Input Uncertainty in Hydrological Models: An Evaluation of Error Models for Rainfall

Hilary McMillan[1*], Bethanna Jackson[2], Martyn Clark[1], Dmitri Kavetski[3], Ross Woods[1]

[1] National Institute of Water and Atmospheric Research, Christchurch, New Zealand

[2] School of Geography, Environment and Earth Sciences, Victoria University of Wellington, New Zealand

[3] School of Engineering, University of Newcastle, Callaghan, Australia.

For Submission to: Journal of Hydrometeorology: Special Collection on the State of the Science of Precipitation Research

Abstract

This paper presents an investigation of rainfall error models used in rainfall-runoff model calibration and prediction. Traditional calibration methods assume input error to be negligible: an assumption which can lead to bias in parameter and estimation and compromise model predictions. In response, a growing number of studies now specify an error model for rainfall input, usually simple in form due to computational constraints during parameter estimation. Such rainfall error models have not typically been validated against experimental evidence: in this study we use data from a dense gauge/radar network to directly evaluate the form of basic statistical rainfall error models. Our results confirm the suitability of a multiplicative error formulation subject to constraints on rainfall intensity. We show that the standard lognormal multiplier distribution provides a relatively close approximation to the true error characteristics but does not fully capture the distribution tails, especially during heavy rainfall where input errors have important consequences for runoff prediction. Our

research highlights the dependency of rainfall error model on model timestep and catchment size.

1. Introduction

Adequate characterisation of rainfall inputs is fundamental to success in rainfall-runoff modelling: no model, however well-founded in physical theory or empirically justified by past performance, can produce accurate runoff predictions if forced with inaccurate rainfall data (Beven, 2004). The impact of rainfall errors on predicted flow has been highlighted by many authors, including Sun et al., 2000; Kavetski et al., 2002, 2006a, Bardossy and Das, 2008, and Moulin et al., 2009. From a management perspective, inaccuracies in rainfall inputs directly compromise model predictions and hence robust decision-making on water and risk management options. Furthermore, errors in rainfall reduce our ability to identify other sources of error and uncertainty, slowing scientific advancement and compromising the reliability of operational applications. This issue is recognized as a major challenge for hydrological modelling science (Kuczera et al, 2006a).

The impact of input uncertainty on streamflow simulations can be quantified by error propagation, either by using conditional simulation or simply by stochastically perturbing the rainfall inputs. Conditional simulation involves simulating ensemble rainfall fields conditioned on the mean and error of spatial rainfall interpolations (e.g., Clark and Slater, 2006; Götzinger and Bárdossy, 2008). Conditional simulation methods do not require many assumptions on rainfall errors (e.g., Clark and Slater, 2006), but can be time consuming to implement. Stochastic perturbation of rainfall inputs is therefore more common (Reichle et al., 2002; Carpenter and Georgakakos, 2004; Crow and van Loon, 2006; Pauwels and de Lannoy, 2006; Komma et al., 2008; Pan et al., 2008; Turner et al., 2008).

In the stochastic perturbation approach it is common to perturb the model rainfall inputs based only on order of magnitude considerations. For example, Reichle et al. (2002) used additive perturbations from a Gaussian distribution, with standard deviation equal to 50% of the rainfall total at each model time step. Given that uncertainty in hydrological simulations directly depends on adequate characterization of input error (e.g., Crow and van Loon, 2006; Götzinger and Bárdossy, 2008), detailed analysis of the observed error of rainfall inputs is a critical research priority.

This paper directly evaluates rainfall error models that are commonly used in rainfall-runoff model calibration and prediction. Research is focused on the 50 km^2 Mahurangi catchment in Northland, New Zealand, where there is detailed space-time information on rainfall from both a dense gauge network (13 stations) and radar rainfall estimates, which provide an unprecedented opportunity to evaluate basic statistical rainfall error models. Our main focus is on understanding uncertainties in raingauge network measurements, since they remain the most common form of model input data. We also provide a comparative analysis based on available high-resolution radar fields, to enhance our understanding of the spatial/temporal rainfall variability and its effects on rainfall uncertainty at both the distributed and point scale, in space and in time. We aim to provide practical guidance on what error models and parameterisations might be appropriate at different temporal and spatial resolutions, and hence to provide the independent information necessary to develop ongoing work on quantifying the impact of errors and uncertainties in rainfall on the quality of calibrated parameters and/or on streamflow estimates (Carpenter and Georgakakos, 2004, Bardossy and Das, 2008, Thyer et al, 2009; Moulin et al, 2009; Kavetski et al 2006b Ajami et al., 2007; Vrugt et al, 2003b).

2. Sources of rainfall input uncertainty in hydrological models

When using raingauge data, as is currently common in hydrological modelling, a major source of uncertainty in streamflow simulations arises from poor representativeness of a discrete set of gauges of the precipitation over the entire catchment (Refsgaard et al., 2006; Villarini et al., 2008, Moulin et al., 2009), and from the assumptions used to interpolate rain rates between these gauges. The commonly used tipping bucket raingauges are themselves subject to both systematic and random errors, with mechanical limitations, wind effects and evaporation losses (Molini et al. 2001, la Barbera et al., 2002; Shedekar et al, 2009).

At the other end of the spectrum, radars offer the potential of providing integrated rainfall estimates over large spatial areas, though several challenges remain in interpreting the raw radar data into quantitative rainfall intensities (Moulin et al., 2009). Weather radar coverage has dramatically increased over the last few decades, giving access to measurements at high spatial and temporal resolutions (Moulin et al., 2009). Although signal treatment methods have significantly improved (Krajewski and Smith, 2002; Chapon et al., 2008), quantitative precipitation estimates still present difficulties.

It is increasingly recognised that uncertainty in rainfall has a critical effect on the accuracy of model predictions, and that efforts to advance scientific understanding through interrogating model parameters and structural hypotheses against data are hampered by errors and incorrect assumptions regarding the quality of the rainfall used to drive the model. Reichert and Mieleitner (2009) recently showed that allowing time dependency in rainfall bias improved model performance more than inclusion of any other time dependent parameter. Kavetski et al. (2006a) note that despite advances in data collection and model construction, remaining issues and the high spatial and temporal variability of precipitation input make it probable that rainfall input uncertainty will remain considerable in the foreseeable future.

3. Error models for Rainfall measurements

All calibration methods are based on hypotheses and assumptions, either explicit or implicit, describing how errors arise and propagate through a hydrological system (Kavetski et al., 2002). In traditional calibration, such as standard least squares (SLS) and equivalent methods based on the Nash-Sutcliffe optimization, input error is assumed negligible and the model and response errors are represented by a pseudo-additive random process (Kavetski et al. 2002; Kuczera et al, 2006). In the last two decades, increasing research effort has been devoted to moving towards more robust, integrated frameworks for separating and treating all sources of uncertainty (Liu and Gupta, 2007).

Beven and Binley (1992) introduced the generalized likelihood uncertainty estimation (GLUE) methodology for model calibration that takes into account the effects of uncertainty associated with the model structure and parameters. However, uncertainties associated with input data and output data (i.e., data errors) are not explicitly considered. Thiemann et al. (2001) introduced the Bayesian recursive parameter estimation (BaRE) methodology that poses the parameter estimation problem within the context of a formal Bayesian framework. BaRE explicitly considers the uncertainties associated with model-parameter selection and output measurements, but input data uncertainty and model structural uncertainty are not specifically separated out and are only implicitly considered, by expanding the predictive uncertainty bounds in a somewhat subjective manner (Liu and Gupta, 2007). A variety of other frameworks that have moved the science forward in recent years include the Shuffled Complex Evolution Metropolis algorithm (SCEM) and extensions (Vrugt et al., 2003a,b), the DYNamic Identifiability Analysis framework (DYNIA) (Wagener et al., 2003), the maximum likelihood Bayesian averaging method (MLBMA) (Neuman, 2003), the dual state-parameter estimation methods (Moradkhani et al., 2005a, 2005b), and the Simultaneous Optimization and Data assimilation algorithm (SODA) (Vrugt et al., 2005). However, these methods do not address all three critical aspects of uncertainty analysis (input error, structural error and output error) in a comprehensive, explicit and cohesive way (Liu and Gupta, 2007).

Despite the challenges in dealing with multiple sources of uncertainty, several important developments have taken place in the last decade. In particular, Kavetski et al. (2002, 2006a, 2006b) introduced the Bayesian total error analysis (BATEA) methodology, which explicitly treats input and output uncertainty, as well as structural errors (Kuczera et al 2006), within a Bayesian framework. BATEA allows the modeller to specify error models for all sources of uncertainty and integrates these models into the posterior inference of model parameters and predictions. Similarly, Ajami et al. (2007) introduced the Integrated Bayesian Uncertainty Estimator (IBUNE), which combines a probabilistic parameter estimator algorithm and Bayesian model combination techniques to provide an integrated assessment of uncertainty propagation within a system. If successful, representing and separating individual sources of error would represent a significant advance in environmental uncertainty analysis.

In current applications of BATEA and IBUNE, input errors have been assumed to be multiplicative and independent: while both frameworks are model-independent, there is currently little understanding of what an appropriate rainfall error model should look like. Other error models, such as *additive* Gaussian errors, have also been used (e.g. Huard and Mailhot, 2006). Almost all rainfall error models to date are empirical in nature, partly due to computational restrictions. However, Moulin et al. (2009) propose a notable exception, developing and calibrating an error model for hourly precipitation rates combining geostatistical tools based on kriging and an autoregressive model to account for temporal dependence of errors. Finally, in order to ensure statistical and computational well-posedness of the inference, typical applications apply the multiplicative assumption either to entire storm events (preserving the pattern but allowing for depth errors, Kavetski et al (2006b)), or to individual days with high leverage on model predictions (determined using sensitivity analysis, Thyer et al. (2009)).

