INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **33**: 746–757 (2013)
Published online 30 March 2012 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.3469

Quantifying parameter uncertainty and assessing the skill of exponential dispersion rainfall simulation models†

Andrew D. Gronewold,[a]* Craig A. Stow,[a] James L. Crooks[b] and Timothy S. Hunter,[a]
[a] *NOAA, Great Lakes Environmental Research Laboratory, Ann Arbor, MI, USA*
[b] *USEPA, National Exposure Research Laboratory, Research Triangle Park, NC, USA*

ABSTRACT: The exponential dispersion model (EDM) has been demonstrated as an effective tool for quantifying rainfall dynamics across monthly time scales by simultaneously modelling discrete and continuous variables in a single probability density function. Recent applications of the EDM have included development and implementation of statistical software packages for automatically conditioning model parameters on historical time series data. Here, we advance the application of the EDM through an analysis of rainfall records in the North American Laurentian Great Lakes by implementing the EDM in a Bayesian Markov chain Monte Carlo (MCMC) framework which explicitly acknowledges historic rainfall variability and reflects that variability through uncertainty and correlation in model parameters and simulated rainfall metrics. We find, through a novel probabilistic assessment of skill, that the EDM reproduces the magnitude, variability, and occurrence of daily rainfall, but does not fully capture temporal autocorrelation on a daily time scale. These findings have significant implications for the extent to which the EDM can serve as a tool for supporting regional climate assessments, for downscaling regional climate scenarios into local-scale rainfall time series simulations, and for assessing trends in the historical climate record. Published in 2012 by John Wiley & Sons Ltd.

KEY WORDS rainfall dynamics; MCMC; parameter uncertainty; exponential dispersion model; Great Lakes

Received 14 October 2011; Revised 23 February 2012; Accepted 25 February 2012

1. Introduction

Stochastic rainfall models have evolved through numerous forms ranging from paired exponential density functions representing dry and wet periods of an alternating renewal process (Gabriel and Neumann, 1962; Green, 1964), to Poisson cluster models, hidden Markov models, and pulse-based models (including Bartlett–Lewis and Neyman–Scott rectangular pulse models, as described by Velghe *et al*., 1994; Cowpertwait, 1994; Cowpertwait, 1995; Onof and Wheater, 1993; Onof and Wheater, 1994). While the range of historical applications for these models is broad (Stern and Coe, 1984), there is an increasing recognition of their suitability for resolving spatial and temporal scale discrepancies between 'output' (such as precipitation and temperature dynamics) from general circulation and regional climate models (GCMs and RCMs, for details see Lofgren *et al*., 2002; Holman *et al*., 2012) and the input required by decision-support, process-based models (the Great Lakes Advanced Hydrologic Prediction System, described in Gronewold *et al*., 2011a, is one example). The spatial extent and temporal resolution of RCM simulations, however, rarely corresponds directly to the input requirements of these regional and local-scale models, additional examples of which include hydrological models (Beven, 2001; Wagener and Wheater, 2006), terrestrial pollutant fate and transport models (Ferguson *et al*., 2003), and water quality models (Reckhow, 1999; Grant *et al*., 2001), which often run at an hourly or daily time step over a specific watershed or subbasin (for further discussion, see Bates *et al*., 1998; Chapman, 1998; Fowler *et al*., 2007; Burton *et al*., 2008; Timbal *et al*., 2009).

Burton *et al*. (2008) and Chapman (1998) note that despite advances in stochastic rainfall simulation models, including improvements in model performance, there is a need for efficient, robust model calibration routines that explicitly acknowledge parameter uncertainty, correlation, and model error, and propagate those features into rainfall simulations. To begin to bridge this research gap we evaluate the performance of an exponential-dispersion rainfall simulation model (EDM, for details see Dunn, 2004) using a Bayesian Markov chain Monte Carlo (MCMC) routine (Berry, 1996; Bolstad, 2004; Gelman *et al*., 2004). A variety of modelling approaches and distributional forms have been explored for simulating rainfall including censored quantile regression (Friederichs and Hense, 2007), generalized linear models (Furrer and Katz, 2007) and Bernoulli-gamma and zero-inflated models (Haylock *et al*., 2006; Cannon, 2008; Fernandes *et al*., 2009), each with some advantages and limitations in their practical application. We suspect our evaluation of benefits associated with explicitly quantifying uncertainty in

* Correspondence to: A. D. Gronewold, NOAA, Great Lakes Environmental Research Laboratory, Ann Arbor, MI, USA.
E-mail: drew.gronewold@noaa.gov
† This article is a US Government work and is in the public domain in the USA.

the EDM and, subsequently, assessing EDM skill in light of that uncertainty, will benefit a wide range of rainfall simulation modelling applications, including (but not limited to) recent applications of the EDM (see, for example, Hasan and Dunn, 2010; Hasan and Dunn, 2011a; Hasan and Dunn, 2011b).

We demonstrate our proposed modelling framework by applying the EDM to precipitation data over a series of subbasins of the North American Laurentian Great Lakes (for the remainder of this paper, we refer to precipitation in terms of equivalent rainfall). We calibrate the model to data from even-numbered years from 1969 to 2008, and then compare the predictive distribution of daily rainfall statistics representing rainfall magnitude and occurrence to data from odd-numbered years.

2. Methods

2.1. The exponential dispersion model

The EDM (Dunn, 2004; Dunn and Smyth, 2005; Dunn and Smyth, 2008; Hasan and Dunn, 2011a) expresses daily rainfall (y, in mm) as a mixture of discrete (i.e. zero) and continuous (i.e. non-zero) values through a single probability density function. The model is based on the assumption that the magnitude of individual rainfall events within a day are well-represented by a gamma $\mathbf{Ga}(-\alpha, \gamma)$ probability distribution with mean $= -\alpha\gamma$ and variance $= -\alpha\gamma^2$, and that the number of rainfall events in any day has a Poisson $\mathbf{Po}(\lambda)$ probability distribution with mean and variance λ. The log-probability density function is:

$$\log f(y|p, \mu, \phi) = \begin{cases} -\lambda & \text{for } y = 0 \\ -y/\gamma - \lambda - \log y + \log W(y, \phi, p) & \text{for } y > 0 \end{cases} \quad (1)$$

where μ is mean daily rainfall (in mm), ϕ is a dispersion parameter, and W is Wright's generalized Bessel function (Wright, 1933; Dunn, 2004) with power parameter p. As indicated in the left-hand side of Equation 1, the probability distribution is characterized by only three parameters (p, μ, ϕ) which are related to Poisson and Gamma distribution parameters λ, α and γ as follows (Dunn, 2004; Dunn and Smyth, 2005):

$$\lambda = \frac{\mu^{2-p}}{\phi(2-p)} \quad (2)$$
$$\alpha = (2-p)/(1-p)$$
$$\gamma = \phi(p-1)\mu^{p-1}$$

