level.*
***FAR** (middle) False Alarm Ratio, fractional chance a forecast of an event of 5mm depth over a 1x1km area within each 10x10km sub region will turn out to be a false alarm. The long term statistics indicate about 40% of forecasts of 5mm depth exceedances turn out to be false alarms.*
***ETS** (right) Equitable Threat score. Measures the fraction of observed and/or forecast events that were correctly predicted, adjusted for hits associated with random chance.*

3 DISCUSSION AND CONCLUSIONS

The analysis for the STEPS Nowcasting procedure can be used to interrogate the expected performance of the nowcasting system for operations. The skill score statistics:

1) Inform the end-user about the **expected skill of the currently available forecast.**

    a) **POD** (probability of detection) is the long term probability the forecast will detect an event, so on a given day the forecast user knows the probability of receiving a correct warning of a rain event (sensitivity) or correct prediction of low or no rainfall conditions likely to lead to compliant water quality (specificity).

    b) **FAR** (False Alarm Ratio) is the long term probability that an issued forecast event will turn out to be a false alarm; so after receiving a forecast of a rain event, the forecast user knows the chance that the predicted event will not occur. (The Success Ratio (SR) is 1-FAR.)

2) Allow the end-user to systematically **compare the skill of different forecasts**.

    Different forecast providers (e.g. Metservice, NIWA) offer a variety of different weather forecast products. The **ETS** score of each forecast product can be used to efficiently, directly and systematically compare a number of forecasts against the identical validation data at the same thresholds, in order to understand if a new forecast type adds any additional, useful, predictive skill.

In the Wellington region, for prediction of > 1 mm rainfall depth over a 1x1km area within a 10x10km sub-region in 30 minute forecast windows, **POD** (probability of detection) ranges from over 0.6 (60%) starting at the observation time, to 0.3 (30%) for the 90-120 minute window. **FAR** (false alarm ratio) increases from under 0.4 (40%) for the 0-30 minute window, to 0.7 (70%) at the 90-120 minute window. In Auckland the nowcasting validation returns better skill statistics: a regional average **POD** of almost 0.7 for the 0-30 minute window decreasing to 0.3 for the 90-120 minute window.

For both the Auckland and Wellington region, the skills scores assessed at the more infrequent 5mm accumulation depth are slightly lower and drop off more rapidly with lead time, with Auckland continuing to slightly outperform Wellington. More intense rainfall is associated with convective systems which are generally more difficult to forecast due to their short evolution time and chaotic nature, so the slightly poorer skill for higher depths is not unsurprising.

There are a number of possible reasons for the better skill statistics in Auckland compared to Wellington.

Firstly, as STEPS is implemented based on the radar rainfall estimates, better calibration of the radar will also tend to improve the STEPS validation statistics. Auckland benefits over Wellington in this regard. Most significantly, as the Auckland radar analysis includes rain gauge data from the Northland Regional Council Network, and there is a directional bias in the weather patterns which sees a significant proportion of rain events traverse the Northland region, affording the opportunity to "calibrate" the radar estimates in advance. The Wellington region is largely surrounded by ocean (even when compared to Auckland) so there is little opportunity to check and adjust the radar rainfall estimates before they make landfall. Furthermore, the Auckland radar has been rigorously calibrated by comparison with Vertically Pointing Radar, meaning, even in the absence of verifying gauge results, the radar derived rainfall estimates exhibit less bias and spread compared to Wellington.

Secondly, the basic implementation of STEPS demonstrated here includes no consideration of local biases in precipitation growth and decay. Rainfall in both the Wellington and Auckland regions is well known to be strongly influenced by orographic effects. The spatially resolved **frequency BIAS** statistics (Figure 5) highlight the orographic precipitation enhancement effects. Around higher elevation areas STEPS tends to underestimate the frequency of rain over the trigger threshold. Wellington has much higher ranges than Auckland. Likewise, low elevation areas in the lee of the dominant flow direction (notably the Wairarapa) exhibit a high **frequency BIAS**, probably as higher rain observations around the ranges are advected over the dryer regions without consideration for the underlying regional rainfall variations.

There are a number of options for correcting the causes for the poorer STEPS skill in the Wellington region, and indeed further improving skill for the Auckland region.

Both NIWA and Metservice run national rain gauge networks with good spatial extent, however due to commercial data licensing constraints, no gauge data from NIWA, and only two gauges from Metservice, are included in the radar analysis step. If such data could be improved it is likely that the accuracy of the rainfall initial conditions on the edges of the analysis (for rain traversing the upper south Island and surrounding North Island regions) would result in improved STEPS forecasts. Continued development of the Vertically Pointing Radar calibration approach is also likely to improve initial conditions.