Fundamentally, progressive disaggregation of individual sources of uncertainty requires more detailed probabilistic models describing the uncertainty in each data source, increasing the parameterization of the inference problem. Meaningful development and application of these hypotheses necessarily require reliable quantitative a priori information. In the absence of such knowledge, unsupported assumptions may be made, undermining the integrity of the inference. In particular, input error is likely to interact with model structural error, making posterior distributions of rainfall model-dependent, as well as affecting the inference of model parameters themselves (Beven, 2004, Balin et al., 2007). The critical significance of developing accurate prior knowledge of rainfall uncertainty is stressed in the findings of Renard et al (2009).

It is important that the data error models should be developed using data analysis that is independent from the hydrological model calibration, to bring genuine independent information into the inference (Renard et al, 2009). This paper is an early step in this direction, where we test the common multiplicative rainfall error model and comment on its suitability for use at varying spatial and temporal scales of the hydrological model.

4. Site and campaign description

The Mahurangi catchment is located in the North Island of New Zealand (Figure 1a). The Mahurangi River drains 50 km^2 of steep hills and gently rolling lowlands; catchment elevations range from 250m above sea level on the northern and southern boundaries, to near sea level at Warkworth on the east coast. Approximately half of the catchment (the central lowlands) is planted in pasture; one quarter of the catchment is in plantation forestry; and one quarter native forest. The catchment's soils have developed over Waitemata sandstones, which typically display alternating layers of sandstone and siltstone. Most soils in the catchment are clay loams, no more than a metre deep; clay and silt loam soils are also present in some parts of the catchment.

The climate is generally warm and humid, with mean annual rainfall of 1628mm and mean annual pan evaporation of 1315mm. Frosts are rare, and snow and ice are unknown. In late summer (February and March), remnants of tropical cyclones occasionally pass over northern New Zealand, producing intense bursts of rain. Convective activity is significant over the summer, whereas the majority of the winter rain comes from frontal systems. Maximum rainfall is usually in July, (the middle of the austral winter) while maximum monthly temperature and pan evaporation occur in January or February.

The catchment was extensively instrumented during the period 1997-2001 (refer to Woods et al., 2001 for further details): data from 29 nested stream gauges and 13 raingauges was complemented by measurements of soil moisture, evaporation and tracer experiments. We describe here only the rainfall data collection. The location of the 13 raingauges is shown in Figure 1b; rainfall depths were measured every two minutes using standard 200 mm collectors and 0.2mm tipping buckets. To augment the rainfall observations, the Physics Department of the University of Auckland deployed a mobile X-band radar for intensive campaigns of 1-2 months duration. This radar was sited in the southwest corner of the catchment and resolves rainfall on a 150m grid, every 5 seconds (typically amalgamated to two-minute average values), for the whole of the catchment.

Figure 1. A. Mahurangi Catchment Location Map.   B. Instrumentation Locations

5. Analysis

Our investigation of the rainfall data available in the Mahurangi catchment is divided into two main themes.

The first part analyses the spatial variability and uncertainty in rainfall when considered solely as a binary (two-state) process. In other words, what is the consistency of wet and dry periods over the catchment? This type of analysis is a crucial check on the assumptions underlying multiplicative rainfall error models, since the latter cannot account for rain events with only partial catchment cover that are hence not recorded by a rain gauge.

The second part examines the consistency of rainfall quantities over the catchment; based both on complete rainfall records and also for individual storm events when correct estimation of rainfall is most crucial. This allows us to estimate the statistical distributions of rainfall multipliers and test if these could form the basis of an adequate rainfall error model. In addition, analysis of rainfall profiles at a given gauge during individual storm events is used to put multipliers into the context of an event and understand interactions between temporal and spatial variability within a storm.

5.1 Rainfall State

Raingauges

Records from the 13 raingauges in the Mahurangi catchment are available as a complete record from 1997-2001. Although some of the records are available at shorter timesteps, all are aggregated to 15 minute intervals: typical of the timestep used in a high resolution hydrological model and designed to be sufficiently long to reduce the influence of random instrument/sampling errors which might otherwise be difficult to distinguish from true spatial variation in rainfall. The analyses are also repeated with rainfall aggregated to 1 hourly measurements, in order to add insight into changes in rainfall variability and hence suitable error formulations according to model timestep.

Our initial hypothesis was that the consistency of rainfall over the catchment is related to the severity of the storm event; i.e. that drizzle or low intensity rainfall might be patchy across the catchment but that heavy rainfall was more likely to be part of an extended weather system covering the complete catchment. The exception might be convective rainfall which could produce intense but localised showers.

To test this hypothesis, rainfall at each timestep over a 4.1 year period was tallied according to mean rainfall intensity (taken over all gauges in the catchment) and number of gauges recording rainfall. These results are plotted in Figure 2 below. The analysis shows that, using the hourly data (Figure 2A), where mean rainfall intensity recorded by the gauges is greater than 1 mm / hour, rainfall occurs at least 12 of the 13 gauges at least 94% of the time. For intensities 1 mm/hour or greater, the consistency of rainfall across the catchment therefore suggests that a multiplicative error model for rainfall would be suitable regardless of the location of the raingauge in the catchment. Comparing the CDF plots for 15-minute data and hourly data, it can be seen that variability in wet/dry states across the catchment becomes more pronounced at shorter time scales- at rainfall intensities greater than 1mm /hour, rainfall is only captured at 12 or 13 gauges 70% of the time and therefore the threshold for suitability of the multiplicative error model would need to be adjusted accordingly. The choice of threshold would be highly dependent on the required level of accuracy- even at 1.5 mm/hour only 81% capture at 12 or more gauges is obtained, by 2 mm/hour capture reaches 86% and 90% capture does not occur until intensities pass 2.6 mm/hour.

Figure 2: Consistency of rainfall state across raingauges for different critical rainfall intensities, at (A) Hourly and (B) 15-minute timesteps

Radar

Radar data provides much more detail about the spatial variation in rainfall state (wet/dry) as it is resolved on a 150 m grid, but the data is only available for specific field campaign periods (approximately 28 days total during the 4-year rainfall measurement period). Hence radar data is used to provide supplementary evidence to compare with conclusions drawn from the raingauge data analysis. The time periods used for this investigation were the 5 storm events captured with the radar, as follows: 7/8/98-12/8/98, 26/8/98-28/8/98, 15/9/98-18/9/98, 3/11/99-4/11/99, 9/11/99-13/11/99. Rainfall intensities measured by the radar were aggregated to timesteps of 15 minutes, 1 hour and 24 hours, for comparison.

As with the raingauge data, the first analysis was intended to investigate the consistency of rainfall estimates over the catchment. Rainfall consistency can be quantified more exactly using radar data than with raingauges, as the proportion of the catchment under rainfall. Figure 3 below shows simple histograms of this proportion, for the three different time steps.

Figure 3: Distributions of percentage of catchment under rainfall at different time steps, for five storm events

As before, we also quantify the relationship between mean intensity of rainfall and consistency of rainfall, by plotting average catchment rainfall against fraction of catchment under rainfall, using the radar data (Figure 4).

Figure 4: Consistency of rainfall state across radar pixels for different critical rainfall intensities, at (A) Hourly and (B) 15-minute timesteps

The results show that, at a 15 minute timestep, there are a significant number of periods during storm events where rainfall is scattered over the catchment (for 35% of timesteps, between 10% and 90% of the catchment is under rainfall). As expected, at an hourly timestep, rainfall appears more uniform in time (42% of timesteps with less than 10% rainfall, 31% of timesteps between 10% and 90% rainfall and 27% of timesteps with greater than 90% rainfall). Finally, at a daily timestep, it was unusual to find significant dry areas of the catchment during a storm day (Figure 3).

Corroborating these observations, the cumulative plots of fraction of catchment under rainfall against intensity (Figure 4) show similar results. At a 15 minute timestep (Figure 4B), rainfall is not consistent across the catchment when mean intensity is below 1 mm/hour, suggesting that a multiplicative rainfall error model would not be suitable. For an hourly timestep this critical intensity is reduced to approximately 0.4 mm/hour (Figure 4B) and for a daily timestep an intensity criterion would not be necessary (not shown). For rainfall with intensity above these thresholds, a multiplicative rainfall error would be suitable. This finding confirms that the approach taken in current applications of the BATEA model calibration strategy (Thyer et al., 2009), with rainfall multipliers applied to those days where the model is most sensitive to rainfall uncertainty (which are generally high-rainfall days), is consistent with observed rainfall spatial variability.

5.2 Rainfall Quantity

Distributions of rainfall totals

Where rainfall occurrence is consistent across the catchment and hence suitable for representation by a multiplicative error model, the distribution (type and parameters) of multipliers must be specified. Previous studies have in general specified multipliers as arising from a lognormal distribution with zero mean (Kavetski *et al.*, 2006a;b), or unknown mean (Thyer et al., 2009). The data at Mahurangi allows us to directly test this hypothesis.

For each of the 13 raingauges, and for each storm timestep (defined as mean catchment rainfall greater than 0.2mm/hour and at least 6 gauges recording rainfall), the corrective multiplier was calculated in order to transform rainfall measured at the gauge into the mean rainfall over the 13 gauges. These distributions are plotted in Figure 5 below, for both 15 minute and hourly timesteps, along with the mean and standard deviation of the empirical multiplier distribution, and the unbiased (zero) line for comparison. An additional plot is shown with all distributions overlaid for comparison.