2.2. Model calibration

We calibrate the EDM using subbasin-average daily rainfall values (calculated using a Thiessen polygon interpolation approach as described in Croley and Hartmann, 1985; Quiroz *et al.*, 2011) in even-numbered years from 1969 to 2008 within each of three subbasins (chosen at random) of the Laurentian Great Lakes (Figure 1). We divide the calibration data into 36 groups with each

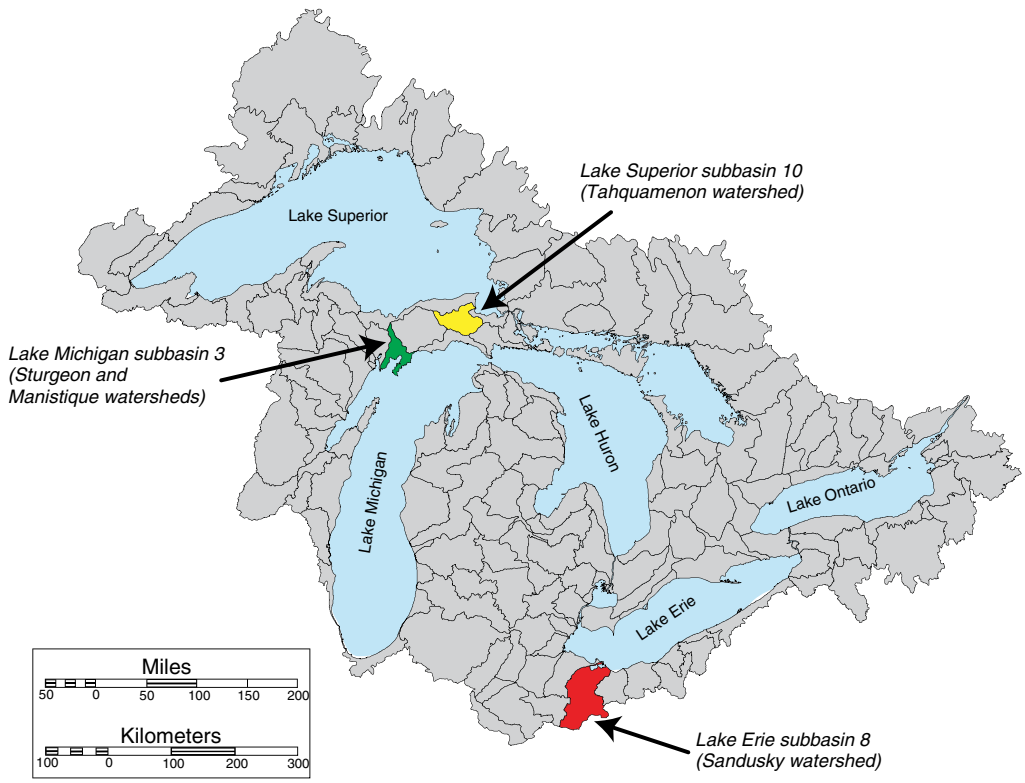

Figure 1. Delineation of the Laurentian Great Lakes drainage basin (grey shaded region) and identification of the three subbasins used in this study.

group comprised of data from one of the three subbasins and from 1 of the 12 months of the year. The first data set, for example, includes 620 daily rainfall values, each observed in the month of January during an even-numbered year from 1969 to 2008 in Lake Superior subbasin 10 (1 subbasin × 20 months × 31 days per month). Rainfall values from odd-numbered years are used for model confirmation (for further discussion of model confirmation routines, see Reckhow and Chapra, 1983; Efron and Tibshirani, 1993; Gronewold *et al.*, 2009).

We estimate parameters of the EDM (i.e. p, μ, ϕ) by drawing samples from the posterior probability distribution for each using a Bayesian MCMC approach. We begin by defining the likelihood of the data given a set of model parameters which, for $\vec{y} \equiv (y_1, \ldots, y_n)$, when applied to the EDM (Equation 1), is $L(\vec{y}|p, \mu, \phi) \equiv \prod_i^n f(y_i|p, \mu, \phi)$. We note that the product form for the likelihood implicitly assumes that daily rainfall values are independent. Previous research (see, for example, Hosseini *et al.*, 2011) suggests that daily rainfall values, however, may be correlated. Our procedure for testing this assumption is described in Section 2.3.1.

We then define a uniform prior distribution on the EDM parameters, $\pi(p, \mu, \phi)$, of the form:

$$\pi(p, \mu, \phi) \propto 1, \text{(where } 1 < p < 2, \mu > 0, \phi \geq 0). \tag{3}$$

We selected a non-informative prior so that the posterior parameter distribution would be minimally influenced by our a priori beliefs about the parameter values. We recognize that our proposed non-informative prior may prove problematic if little data are available to inform the parameters. However, we found that sufficient data are available to yield robust parameter inference with any vague prior. More specifically, we found that, in our study, model parameters are identifiable with as few as 20 observations, but not with 10. It is possible that studies in other regions (such as those with drier climes) might require alternative priors or larger data sets (or both) in order to identify model parameters. For further discussion on selection of prior probability distributions, see Press (2003).

Following Bayes' theorem, the posterior probability distribution of the EDM parameters is then:

$$\begin{aligned}\pi(p, \mu, \phi|\vec{y}) &\propto L(\vec{y}|p, \mu, \phi) \times \pi(p, \mu, \phi) \\ &\propto \prod_i^n f(y_i|p, \mu, \phi).\end{aligned} \tag{4}$$

For each iteration in the MCMC chain, we calculate the joint posterior probability density of the data and candidate model parameters (i.e. Equation 4) using the `dtweedie` function in the `tweedie` package (for details, see Dunn and Smyth, 2005; Dunn and Smyth, 2008) in the statistical software program `R` (Ihaka and Gentleman, 1996). Details of the MCMC algorithm are included in the Appendix.

We ran the MCMC algorithm three times with different initial parameter values, leading to three parallel 'chains' for each parameter. We ran all three chains for 20,000 iterations, and removed the first 10,000 as a 'burn-in' period (Gelman *et al.*, 2004). We then thinned the remaining 10,000 iterations at a 1 : 10 ratio, leaving a total of 3,000 simulated samples (1,000 per chain × 3 chains) from the posterior distribution for each parameter. We tested each MCMC run for convergence by calculating the potential scale reduction factor $\hat{R}$ and verifying that it was close to 1.0 for all MCMC chains. For details, see Gelman *et al.* (2004, p. 297).

2.3. Model confirmation

One goal of this paper is to assess the EDM as a potential tool for documenting historical dynamics, and simulating future dynamics (perhaps based on results of regional-scale climate models). However, doing so presupposes that the EDMs can accurately hindcast observed aggregate rainfall patterns. Therefore, a confirmatory study is necessary. While we recognize many statistics could be used to evaluate model accuracy (for further discussion, see Stow *et al.*, 2009), we base our model confirmation on an assessment of two metrics which correspond to rainfall frequency and magnitude, respectively.