In this work spatially resolved **frequency BIAS** is approached by identifying the optimal ensemble member agreement for triggering of alarming for the maximum rainfall depth in a 1x1km area over a 10x10km sub-region. In effect, the approach applied a local climatological bias correction for STEPS's inability to model rainfall growth and decay processes. Work in Switzerland has suggested that further benefit might be realised by considering the direction of the weather system, for example different climatological means are likely to be diagnosed for northerly compared to southerly flows. Additionally, the sub-region size could be expanded to increase the skill of the STEPS analysis for forecasting maximum rainfall depth *somewhere* within a target region.

Perhaps most importantly, the systematic verification of precipitation forecasts provides the opportunity for end users to compare the relative skill of different forecasting systems. Numerical Weather Prediction models, for example, can also generate high resolution rainfall forecasts. In the future, **ETS** metrics for STEPS should be compared to the forecasts available from NIWA and Metservice in order to better understand the benefits and limitations of the respective precipitation forecasting approaches.

The next step in this work is to ingest and analyse all of the historic radar dataset into Auckland and Wellington systems (about 10 years observations) in order to test the skill of STEPS for predicting high high intensity, rare events. If a high **Probability of Detection** and minimal **False Alarm Ratio** score combination can be consistently achieved, real time alerting to the community of an imminent flooding hazard may be possible.

ACKNOWLEDGEMENTS

Radar observations were provided by NZ MetService Limited. STEPS is licenced from the [Australian] Bureau of Meteorology: Alan Seed's suggestions about the implementation of STEPS in New Zealand are much appreciated.

REFERENCES

ARC, (1999) *Guidelines for stormwater runoff modelling in the Auckland region* Auckland Regional Council, Auckland.

Berne, A., Delrieu, G., Creutin, J. and Obled, C. (2004) *Temporal and spatial resolution of rainfall measurements required for urban hydrology* Journal of Hydrology, 299, 166-179.

Bowler N. E., Pierce C. E. and Seed A. W. (2006) *STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP* Quarterly Journal of the Royal Meteorological Society, 132, 2127-2155

Carey-Smith T., Henderson R. and Shailesh S. (2018) High Intensity Rainfall Design System (Version 4). National Institute of Water & Atmospheric Research Ltd.
https://www.niwa.co.nz/sites/niwa.co.nz/files/2018022CH_HIRDSv4_Final.pdf

Cooper, M.R. and Fernando D.A.K. (2009) *The effect of the rain gauge distribution on stormwater models* in proceedings 18th World IMACS/MODSIM congress, Cairns, Australia 13-17 July 2009.

Crouch, J. (2003) *Using radar to diagnose weather systems* Tephra, 20 (June), 46–48
http://www.civildefence.govt.nz/assets/Uploads/publications/tephra-june-2003.pdf

Golding B. W. (1998) *Nimrod: A system for generating automated very short range forecasts.* Meteorological Applications, 5, 1-16

Harris, D., Foufoula-Georgiou, E., Droegemeier, K. K. and Levit, J. J. (2001) *Multiscale statistical properties of a high-resolution precipitation forecast* Journal of Hydrometeorology , 2, 406-418.

Heinonen, M., Jokelainen M., Fred T., Koistinen J. and Hohti H. (2013*) Improved wet weather wastewater influent modelling at Viikinmäki WWTP by on-line weather radar information* Water Science & Technology, 68, 3, 499-505.

Joseph, T., Sutherland-Stacey L., Donald N., Austin G., (2014) Overcoming spatial-temporal rainfall variation using rainfall radar. NZ Water Conference. Hamilton.

Lin, C., S. Vasic, A. Kilambi, B. Turner, and I. Zawadzki, 2005: *Precipitation forecast skill of numerical weather prediction models and radar nowcasts.* Geophys. Res. Lett., 32,

Löwe R., Thorndahl S., Mikkelsen P. S., Rasmussen M.R. and Madsen H., (2014) *Probabilistic online runoff forecasting for urban catchments using inputs from rain gauges as well as statically and dynamically adjusted weather radar,* Journal of Hydrology, 512, 397-407

Mandapaka, P. V., U. Germann, L. Panziera, and A. Hering, (2012): *Can Lagrangian extrapolation of radar fields be used for precipitation nowcasting over complex alpine orography?* Wea. Forecasting, 27, 28–49,

Manson, I.B. 2003 Binary events. Forecast Verification—A Practitioner's Guide in *Atmospheric Science*, I. T. Jolliffe and D. B. Stephenson, Eds., Wiley, 37–76

Milsom, G. D Ward, D., Lind M., Chandramohan P., Mills C., Wilson D. and Patil R. (2007*) Project Storm 2 - A "world class" strategic planning tool* Water New Zealand Conference. 19-21 September, Energy Events Centre, Rotorua.