Figure 5: Multipliers required to convert raingauge readings to catchment mean rainfall, at (A) Hourly and (B) 15-minute timesteps

Following the same methodology as for the raingauges, the radar data analysis was repeated using the 28 days, at both 15 minute and hourly timescales. Gauged records were approximated using the radar pixel closed to the gauge location, and multipliers were calculated to transform this value to the catchment mean calculated from the complete radar data set (not just gauge locations). Results are shown in Figures 6A and 6B.

Figure 6: Multipliers required to convert radar readings at raingauge locations to catchment mean rainfall measured by radar, at (A) Hourly and (B) 15-minute timesteps

The log multipliers calculated using both 15-minute data and hourly data were also summarised using normal quantile-quantile plots (Figure 7A;B), and via the Lilliefors test to test for normality. The normality hypothesis was rejected for all 13 sites using the raingauge data, and for 12 of the 13 sites using the radar data (the exception being Upper Goatley). The qq-plots show the fat negative tail and excessive kurtosis causing this result, and show that a lognormal distribution does not fully capture the empirical multiplier distribution.

Figure 7: QQ-Plot of Multiplier Distribution against Normal Distribution for raingauge and radar data at (A) Hourly and (B) 15-minute timestep

When multiplicative error models for rainfall are used in hydrological modelling applications, the assumptions are typically made that rainfall multipliers are uncorrelated in time and have an invariant distribution (Kavetski 2006a;b). Autocorrelations of the empirical multiplier series were calculated: at an hourly timestep the maximum autocorrelation occurs at lag 1, with a mean value of 0.15 over all gauges, and decreases rapidly for higher lags. At a 15-minute timestep the value is increased to 0.34. Therefore for models operating at a 15-minute timestep where multipliers are applied to consecutive timesteps, an error model including an autocorrelation term would improve the representation of errors; however for models at an hourly timestep or where multipliers are applied only to selected heavy-rainfall timesteps, autocorrelation would be less important.

To test for invariance in multiplier distribution, comparisons were made amongst storm timesteps of different rainfall depth (Figure 8A) and according to the season during which the rain fell (Figure 8B). The figures show that the multiplier distribution does not vary significantly with season, although there is slightly more variation between gauges during the wetter seasons (winter and spring) than in the drier seasons (summer and autumn). The depth of rainfall does however affect the multiplier distribution: during light rain the distribution is close to lognormal; but during heavy rainfall multipliers have increased skew and the distribution has heavier tails with more outliers. Although it is not unexpected that heavy rain events show more variation over the catchment (e.g. during convective rain), this result shows that high multiplier values are caused by true variation in rainfall processes and not by noise (i.e. sampling and instrument error) obscuring the signal at low rainfall depths.

Figure 8: QQ-Plot of Multiplier Distribution against Normal Distribution at hourly timestep, compared by (A) Rainfall Depth and (A) Season

Raingauge versus radar data

Comparing Figures 7A and 7B significant differences are shown between using raingauge data (improved temporal coverage; reduced spatial coverage) versus radar data (shorter temporal coverage, improved spatial coverage). In particular, the skew is more pronounced in the radar data than in the raingauge measurements, especially at 15-minute timescale. In addition, the radar data shows noticeable changes in error distribution between gauges, unlike the raingauge data, where the errors tended to follow a more consistent shape.

It is likely that the more complete sampling of catchment rainfall provided by the radar, including the more inaccessible hillcountry and bush areas where raingauges are more difficult to site, is in part responsible for these differences. However, the susceptibility of radar data to transient errors from sources such as measuring the field at some distance above the ground and recording the reflectivity data with a limited radiometric resolution (e.g. Fabry *et al*., 1994; Nicol and Austin, 2003; Jordan *et al*., 2000) may also play a role.

Another important distinction is that the radar data is concentrated during storm periods, and multiplier distributions are more skewed during large storms (Figure 8A). For example, if only timesteps where the raingauge reading is greater than 2 mm per 15 mins are considered, skewness typically triples (although this effect may be partly due to the resulting change in sample size). The result is particularly pronounced in atypical areas of the catchment, such as Moirs Hill, which lies in the hills in the South-West of the catchment and consistently records higher than average rainfall in response to orographic rainfall effects. Increased skewness during storm events must be accounted for if rainfall multipliers are to be applied to high-rainfall days only (as in the BATEA case study of Thyer et al, 2009), and suggests that a skew distribution together with careful siting of the raingauge would be necessary to capture the multiplier distribution even at a relatively small basin such as the Mahurangi.

It is also interesting to note the differences in the representativeness of individual gauges. For example, the Toovey Road gauge is highly representative of the catchment mean rainfall, showing a strongly peaked distribution with mean close to zero, while other sites such as Falls Road are less representative for both raingauge and radar analyses. This observation echoes previous studies into representative areas of a catchment for given variables such as soil moisture, designed to reduce gauging requirement requirements, e.g. Martinez-Fernandez and Ceballos (2005), Vachaud *et al.* (1985).

Consistency in rainfall profiles

In response to the differences between raingauge and radar data, highlighted in the previous section, further investigation was carried out into the spatial and temporal variation in rainfall volumes. In particular, this analysis may provide insights into the increased variability in multiplier distributions in the radar data, not seen in the raingauge totals.

Firstly, spatial variation in rainfall totals was investigated. Hourly and 15-minute rainfall profiles were extracted for all raingauges, for the nine largest storms over the study period (Figure 9A;B).

Figure 9: Comparison of storm profiles between raingauges, at (A) Hourly and (B) 15-minute timesteps

These profiles show remarkably little variation in storm profile between raingauges, with peaks generally occurring at all gauges within a timespan of 1 hour. This result may be due to an underlying lack of variability in rainfall or may rather be caused by the averaging effect of recording raingauge totals at hourly intervals, which could hide rainfall variability at short temporal scales. One storm where both 2 min and 15 min data are available from the radar is investigated in more detail: Figure 10A below shows the storm profile at both 2 min and 15 min timesteps, for 3 of the raingauge locations. In addition, the multipliers from each of those gauges to the mean catchment rainfall are calculated, again at both 2 min and 15 min timesteps, and shown in Figure 10B.

Secondly, the ability to capure peak rainfalls was examined. The peaks of rainfall intensity are significantly damped (Figure 10A; max intensity reduced by more than 60% for some peaks) and this effect would lead to important changes in the tails of the multiplier distributions. Multiplier distributions must therefore be considered dependent on timestep length. The fluctuations caused by rain cells tracking over the catchment occur in the 2 min data, but much of the variation is lost in the 15 min data. The time for a typical rain cell to travel across the catchment has been estimated as close to 15 minutes, assuming raincell size of 1-5 km, a speed of the order of 10 ms^{-1} and a catchment width of 5 – 10 km (Woods *et al*, 2001). We conclude that in this case, consistency between raingauges at 15 minutes is due mainly to the averaging effect seen at timescales longer than the time taken for a raincell to traverse the catchment. When using spatial maps of radar data to calculate multipliers, as oppose to raingauges, more of this variation may be captured as areas both directly under the track of the raincell, and those on the edge of the rain area, are fully sampled.

Figure 10: Comparison of (A) storm profile and (B) log multipliers calculated at 2 minute and 15 minute timesteps for the rainstorm of 10-11 Aug 1998, for 3 raingauge sites.

6. Effects of Rainfall Variability on Rainfall-Runoff Processes

Additional knowledge of spatial rainfall variability has the potential not only to inform statistical models of rainfall uncertainty, but also to change our understanding of the catchment processes and hence our conceptual model of the catchment. The dense monitoring network at Mahurangi draws to our attention the intense bursts of rainfall at short space or time scales (as demonstrated by the long tails of the multiplier distributions (Figures 5;6) and increased variability in multipliers at the 2 minute timescale (Figure 10)) which would not be apparent when using catchment average data at typical model timescales. A range of 'fast-response' processes such as infiltration excess flow, transient overland flow or macropore flow might therefore be under-represented in models where catchment average rainfall is used; or cause compensatory effects in model calibration to correct the bias.

Figure 11 below shows a simplified example where the Mahurangi catchment is classified into two soil texture zones (loam vs. clay; Figure 11A). We then plot (Figure 11B) the mean catchment rainfall against the percentage of the catchment where radar-measured rainfall intensity exceeds the estimated saturated hydraulic conductivity of the soil (estimated using the Clapp and Hornberger (1978) soil parameters for the two soil texture zones). Infiltration excess conditions in some areas can start to occur with average catchment rainfall as low as 1 mm/hour, although that average rate would be insufficient to activate this process. Areas under infiltration excess are likely to contribute an over-representative fraction of channel flow and therefore be essential in the rainfall-runoff mechanism. We conclude that while multiplicative error hypotheses may capture some of the rainfall error structure and enable point measurements of rainfall to be corrected to the catchment average, multipliers alone cannot capture the interaction of soil and land cover variability with rainfall variability. As the latter is progressively elucidated using spatially-distributed rainfall measurements, more complex error models could and should be derived.

Figure 11: (A) Catchment soil texture and (B) Estimated catchment area under infiltration excess conditions vs. average rainfall rate

7. Conclusions

Our investigation of rainfall variability within the dense gauge/radar network at Mahurangi has highlighted some important results for understanding rainfall uncertainty and deriving data-based probabilistic error models for use in hydrological calibration.