2.3.1. Rainfall frequency

We begin by calculating the predictive distribution of the number of days with no measurable rainfall ($z_{i,j}$) in month i ($i \in 1, \ldots, 12$) and subbasin j ($j \in 1, \ldots, 3$) by following the common assumption that $z_{i,j}$ has a binomial $\mathbf{Bi}(z_{i,j}|n_i, \theta_{i,j})$ probability distribution with n_i equal to the total number of days in month i, and $\theta_{i,j}$, the posterior probability of no measurable rainfall in month i and subbasin j. Following Dunn (2004) and equation 1, θ is equal to $\exp(-\lambda)$. The predictive distribution for $z_{i,j}$ given a set of observed daily rainfall values $\vec{y}_{i,j}$ in even-numbered years from 1969 to 2008 is then:

$$\begin{aligned}p(z_{i,j}|\vec{y}_{i,j}, n_i) = \int_0^1 \binom{n_i}{z_{i,j}} \theta_{i,j}^{z_{i,j}} \times \\ (1-\theta_{i,j})^{n_i - z_{i,j}} \pi(\theta_{i,j}|\vec{y}_{i,j}) d\theta_{i,j}\end{aligned} \tag{5}$$

where $\pi(\theta_{i,j}|\vec{y}_{i,j})$, the posterior probability distribution of $\theta_{i,j}$ given $\vec{y}_{i,j}$ (based on m MCMC samples $\theta_{i,j,1}, \ldots, \theta_{i,j,m}$) is:

$$\pi(\theta_{i,j}|\vec{y}_{i,j}) = \frac{1}{m} \sum_{k=1}^{m} \delta(\theta_{i,j,k} - \theta_{i,j}) \tag{6}$$

and $\delta(\ldots)$ is the Dirac delta function with unit probability mass. The predictive probability distribution of $z_{i,j}$ is then:

$$p(z_{i,j}|\vec{y}_{i,j}, n_i) = \frac{1}{m} \binom{n_i}{z_{i,j}} \sum_{k=1}^{m} \left[\theta_{i,j,k}^{z_{i,j}} (1-\theta_{i,j,k})^{n_i - z_{i,j}}\right] \tag{7}$$

$$= \frac{1}{m} \sum_{k=1}^{m} \mathtt{dbinom}(z_{i,j}|n_i, \theta_{i,j,k}) \tag{8}$$

where `dbinom` is the binomial probability distribution density function in the statistical software program `R`. We then calculate the 95% prediction set for the number of days in a month with no measurable precipitation. Here, following Gronewold and Wolpert (2008), we define a 95% prediction set as the set of 'highest probability' integer values between 0 and n_i (the number of days in month i) for which the cumulative probability mass is at least 0.95. For illustrative examples of 95% prediction sets, see Figures 3 and 4 in Gronewold and Wolpert (2008) and Figures 5 and 6 in Gronewold *et al.* (2011b). We then calculate the corresponding 95% prediction intervals for the fraction of days in a month with no measurable precipitation by dividing the bounds of the 95% set by the respective number of days in each month.

To test our assumption of conditional independence of daily rainfall values, we calculate the predictive distribution for two quantities; the number of days with no measurable rainfall in a given month *given that* the previous day had no measurable rainfall (which we identify as z^0), and the number of days with no measurable rainfall in a given month *given that* the previous day had measurable rainfall (which we refer to as z'). These two predictive distributions can be expressed (following the logic of Equations 5 through 8) as (for simplicity, we have removed i and j from Equations 9 and 10):

$$p(z^0|\vec{y}, n^0) = \frac{1}{m}\sum_{k=1}^{m} \texttt{dbinom}(z^0|n^0, \theta_k) \qquad (9)$$

and,

$$p(z'|\vec{y}, n') = \frac{1}{m}\sum_{k=1}^{m} \texttt{dbinom}(z'|n', \theta_k) \qquad (10)$$

where n^0 is the total number of days in a given month preceded by a day with no rainfall, and n' is the total number of days in a given month preceded by a day with rainfall. Following the procedure described above, we then calculate the 95% prediction set for both z^0 and z', and compare both to observed values from the calibration and confirmation periods.

2.3.2. *Rainfall magnitude*

To assess the EDM's potential for simulating daily rainfall magnitude, we calculate the probability distribution of *measurable* daily rainfall (i.e. daily rainfall amounts greater than zero) for each subbasin-month combination and compare it to the observed probability distribution of *measurable* daily rainfall. We do this by entering all 3000 MCMC samples from each EDM parameter posterior distribution into the `rtweedie` function in the `tweedie` package (Dunn and Smyth, 2005; Dunn and Smyth, 2008), an approach that generates (for one iteration of the `rtweedie` function) 3,000 samples from the posterior predictive distribution of y. We note here that this approach is analogous to simulating a 3,000-day long time series of daily rainfall values.

To fully capture intrinsic variability in the probability distribution of daily rainfall values (due, in part, to a finite number of samples from the EDM parameter joint posterior distribution), we repeat this procedure 10,000 times, excluding from the final set of simulated daily rainfall values those which equal zero (a schematic, and a slightly more detailed description of this procedure, are included in Figure 8 and the Appendix, respectively). We then calculate the quantiles of the observed measurable rainfall times series (separately for both calibration and confirmation periods) and the quantiles from each of the 10,000 sets of simulated rainfall values. Finally, we compare each observed rainfall value to the set of simulated rainfall values from the corresponding quantile of the simulated sets. Our approach differs from conventional quantile–quantile comparisons because it explicitly depicts uncertainty and (unlike comparisons of quantile residuals) identifies where potential sources of bias are likely to arise from the simulated time series.

3. Results

3.1. Model calibration

Our model calibration results indicate several patterns in the EDM parameters both within and across each of the subbasins (Figure 2). For example, the 95% highest posterior density (HPD) regions (Figure 2) suggest that the EDM power parameter, p, varies throughout the year when fit to the Lake Superior and Lake Michigan subbasin data, but that it is relatively consistent and somewhat higher throughout the year when fit to the Lake Erie subbasin data. The relatively high values and low variability of the power parameter for the Lake Erie watershed likely reflect a combination of empirical evidence (Figure 4) indicating that the fraction of days with no rainfall in the Lake Erie watershed is relatively consistent throughout the year, and that the mean daily rainfall is also relatively high and relatively consistent for the Lake Erie watershed as well. For a more rigorous assessment of EDM power parameter dynamics, see Dunn (2004) and Hasan and Dunn (2011a).

We also find that the expected value of daily rainfall μ follows a seasonal pattern with peak values for Lake Superior subbasin 10 in April, for Lake Michigan subbasin 3 in September, and for Lake Erie subbasin 8 in June. We could assess mean daily rainfall values (i.e. calculate a value for μ) through an empirical data assessment as well, however assessing the posterior distribution of μ in a Bayesian framework explicitly propagates data variability through the posterior distribution into uncertainty in rainfall forecasts. Finally, we find that the dispersion parameter, ϕ, varies from season to season, although the pattern differs among subbasins (bottom row, Figure 2).