Morrissey, M.L., Maliekal, J.A., Greene, J.S. and Wang, J. (1995). *The uncertainty of simple spatial averages using rain gauge networks*. Water Resources Research, 31, 2011-2017.

Nystuen, J.A. (1998) *Temporal sampling requirements for automatic rain gauges* Journal of Atmospheric and Oceanic Technology, 15, 1253-1260.

Seed A. W., (2003) *A dynamic and spatial scaling approach to advection forecasting* Journal of Applied Meteorology, 42, 381-388

Sempere-Torres, D., Corral, C., Raso, J. and Malgrat, P. (1999*) Use of weather radar data for combined sewer overflow monitoring and control*. Journal of Environmental Engineering, April, 372−380

Shaw, S.B., Stedinger, J.R. and Walter, M.T. (2010) *Evaluating Urban Pollutant Buildup/Wash-Off Models Using a Madison, Wisconsin Catchment* Journal of Environmental Engineering, 136, 2, 194-203

Steiner, M. (1996*) Uncertainty of estimates of monthly areal rainfall for temporally sparse remote observations* Water Resources Research, 32, 373-388.

Sutherland-Stacey, L., Austin, G. and Shucksmith, P. (2011*) A mobile rain radar for high resolution hydrological observations in New Zealand*. New Zealand Journal of Hydrology 50, 2, 339–360. http://hydrologynz.co.nz/journal.php?article_id=752

Sutherland-Stacey L., Joseph T., Austin G. and Brown N. (2016) *Using Rain Radar to Estimate the Size of Flood Events*, NZ Water Conference, Nelson

Sutherland-Stacey, L., Joseph, T., Williams, K., Nicol, J., Brown, N. and Austin G. (2017) *Operational use of Rain Radar in Auckland*, New Zealand Water Conference proceedings, Claudelands Event Centre, Hamilton, 21 September 2017 https://www.waternz.org.nz/Article?Action=View&Article_id=1340

Sutherland-Stacey L. Williams K. Nicol J., Austin G. (2018) *Improving Operational use of Scanning Rain Radar Estimates with Vertically Pointing Radar*, New Zealand Stormwater Conference proceedings, Queenstown, 24 May 2018 https://www.waternz.org.nz/Attachment?Action=Download&Attachment_id=3245

Villarini, G., Mandapaka, P.V., Krajewski, W.F. and Moore, R.J. (2008*) Rainfall and sampling uncertainties: A rain gauge perspective* Journal of Geophysical Research D: Atmospheres, 113, D11102

WaPUG (2002), *Code of practice for the hydraulic modelling of sewer system* Wastewater Planning Users Group, United Kingdom.

World Meteorological Organization (2008*) Guide to Meteorological Instruments and Methods of Observation* Chapter 6 Measurement of Precipitation, Seventh Edition No-8.
https://www.wmo.int/pages/prog/gcos/documents/gruanmanuals/CIMO/CIMO_Guide-7th_Edition-2008.pdf

APPENDIX

Table 2: Tabulated skill score results at different depth/durations for STEPS nowcasts in Wellington and Auckland regions. The POD, FAR, and ETS metrics are described in the text.

WLG radar, >1mm depth

	0-30min	30-60min	60-90min	90-120min
POD	0.60	0.43	0.34	0.30
FAR	0.40	0.57	0.66	0.70
ETS	0.41	0.26	0.19	0.16

	0-60min	60-120min
POD	0.61	0.44
FAR	0.38	0.56
ETS	0.43	0.26

	0-120min
POD	0.64
FAR	0.36
ETS	0.44

WLG radar, >5mm depth

	0-30min	30-60min	60-90min	90-120min
POD	0.29	0.12	0.06	0.05
FAR	0.72	0.88	0.94	0.95
ETS	0.17	0.07	0.03	0.02

	0-60min	60-120min
POD	0.34	0.14
FAR	0.65	0.85
ETS	0.21	0.07

	0-120min
POD	0.37
FAR	0.61
ETS	0.23

AKL radar, >1mm depth

	0-30min	30-60min	60-90min	90-120min
POD	0.73	0.61	0.53	0.46
FAR	0.27	0.39	0.47	0.54
ETS	0.56	0.42	0.34	0.28

	0-60min	60-120min
POD	0.73	0.58
FAR	0.27	0.42
ETS	0.55	0.38

	0-120min
POD	0.74
FAR	0.26
ETS	0.55

AKL radar, >5mm depth

	0-30min	30-60min	60-90min	90-120min
POD	0.54	0.37	0.27	0.21
FAR	0.46	0.62	0.73	0.79
ETS	0.37	0.23	0.15	0.11

	0-60min	60-120min
POD	0.57	0.37
FAR	0.43	0.63
ETS	0.40	0.22

	0-120min
POD	0.57
FAR	0.42
ETS	0.39