Our examination of the variability of wet/dry states over the catchment has confirmed that multiplicative error is a suitable formulation for correcting mean catchment rainfall values during high-rainfall periods (e.g. intensities over 1 mm/hour); or for longer timesteps at any rainfall intensity (timestep 1 day or greater). We suggest that the effect of timestep on multiplier suitability is regulated by catchment size: specifically the time required for typical raincells to cross the catchment could be used as a first estimate of critical timestep.

We found that the standard distribution used for rainfall multipliers, the lognormal, provides a relatively close fit to the empirical multiplier distributions. However the empirical distributions have greater excess kurtosis and positive skew than the lognormal, and therefore alternative distributions should be considered where the tails of the multiplier distribution are considered particularly important. Since heavy rainfall events display multiplier distributions differing most significantly from the lognormal, a skewed and heavier-tailed distribution to be used for times of high rainfall would more faithfully reproduce the observed error characteristics. We found that the error distributions do not vary significantly with season and hence an invariant distribution is sufficient.

Lastly, the high resolution of the data available demonstrated the time/space complexity of rainfall behaviour that cannot be corrected by a simple multiplicative error on measured rainfall. A hydrological model that aims to capture the full effects of rainfall variability (and its interaction with topographical and ecological variability) would therefore need an additional mechanism, such as a distribution function approach to rainfall input; or a blurred threshold for processes such as infiltration excess.

Acknowledgements Financial support for this project was provided by the New Zealand Foundation for Research Science and Technology grant CO1X0812. The authors wish to thank all participants of the MARVEX campaign and particularly Prof. Geoff Austin of the University of Auckland for use of the X-band radar data.

References

Ajami, N. K., Duan, Q. and Sorooshian, S., 2007. An Integrated Hydrologic Bayesian Multi-Model Combination Framework: Confronting input, parameter and model structural uncertainty in Hydrologic Prediction, *Water Resour Res*, **43**, W01403,2 doi:10.1029/2005WR004745.

Balin, D, Shbaita, H., Lee, H. and Rode, M., 2007. Extended Bayesian uncertainty analysis for distributed rainfall-runoff modelling: application to a small lower mountain range catchment in central Germany, Proc CAIWA Conference on "Coping with Complexity and Uncertainty", Basel, Switzerland, November 2007.

Bárdossy, A. and Das, T., 2008. Influence of rainfall observation network on model calibration and application, *Hydrol. Earth Syst. Sci*., **12**, 77–89, 2008

Beven, K., 2004. *Rainfall-runoff modelling – the primer*. John Wiley & Sons Ltd, New York.

Beven, K and Binley, A. M., 1992. The future of distributed models: model calibration and uncertainty prediction, *Hydrological Processes* **6**:279-298.

Carpenter, T. M., and K. P. Georgakakos, 2004, Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed hydrologic model, *Journal of Hydrology*, **298**(1-4), 202-221.

Chapon, B., Delrieu, G., Gosset, M., and Boudevillain, B., 2008. Variability of rain drop size distribution and its effect on the Z-R relationship: A case study for intense Mediterranean rainfall, *Atmos.Res*., **87**, 52–65,

Clark, M. P., and A. G. Slater, 2006, Probabilistic quantitative precipitation estimation in complex terrain, *Journal of Hydrometeorology*, **7**(1), 3-22.

Clapp, R.B., Hornberger, G.M., 1978. Empirical equations for some soil hydraulic properties. *Wat. Resour. Res.* 14, 601-604.

Crow, W. T., and E. Van Loon, 2006, Impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture, *Journal of Hydrometeorology,* **7**(3), 421-432.

Fabry, F., Bellon, A., Ducan, MR, Austin GL, 1994. High-resolution rainfall measurements by radar for very small basins – the sampling problem re-examined. *Journal of Hydrology* **161**: 415-428.

Huard, D. and A. Mailhot, 2006, A Bayesian perspective on input uncertainty in model calibration: Application to hydrological model \u201cabc\u201d, *Water Resour. Res.*, **42**, W07416, doi:10.1029/2005WR004661.

Gotzinger, J., and A. Bardossy, 2008, Generic error model for calibration and uncertainty estimation of hydrological models, *Water Resour. Res.*, 44.

Jordan P, Seed A, Austin G, 2000. Sampling errors in radar estimates of rainfall. *Journal of Geophysical Research – Atmospheres.* **105**: 2247-2257.

Kavetski, D. Franks, S. W. and G. Kuczera, 2002. Confronting input uncertainty in environmental modelling. In: Duan, Q., Gupta, H.V., Sorooshian, S., Rousseau, A.N., Turcotte, R. (Eds.), Calibration of Watershed Models. Water Science and Application 6, American Geophysical Union. pp. 49-68.

Kavetski, D., Kuczera, G., Franks, S.W., 2006a. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. *Water Resour. Res*. **42** (3). doi:10.1029/2005WR00436.

Kavetski D., G. Kuczera, S. W. Franks, 2006b, Bayesian analysis of input uncertainty in hydrological modeling: 2. Application, *Water Resour. Res*., **42**, W03408, doi:10.1029/2005WR004376.

Komma, J., G. Bloschl, and C. Reszler, 2008, Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting, *J. Hydrol.*, **357**(3-4), 228-242.

Krajewski, W. F. and Smith, J., 2002 Radar hydrology : rainfall estimation., *Adv. Water Resour.*, **25**(8-12), 1387–1394.

Kuczera, G., Kavetski, D., Franks, S. and Thyer, M., 2006. Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters', *J. Hydrol*, **331** 161-177.

La Barbera, P., Lanza, L.G., Stagi, L., 2002. *Influence of systematic mechanical errors of tippingbucket rain gauges on the statistics of rainfall extremes*, *Water Sci. Techn*., **45**(2), 1-9.

Liu Y and Gupta HV., 2007, Uncertainty in Hydrologic Modeling: Towards an Integrated Data Assimilation Framework, *Water Resour. Res*.

Martinez-Fernandez, J; Ceballos, A, 2005. Mean soil moisture estimation using temporal stability analysis. *J. Hydrol.* **312**:28-38.

Molini, A., La Barbera, P., Lanza, L. G., Stagi, L., 2001. Rainfall intermittency and the sampling error of tipping-bucket rain gauges. *Phys. Chem. Earth (C)*, **26**(10-12), 737-742.

Molini, A., Lanza, L.G., and La Barbera, P., 2005. The impact of tipping-bucket raingauge measurement errors on design rainfall for urban-scale applications. *Hydrol. Proc.*, **19**, 1073-1088.

Moradkhani, H., S. Sorooshian, H. V. Gupta, and P. Houser, 2005a. Dual state parameter estimation of hydrologic models using ensemble Kalman filter, *Advances in Water Resources*, **28**, 135-147.

Moradkhani, H., K.-L. Hsu, H. V. Gupta, and S. Sorooshian, 2005b. Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, *Water Resour. Res.*, **41**, doi:10.1029/2004WR003604.

Moulin, L., Gaume, E., and Obled, C., 2009 Uncertainties on mean areal precipitation: assessment and impact on streamflow simulations, *Hydrol. Earth Syst. Sci.*, **13**, 99-114.

Nicol JC, Austin GL, 2003. Attenuation correction constraint for single-polarisation weather radar. *Meteorological Applications* **10**: 345-354.

Neuman, S. P., 2003, Maximum likelihood Bayesian averaging of uncertain model predictions, *Stochastic Environmental Research and Risk Assessment*, **17**, 291-305.

Pan, M., E. F. Wood, R. Wojcik, and M. F. McCabe, 2008, Estimation of regional terrestrial water cycle using multi-sensor remote sensing observations and data assimilation, *Remote Sensing of Environment*, **112**(4), 1282-1294.

Pauwels, V. R. N., and G. J. M. De Lannoy, 2006, Improvement of modeled soil wetness conditions and turbulent fluxes through the assimilation of observed discharge, *J. Hydrometeorol.*, **7**(3), 458-477.

Refsgaard, J. C., van der Sluijs, J. P., Brown, J., and van der Keur, P, 2006. A framework for dealing with uncertainty due to model structure error, *Adv. Water Resour.*, **29**, 1586–1597, 2006.

Reichert, P., and J. Mieleitner, 2009, Analyzing input and structural uncertainty of nonlinear dynamic models with stochastic, time-dependent parameters, *Water Resour. Res.*, **45**, W10402, doi:10.1029/2009WR007814

Reichle, R. H., J. P. Walker, R. D. Koster, and P. R. Houser, 2002, Extended versus ensemble Kalman filtering for land data assimilation, *J. Hydrometeorol.*, **3**(6), 728-740.

Renard, B. Kavetski, D., Kuczera, G., M. Thyer and S. W. Franks, 2009. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, *Wat Resour Res*, in revision.

Shedekar, V, Brown, L, Heckel, M, King, K, Fausey, N and Harmel, R, 2009. Measurement Errors in Tipping Bucket Rain Gauges under Different Rainfall Intensities and their implication to Hydrologic Models, Proc 7th World Congress of Computers in Agriculture and Natural Resources, available from American Society of Agricultural and Biological Engineers, St. Joseph, Michigan.

Sun, X., Mein, R. G., Keenan, T. D., and Elliott, J. F., 2000. Flood estimation using radar and raingauge data, *J. Hydrol*., **239**, 4–18

Thiemann, T., M. Trosset, H. Gupta, and S. Sorooshian, 2001. Bayesian recursive parameter estimation for hydrologic models, *Water Resour. Res.*, **37**(10), 2521-2535.

Thyer M, Renard B, Kavetski D, Kuczera G, Franks SW, Srikanthan S, 2009. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis, Water Resources Research, 45, W00B14, doi:10.1029/2008WR006825.