Our parameter correlation assessment indicates that the expected value of daily rainfall μ is independent of

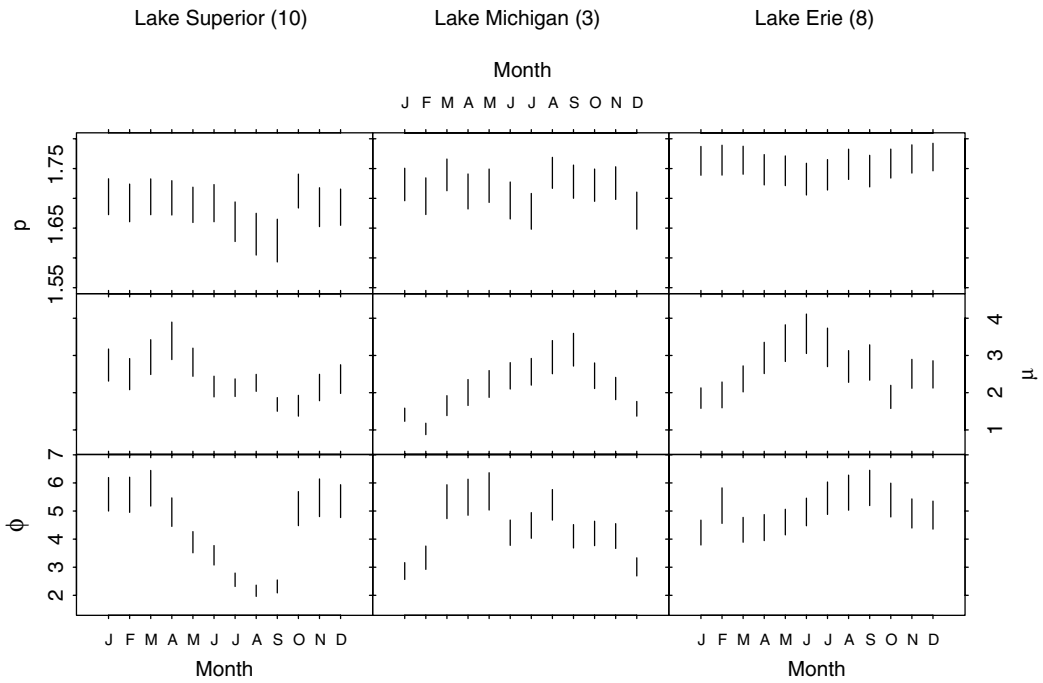

Figure 2. Model calibration results, including posterior distribution 95% highest posterior density (HPD) regions (vertical lines) for each combination of EDM parameter (rows) and subbasin (columns). Calibration results are based on conditioning the EDM to rainfall data in even-numbered years from 1969 to 2008.

both the dispersion parameter ϕ and power parameter p, based on the orientation of the marginal posterior density contours in the $\{\mu, p\}$- and $\{\mu, \phi\}$-planes (middle left and bottom centre panels in Figure 3, respectively). We find, however, that ϕ and p are positively correlated (middle panel Figure 3). The patterns in Figure 3 are based on calibrating the EDM to January rainfall data in Lake Superior subbasin 10, however we find (results not shown) they represent parameter relationships when the EDM is calibrated using rainfall data from other months and other subbasins.

3.2. Model confirmation

3.2.1. Rainfall frequency

A comparison between the observed fraction of days in a given month with no measurable rainfall and the predictive distribution indicates that the EDM slightly underestimates the probability of zero rainfall (Figure 4). The 95% prediction regions contain approximately 84% of the observations from the calibration period and 82% from the confirmation period across all subbasin–month combinations. This proportion does not vary systematically within subbasins across months, nor is there a systematic tendency to miss either high or low extreme values. We recognize that this incomplete coverage could occur for several reasons, including temporal dependence in the EDM, as well as the use of Thiessen-weighting to synthesize gauge-based observations. Future work will focus on differentiating these and other possibilities.

Our results also suggest that the EDM does not fully represent the temporal dependence of daily rainfall values. For example, we found that the 95% prediction interval for z^0 (using EDM parameter values conditioned under an assumption of conditional independence) included 71% of observations from the calibration period, and 69% of observations from the confirmation period. Similarly, we found that the 95% prediction interval for z' included 72% and 74% of the observations from the calibration and confirmation periods, respectively. The relatively lower skill of the EDM when forecasting rainfall values conditioned on their antecedent conditions implies that $p(z|\vec{y}, n) \neq p(z^0|\vec{y}, n^0) \neq p(z'|\vec{y}, n')$, and suggests that the EDM in its current form is more suitable for modelling monthly precipitation values (assuming less temporal dependence when aggregating from daily to monthly scales) or that it may need to be modified if applied to daily rainfall values. We find that previous applications of the EDM (Dunn, 2004; Dunn and Smyth, 2005; Hasan and Dunn, 2011a), which progressively gravitate towards a focus on monthly rainfall data, do not explicitly acknowledge (through, for example, the type of skill assessment we present here) nor emphasize (perhaps in qualitative terms) this important distinction. One possibility to incorporate the dependence structure would be to use the Tweedie distribution in a generalized linear modelling framework (Hasan and Dunn, 2011a), incorporating covariates such as lags in the observed daily rainfall.

3.2.2. Rainfall magnitude

Generally, the EDM provides a reasonable reproduction of the probability distribution of measurable daily rainfall for each subbasin–month combination, indicated by the

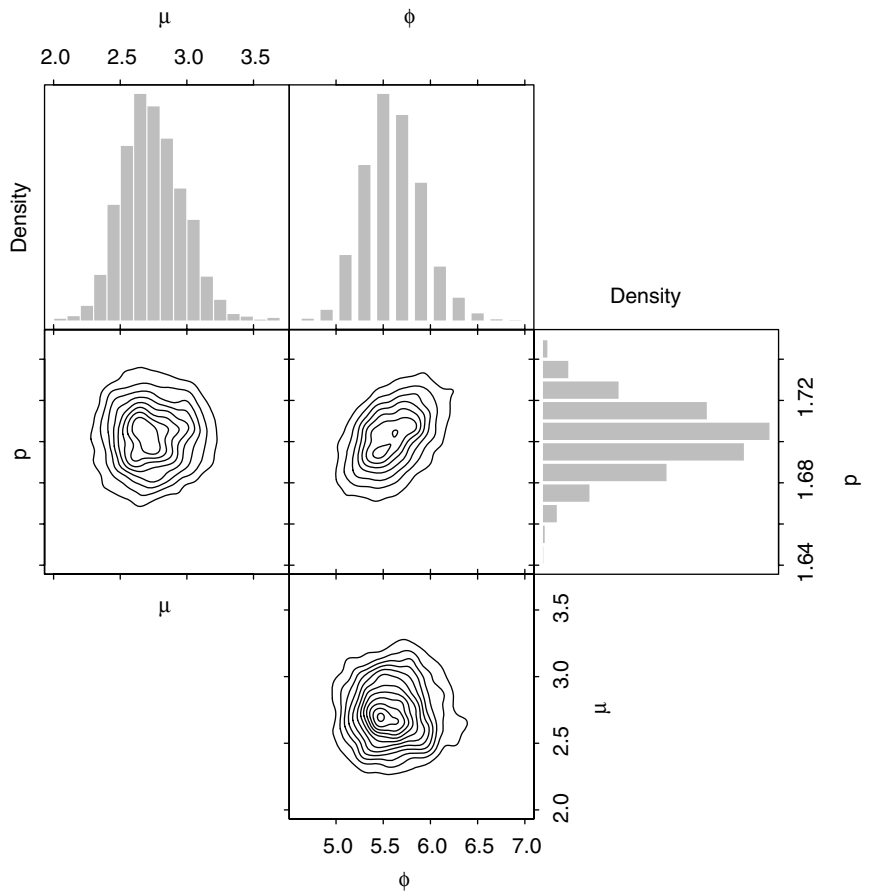

Figure 3. Histograms and contour plots of the marginal and joint posterior probability distributions (respectively) for EDM parameters. Results shown are based on calibrating the EDM to daily rainfall values from January across even-numbered years from 1969 through 2008 in Lake Superior subbasin 10 (Figure 1). The scale of the probability density axis (labeled 'Density') is intentionally removed from each histogram.

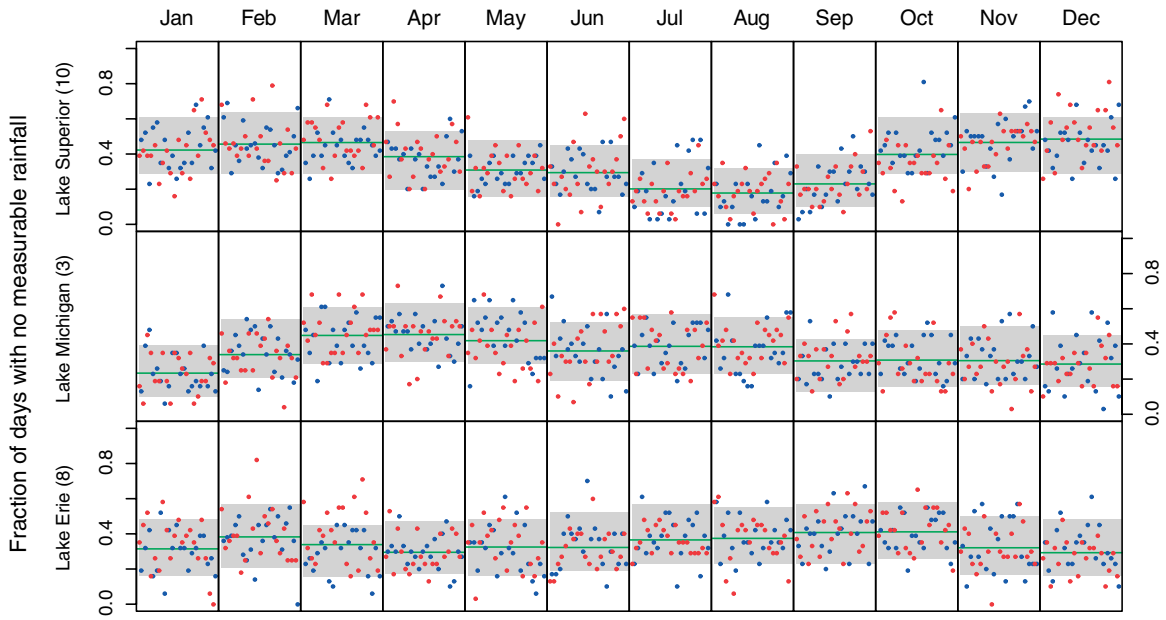

Figure 4. Model confirmation results, including 95% prediction intervals (grey regions) for the fraction of days with no measurable rainfall in each month. In each panel, dots represent the observed fraction of days with no measurable rainfall in a given month from 1969 (left-most dot in each panel) through 2008 (right-most dot in each panel). Blue dots represent the fraction of days with no measurable rainfall in a given month in even years (calibration data), and red dots represent the fraction of days with no measurable rainfall in a given month in odd years (confirmation data).