Turner, M. R. J., J. P. Walker, and P. R. Oke, 2008. Ensemble member generation for sequential data assimilation, Remote Sensing of Environment, **112**(4), 1421-1433.

Vachaud, G., A. Passerat de Silans, P. Balabanis and M. Vauclin, 1985. Temporal stability of spatially measured soil water probability density function, *Soil Sci. Soc. Am. J.* **49**, pp. 822–828.

Villarini, G., Mandapaka, P., Krajewski, W., and Moore, R. 2008: Rainfall and sampling uncertainties: A rain gauge perspective., *J. Geophys. Res*., **113**, D11102. 12, pp. 10.1029/2007JD009214

Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian, 2003a. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, **39**(8), 1201,doi:10.1029/2002WR001642.

Vrugt, J A., H. V. Gupta, L. A. Bastidas, W. Bouten, and S. Sorooshian, 2003b, Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.*, **39**(8), doi: 10.1029/2002WR001746.

Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten, 2005. Improved treatment of uncertainty in hydrologic modeling: Combining the strength of global optimization and data assimilation, *Water Resour. Res.*, **41**, W01017, doi:10.1029/2004WR003059.

Wagener, T., N. McIntyre, M. J. Lees, H. S. Wheater, and H. V. Gupta, 2003. Towards reduced uncertainty in conceptual rainfall-runoff modeling: Dynamic identifiability analysis, *Hydrol. Process.*, *17*, 455-476.

Woods RA, Grayson RB, Western AW, Duncan MJ, Wilson DJ, Young RI, Ibbitt RP, Henderson RD, McMahon TA, 2001. Experimental Design and Initial Results from the Mahurangi River Variability Experiment: MARVEX. In: Lakshmi, V.; Albertson, J.D.; Schaake, J. (eds). Observations and Modelling of Land Surface Hydrological Processes, pp. 201-213. Water Resources Monographs. American Geophysical Union, Washington, DC.

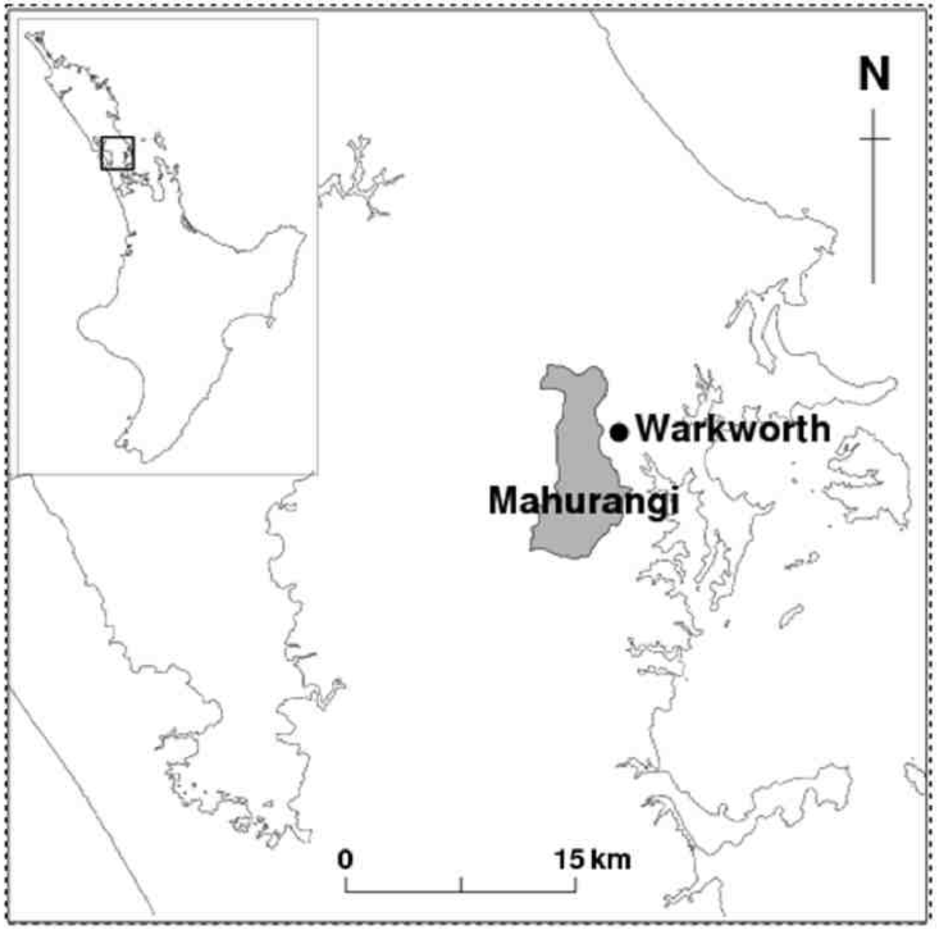

Figure 1a: Mahurangi Catchment Location Map

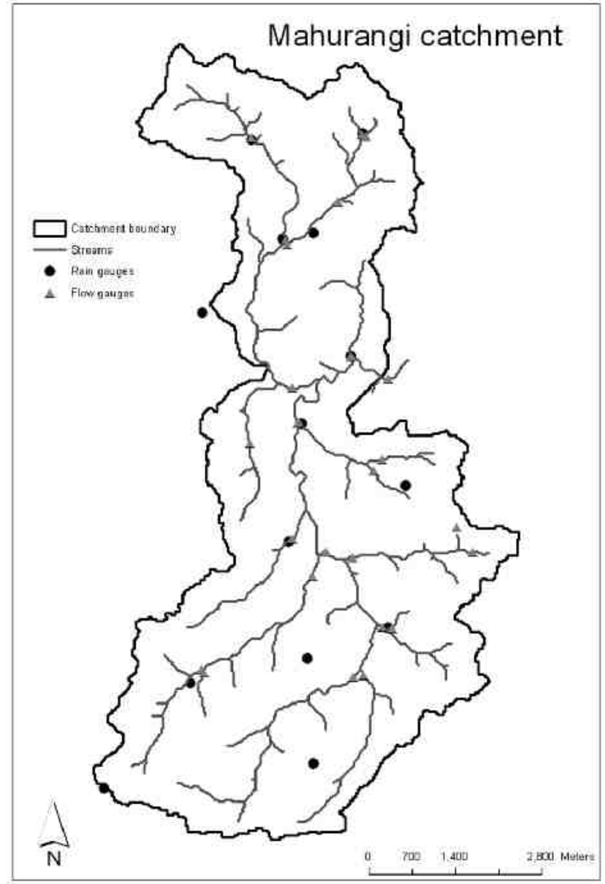

Figure 1b: Instrumentation Locations

A. Hourly timestep

B. 15 minute timestep

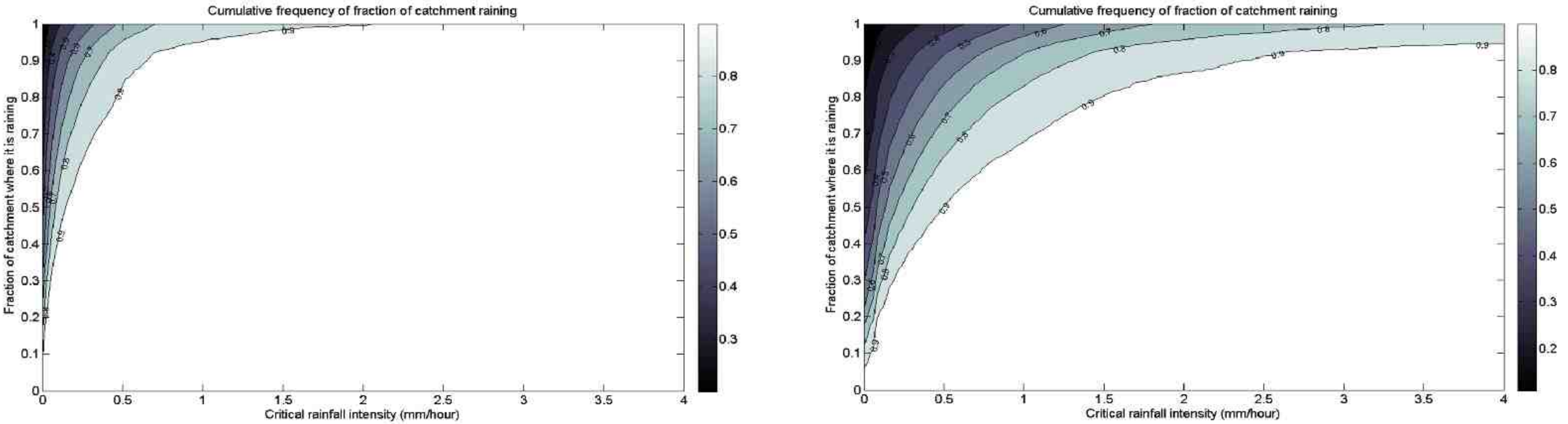

Figure 2: Consistency of rainfall state across raingauges for different critical rainfall intensities, at (A) Hourly and (B) 15-minute timesteps