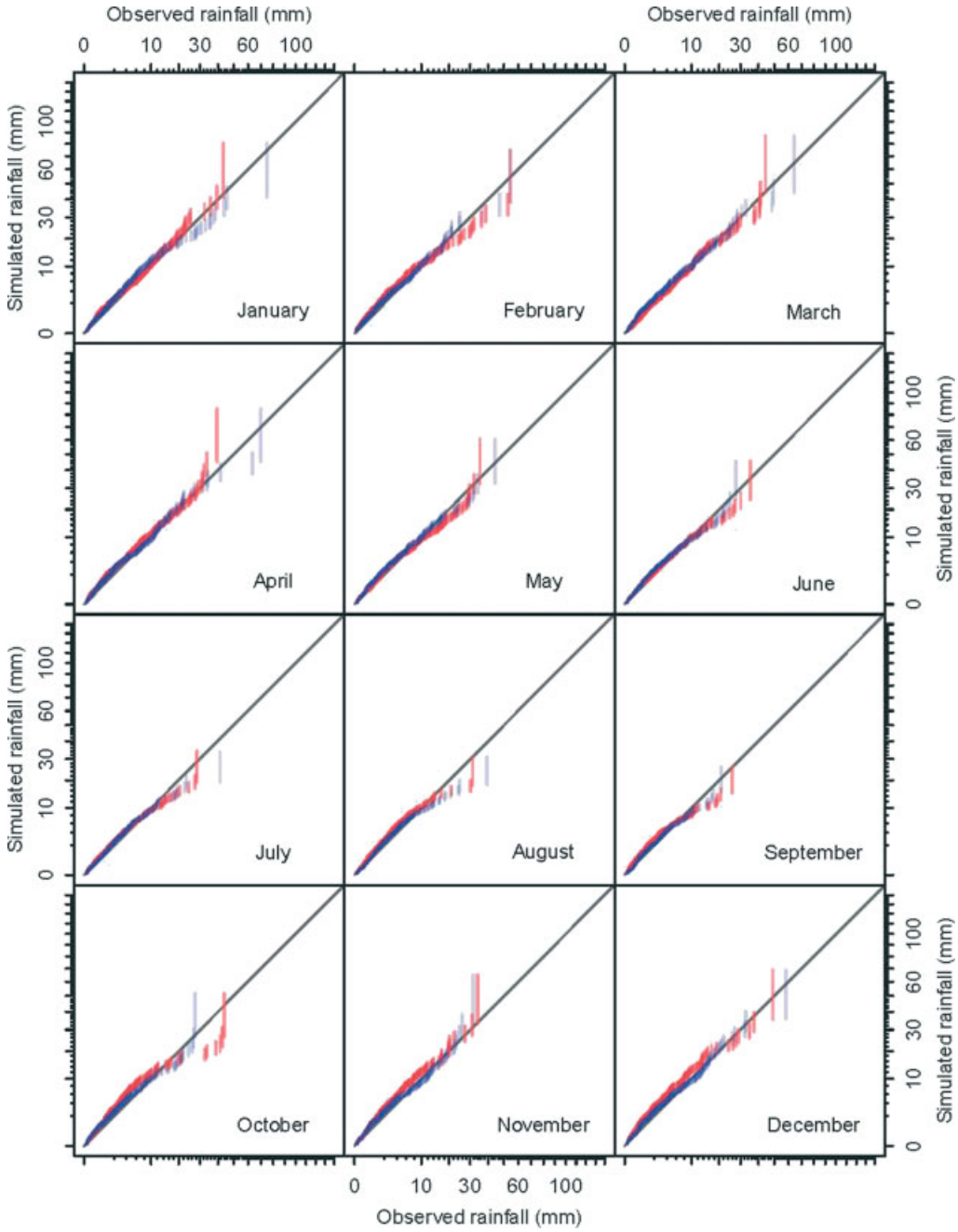

Figure 5. Comparison between observed and simulated rainfall values over Lake Superior subbasin 10 from the same quantile. Vertical lines indicate the range of simulated values from a particular quantile generated using all MCMC samples from the EDM joint parameter posterior distribution. Blue lines represent calibration data, and red lines represent confirmation data. The 1 : 1 line (black) is shown for reference. Axes are presented at square-root scale to improve clarity.

fact that most of the vertical blue (calibration period) and red (confirmation period) lines in Figures 5 through 7 intersect the 1 : 1 line. Each vertical line indicates the range of simulated daily rainfall values for each quantile from the corresponding set of observed non-zero daily rainfall values. Put differently, the location of each vertical line along the x-axis in each panel corresponds to an observed daily rainfall value. The corresponding height of each vertical line reflects the uncertainty in the predicted daily rainfall value from the quantile of the observed daily rainfall value. Within a particular panel in Figures 5 through 7, relatively short vertical lines that intersect the 1 : 1 line indicate a predictive distribution for daily rainfall (for the subbasin–month combination represented by that panel) which is similar to the observed rainfall probability distribution. Wide vertical lines in any given panel that intersect the 1 : 1 line indicate that there is significant uncertainty in a particular quantile of the predictive distribution, but that the observed rainfall value from the same quantile is within the predicted range of values.

While the EDM appears to reproduce the general features of the probability distribution of daily rainfall, we observe, as expected, significant variability in the EDM-derived upper quantiles of the daily rainfall distribution.

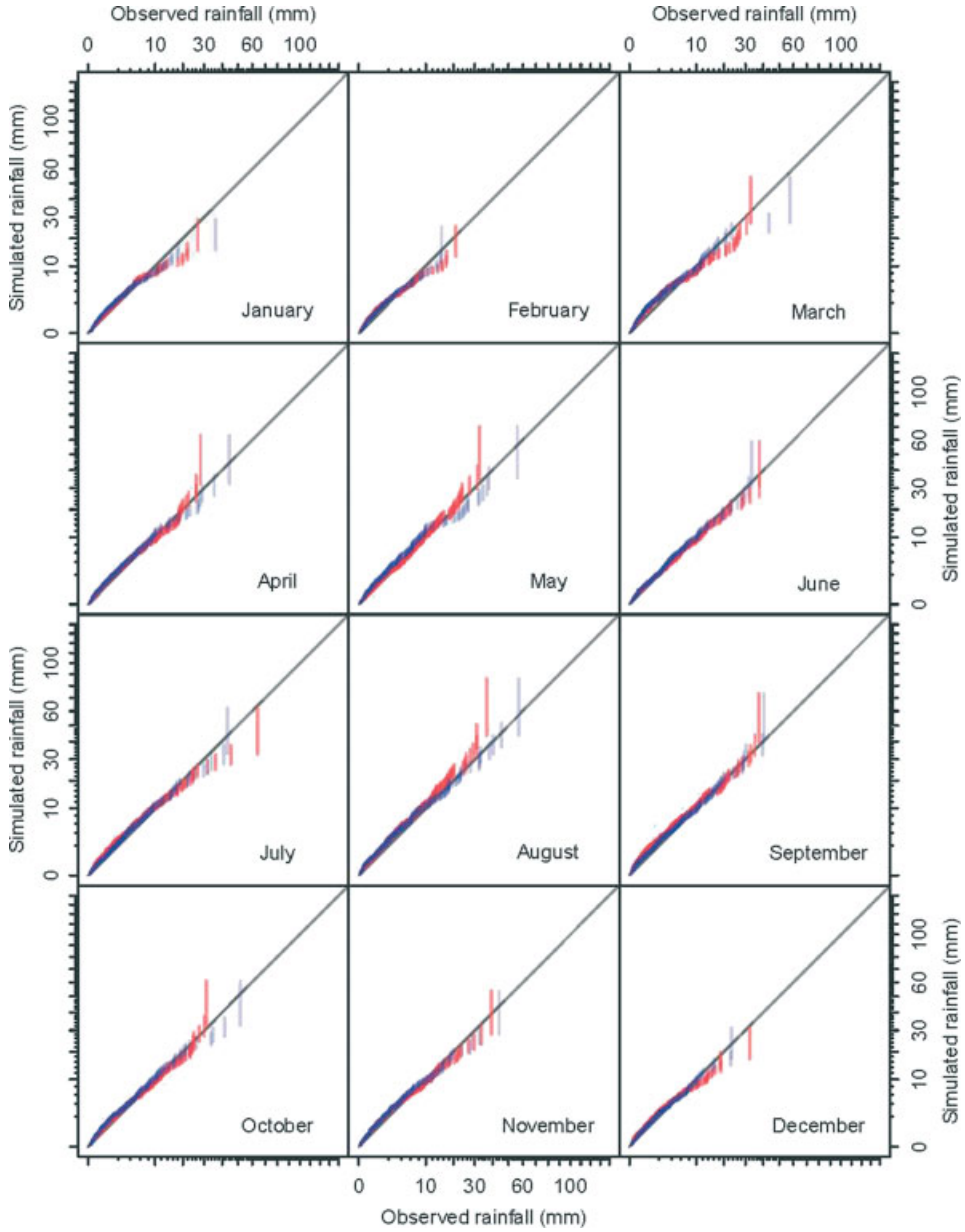

Figure 6. Comparison between observed and simulated rainfall values over Lake Michigan subbasin 3 from the same quantile. Vertical lines indicate the range of simulated values from a particular quantile generated using all MCMC samples from the EDM joint parameter posterior distribution. Blue lines represent calibration data, and red lines represent confirmation data. The 1 : 1 line (black) is shown for reference. Axes are presented at square-root scale to improve clarity.

For example, the upper left-hand panel of Figure 5 indicates that our calibrated model (for January in Lake Superior subbasin 10) simulates extreme daily rainfall ranging between roughly 28 and 135 mm (as indicated by the extent of the right-most red and blue vertical lines) while the corresponding rainfall values from the same quantile in the observed data sets were about 73 mm (calibration years) and 41 mm (validation years).

4. Discussion and conclusions

This paper assesses the EDM in a Bayesian MCMC framework for simulating daily rainfall in three subbasins of the North American Laurentian Great Lakes. We have shown that, within this framework, explicitly acknowledging variability in daily rainfall time series data through uncertainty and correlation in the EDM parameter joint posterior probability distribution (Figure 3) leads to appropriate representation of uncertainty in simulated daily rainfall magnitude and occurrence. By 'appropriate representation of uncertainty' we mean that the uncertainty expressed in EDM forecasts neither greatly exceeds, nor significantly underestimates the variability observed in independent rainfall time series. We base this assessment on two measures of model skill; (1) the fraction of 95% prediction intervals which include the observed fraction of days with no measurable rainfall in an independent confirmation data set, and (2) a quantile-based comparison of simulated and observed

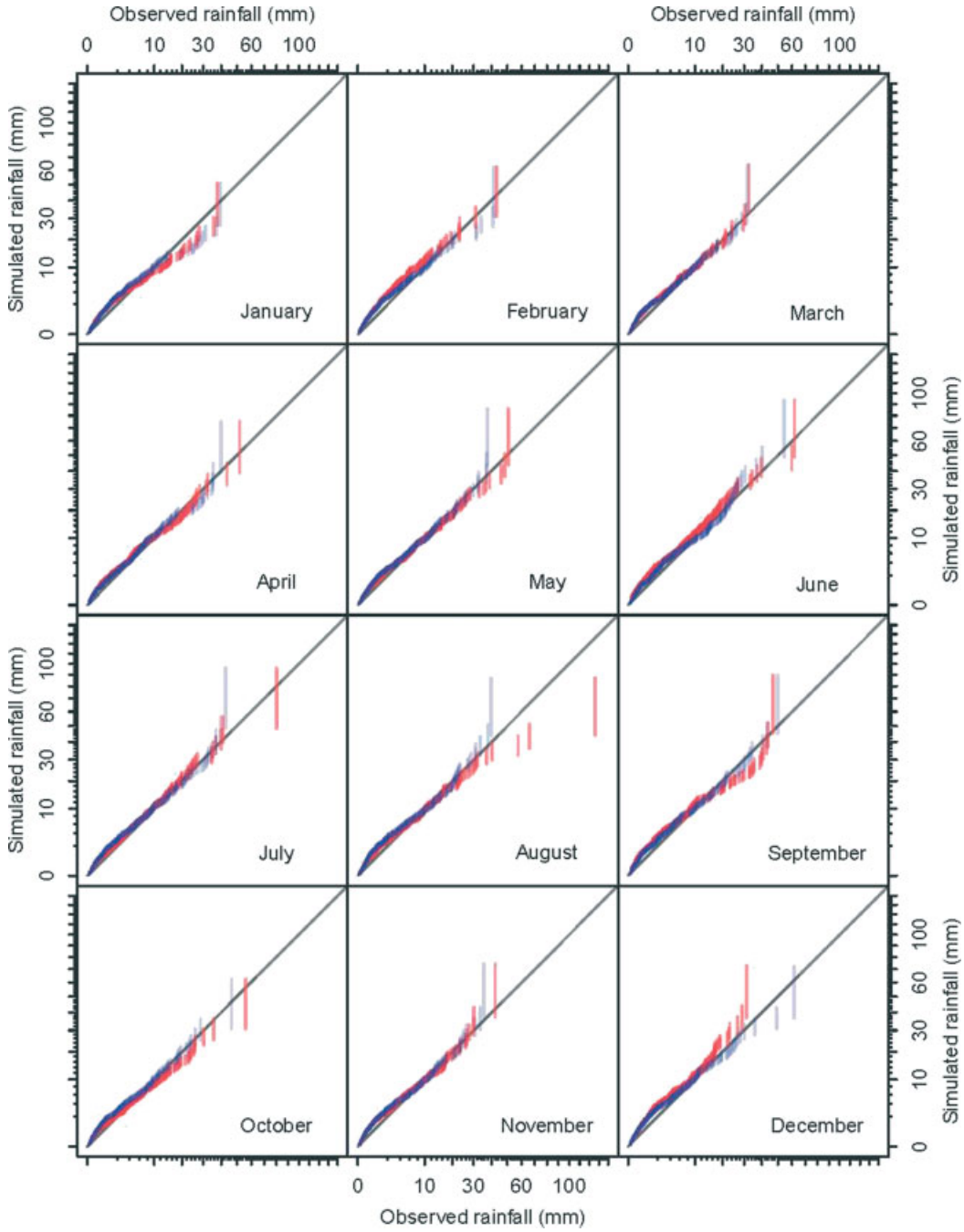

Figure 7. Comparison between observed and simulated rainfall values over Lake Erie subbasin 8 from the same quantile. Vertical lines indicate the range of simulated values from a particular quantile generated using all MCMC samples from the EDM joint parameter posterior distribution. Blue lines represent calibration data, and red lines represent confirmation data. The 1 : 1 line (black) is shown for reference. Axes are presented at square-root scale to improve clarity.

daily rainfall time series. Other metrics and assessment techniques could be used, including total monthly or annual rainfall amounts, coupled, perhaps, with an analysis of the histograms of Bayesian posterior predictive p-values for these metrics. For a detailed description of these and similar assessment metrics, including probability integral transform and verification rank histograms, see Raftery *et al*. (2005); Elmore (2005); Gronewold *et al*. (2009). We plan to explore these alternatives in future applications of the EDM (including, for example, those which support probabilistic approaches to resource management, as described in Gronewold and Borsuk, 2009).

In addition to providing a robust basis for quantifying uncertainty in EDM parameters and forecasts, our Bayesian MCMC calibration approach provides a convenient alternative to conventional model calibration and forecasting schemes. For example, stochastic rainfall simulation models are often calibrated through optimization algorithms that yield parameter point estimates and produce deterministic comparisons between simulated and observed time series metrics (which, in some cases,

include procedures for 'matching' observed and simulated statistics using parameter perturbations, as described in Burton *et al.*, 2008). Our approach provides an alternative that explicitly acknowledges uncertainty and variability in rainfall dynamics. In doing so, our application of the EDM allows us to choose between either reflecting the same degree of variability observed in historic time series in future simulations, or modifying the location and scale of EDM parameters (based, perhaps, on regional climate model outputs). The latter approach acknowledges and meets the growing need for models which reflect changing conditions over time (Milly *et al.*, 2008) and provides an alternative to scenario-weighting approaches (for similar applications and further discussion, see (Lofgren *et al.*, 2002). Furthermore, we recognize that while we have not rigorously tested our assumption of decadal stationarity in this particular application of the EDM, the rainfall data from 1969 through 2008 in the three subbasins studied does not appear to demonstrate a significant trend over this 40 year period (Figure 4), with the exception of the frequency of daily rainfall events in September in the Lake Superior Tahquamenon watershed (subbasin 10), which appears to be decreasing (i.e. the fraction of days with no measurable rainfall is increasing). Regardless, our application of the EDM can easily be transferred to a more comprehensive regional or hierarchical analysis of precipitation patterns within the Great Lakes basin, and we view this as an area for future research. A regional frequency analysis, in particular, could potentially reduce some of the variability in our estimates (Figures 5–7) of extreme rainfall quantiles (for further discussion, see Trefry *et al.*, 2005; Ribatet *et al.*, 2007).