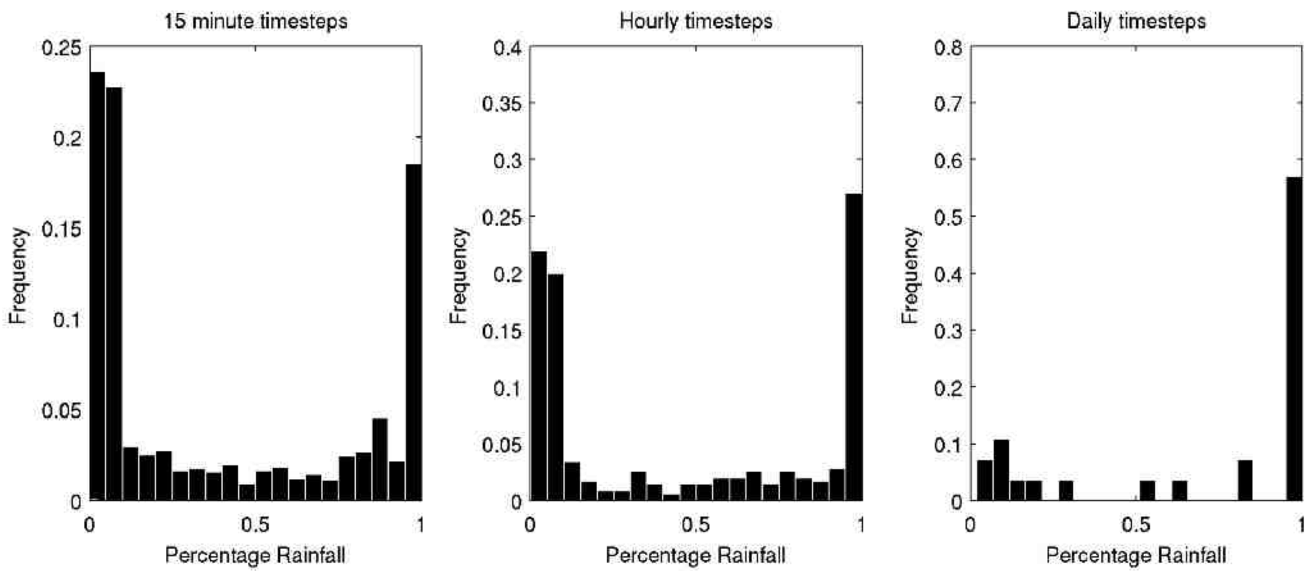

Figure 3: Distributions of percentage of catchment under rainfall at different time steps, for five storm events

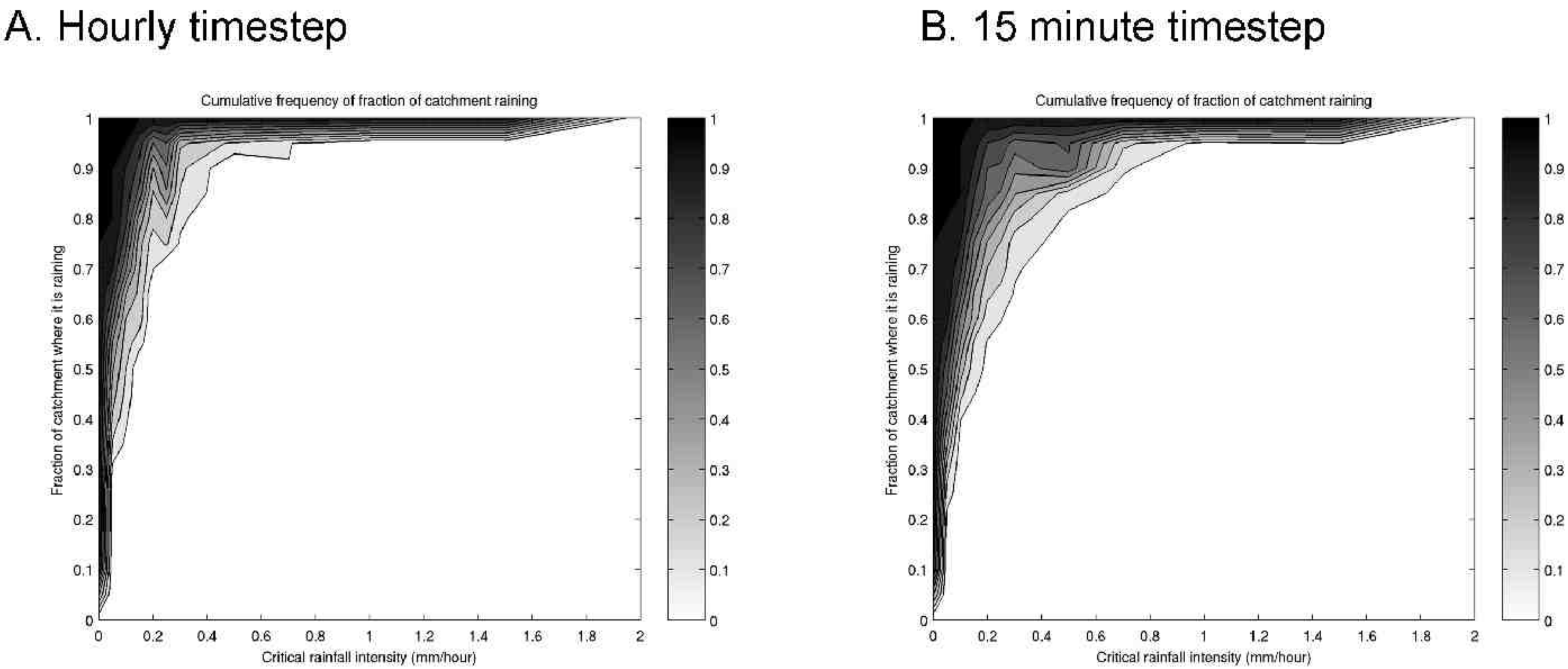

Figure 4: Consistency of rainfall state across radar pixels for different critical rainfall intensities, at (A) Hourly and (B) 15-minute timesteps

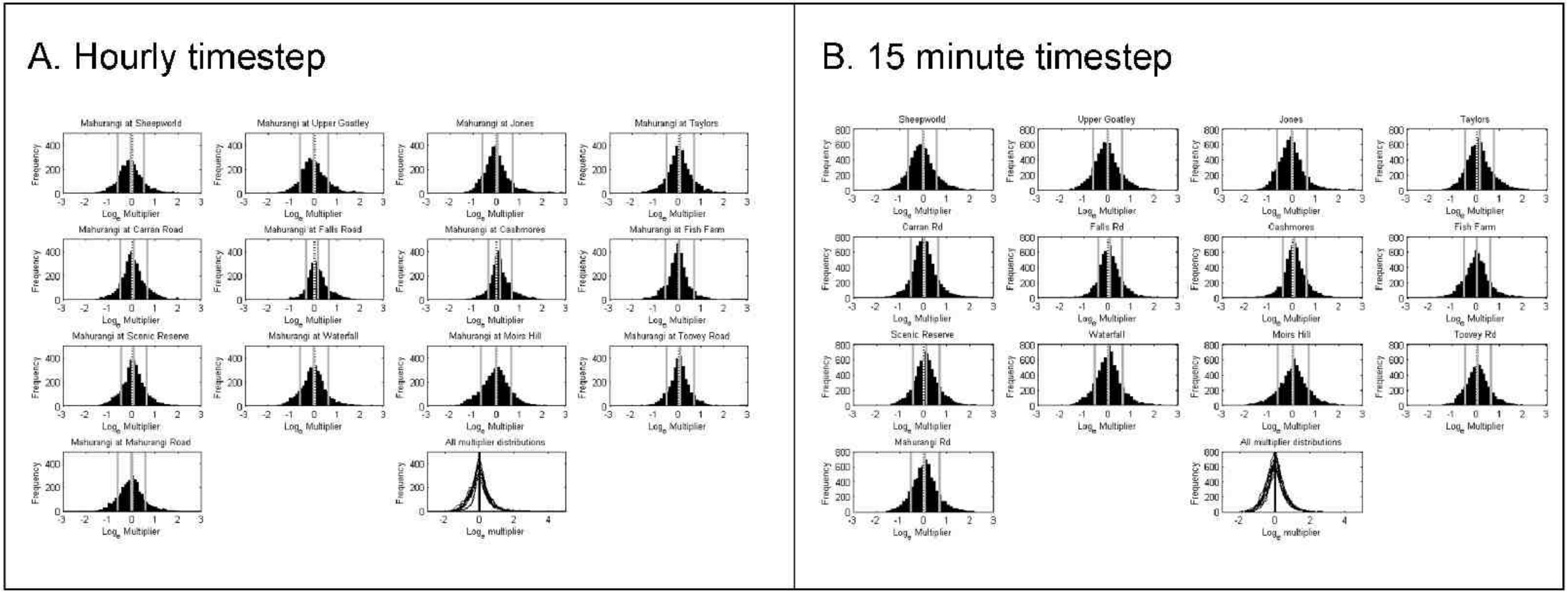

Figure 5: Multipliers required to convert raingauge readings to catchment mean rainfall, at (A) Hourly and (B) 15-minute timesteps

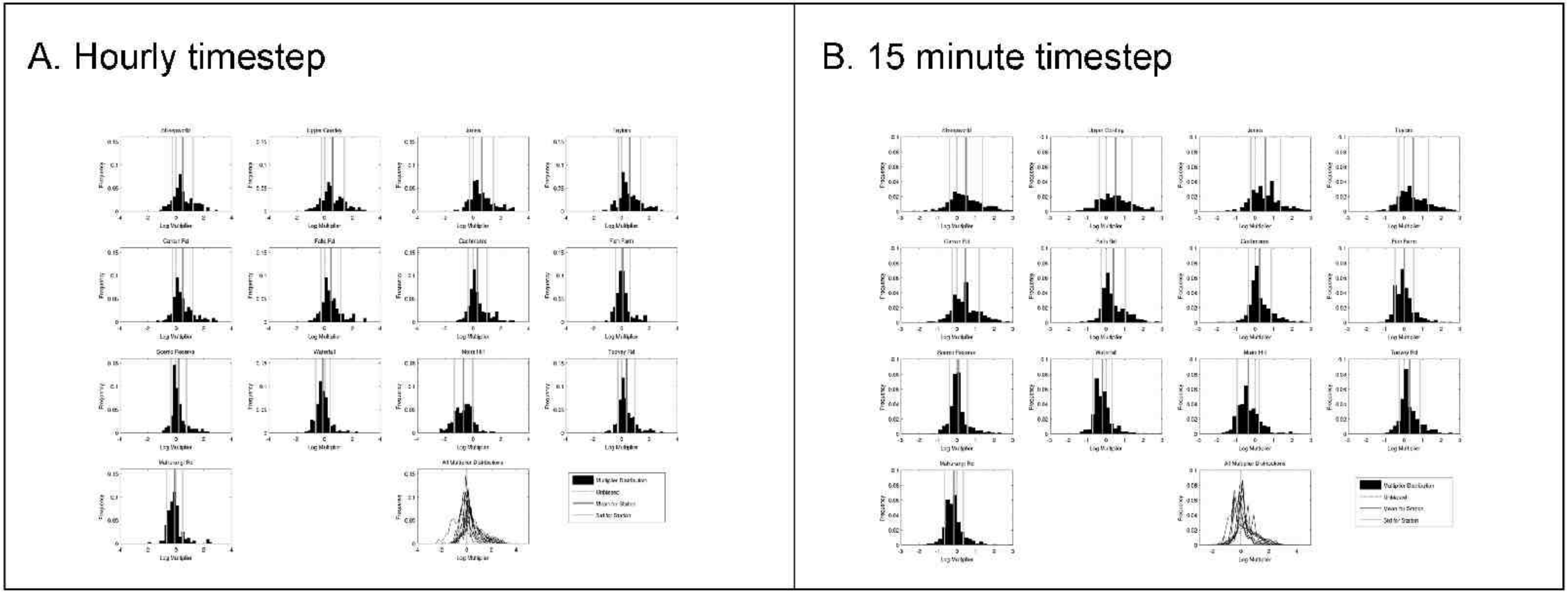

Figure 6: Multipliers required to convert radar readings at raingauge locations to catchment mean rainfall measured by radar, at (A) Hourly and (B) 15-minute timesteps

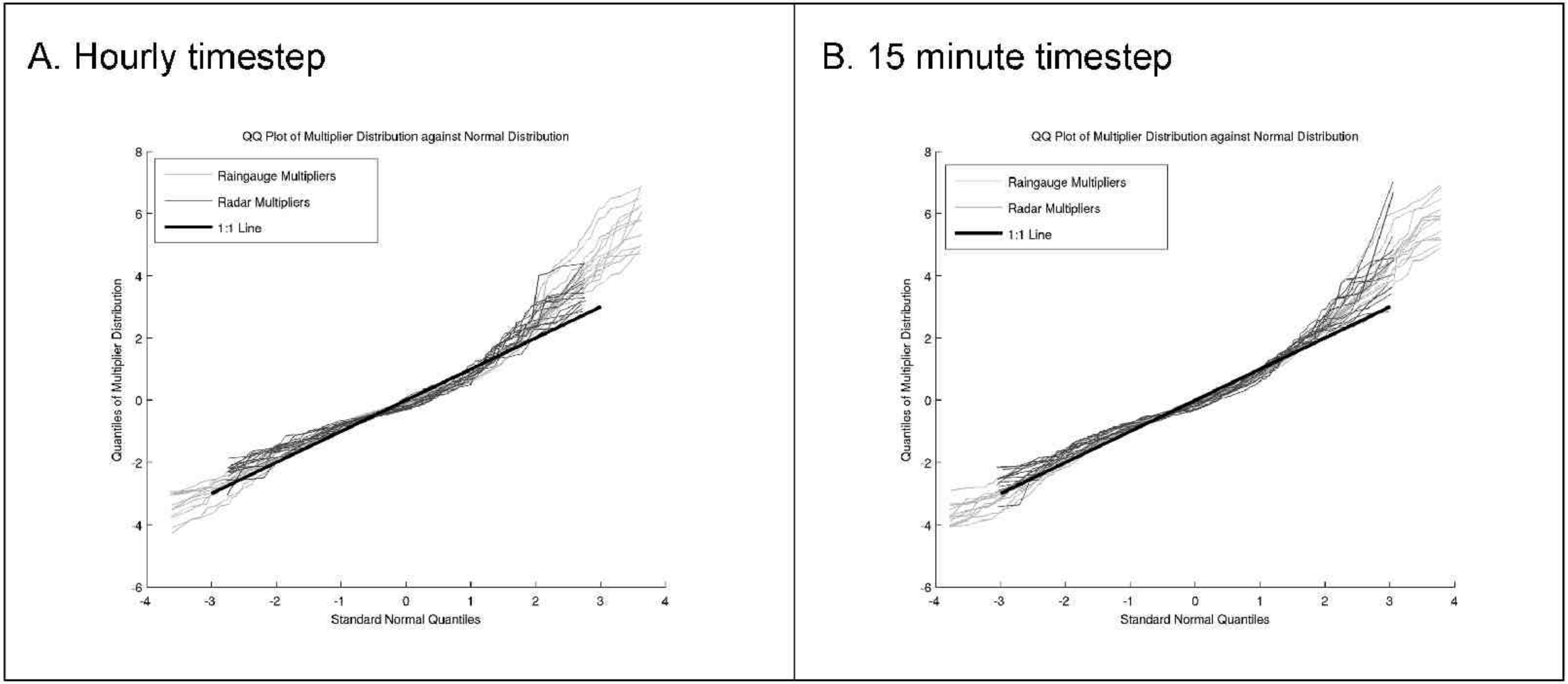

Figure 7: QQ-Plot of Multiplier Distribution against Normal Distribution for raingauge and radar data at (A) Hourly and (B) 15-minute timestep

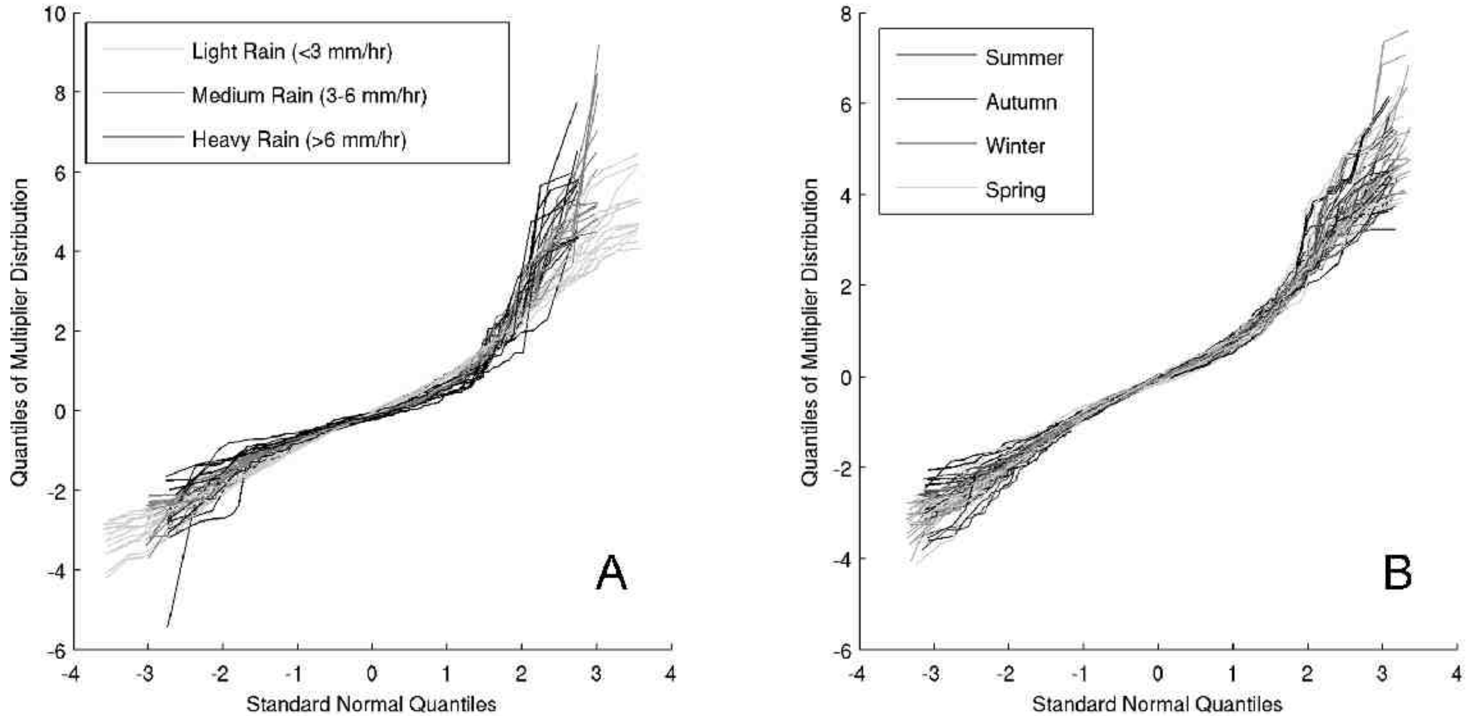

Figure 8: QQ-Plot of Multiplier Distribution against Normal Distribution at hourly timestep, compared by (A) Rainfall Depth and (A) Season

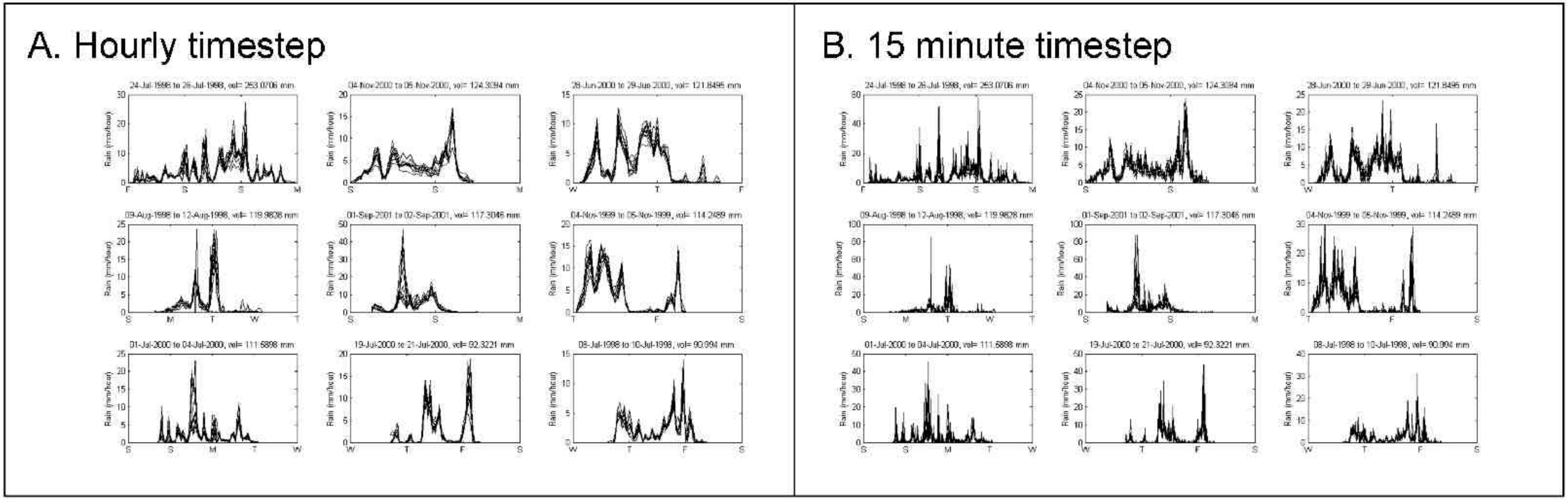

Figure 9: Comparison of storm profiles between raingauges, at (A) Hourly and (B) 15-minute timesteps

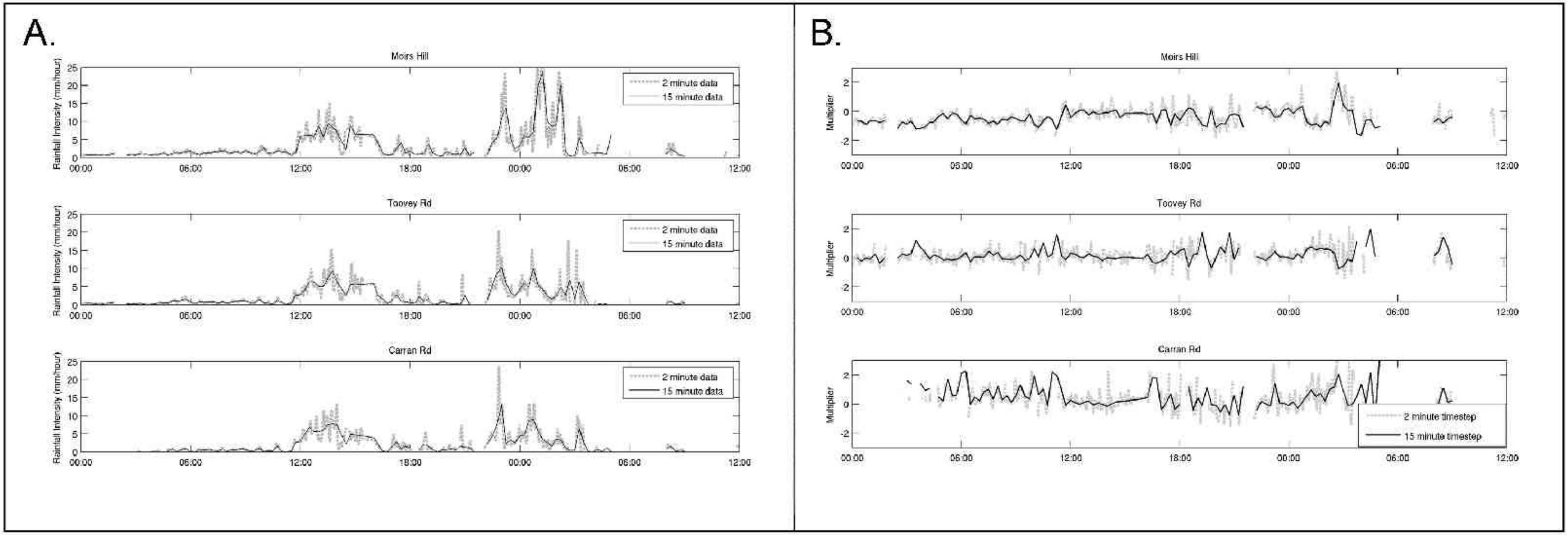

Figure 10: Comparison of (A) storm profile and (B) log multipliers calculated at 2 minute and 15 minute timesteps for the rainstorm of 10-11 Aug 1998, for 3 raingauge sites.

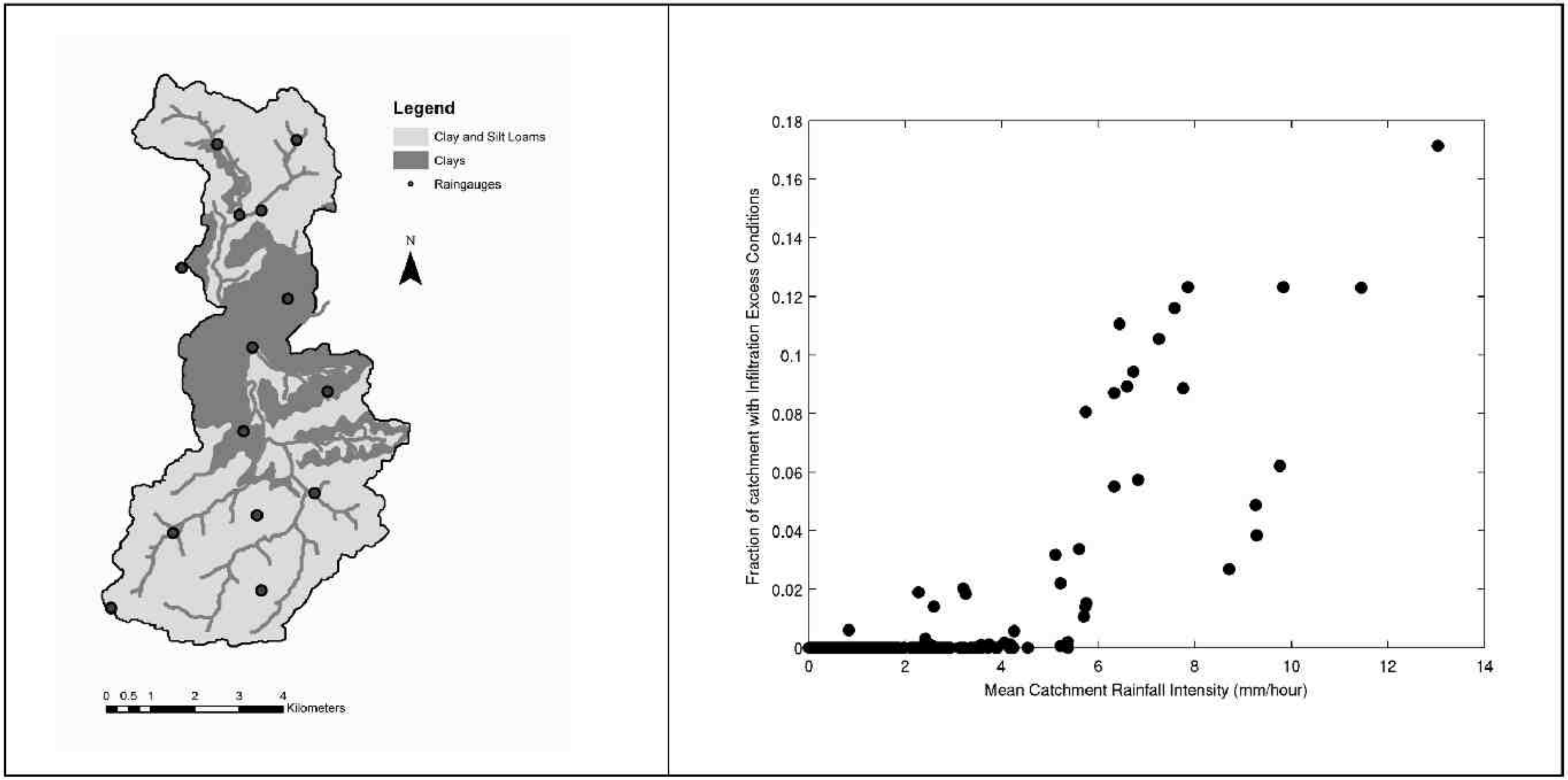

Figure 11: (A) Catchment soil texture and (B) Estimated catchment area under infiltration excess conditions vs. average rainfall rate