Our study also underscores pending difficulties associated with downscaling regional climate change scenarios into local scale dynamics, particularly in regions (such as the Laurentian Great Lakes) with significant local-scale spatial climate variability. For example, the Lake Superior Tahquamenon watershed (subbasin 10) and the Lake Michigan Sturgeon and Manistique watershed (subbasin 3) are within 50 miles (roughly 80 kilometers) of each other (Figure 1), yet the rainfall dynamics in each differ (Figure 4), due in part to different weather and wind patterns over and adjacent to Lake Superior. By calibrating the EDM to local scale climate data, our approach serves as an ideal cornerstone for a future coupled RCM-EDM which propagates regional climate patterns into local-scale rainfall (and other weather component) dynamics.

In this process, we have identified three areas for improvement. First, we recognize that EDM performance could be improved through an analysis of potential thresholds for 'measurable' rainfall amounts (Burton *et al.*, 2008) and corrections for potential bias these thresholds might introduce. Second, we acknowledge that temporal dependencies in daily rainfall data are not routinely captured by the EDM, and therefore the EDM is likely most suitable for application to monthly time series data. We underscore here how our assessment provides a robust quantitative basis for making this distinction, and that future research might focus on supplementing the EDM with algorithms for expressing autocorrelation on a daily time scale. Third, the daily rainfall values we used here are, in fact, averaged (based on multiple individual rain gauges) over each subbasin, an approach which could overestimate the frequency and underestimate the intensity of daily rainfall dynamics (Bates *et al.*, 1998). We intend, in future research, to calibrate the EDM to individual rain gauge data and then combine the EDM parameters using a Bayesian model averaging or hierarchical approach (Raftery *et al.*, 2005; Gelman and Hill, 2007; Ancelet *et al.*, 2010).

Acknowledgements

The authors thank Brent Lofgren, Anne Clites, and two anonymous reviewers whose comments improved the clarity and overall quality of this paper. The authors also thank Cathy Darnell for providing editorial and graphics support. This paper is GLERL contribution number 1624.

A1. Appendix

A1.1. Markov chain Monte Carlo (MCMC) algorithm

The algorithm used in this paper is a three-parameter Metropolis-Hastings algorithm, which is a type of Markov chain Monte Carlo (MCMC) algorithm used to generate random samples from intractable probability densities. The algorithm is implemented iteratively with k being the current iteration, and the parameter set (or 'state') for a particular iteration is defined as $\{p_k, \mu_k, \phi_k\}$. A trial state $\{p', \mu', \phi'\}$ is generated using a trivariate normal distribution centred at the current state:

$$(p', \mu', \phi') \sim N_3\left(\{p_k, \mu_k, \phi_k\}, \Sigma\right)$$

where Σ is a covariance matrix.

The likelihood and prior distribution are evaluated at this trial state. Alternatively, one can change a random subset of the parameters at each iteration rather than the whole set. The trial state is accepted with probability

$$\Pr\{p', \mu', \phi' \rightarrow p_{k+1}, \mu_{k+1}, \phi_{k+1}\} = \min\left(1, \frac{L(\vec{y}|p', \mu', \phi') \times \pi(p', \mu', \phi')}{L(\vec{y}|p_k, \mu_k, \phi_k) \times \pi(p_k, \mu_k, \phi_k)}\right)$$

and if the trial state is not accepted, the current state $\{p_k, \mu_k, \phi_k\}$ is retained at iteration $k + 1$.

The set of states over all iterations constitutes a 'chain'. The acceptance–rejection step ensures that the chain will eventually converge to the posterior distribution, $\pi(p, \mu, \phi|\vec{y})$. Convergence was checked using the `CODA` package in `R`, and we found no evidence of non-convergence.

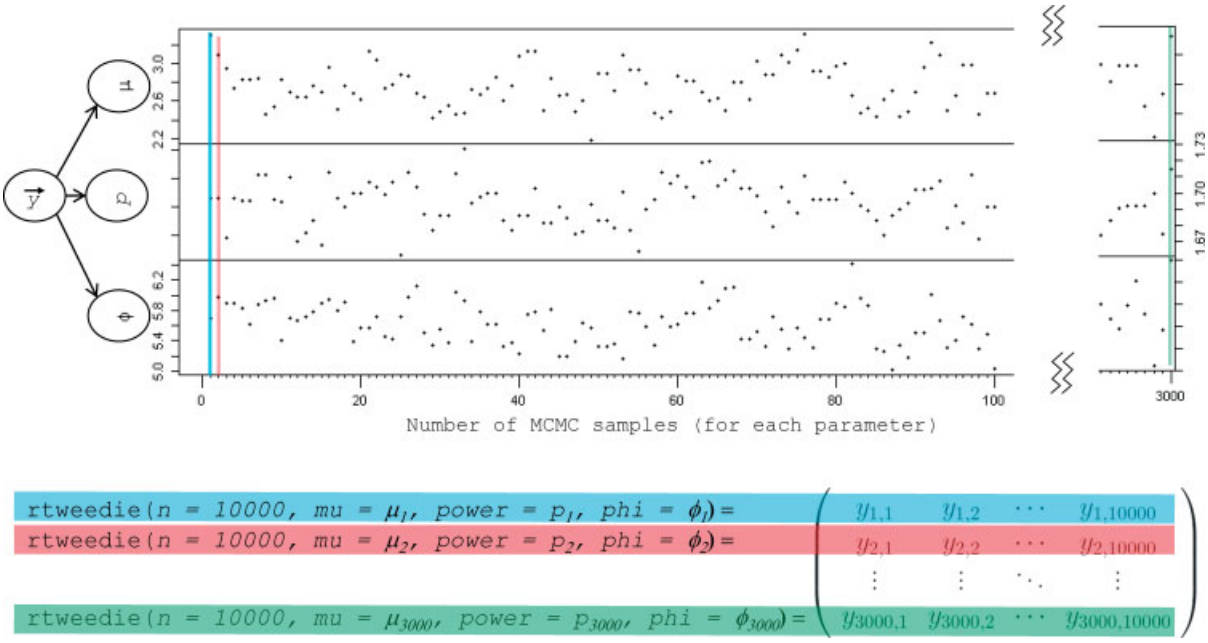

Figure 8. Schematic representation of procedure for simulating the probability distribution of measurable daily rainfall amounts (see Appendix A2 for details).

Generally, one wants the Metropolis-Hastings algorithm to accept between 20% and 50% of the trial states. Too many or too few mean the chain is not sampling the posterior efficiently. This fraction can be tuned by trying different values of Σ. After some experimentation, we found that a diagonal covariance with $\Sigma = \mathbf{diag}(0.002, 0.07, 0.07)$ yields a 32% acceptance fraction.

A2. Simulating the probability distribution of measurable daily rainfall

Figure 8 provides a schematic representation of our procedure for simulating the probability distribution of the magnitude of measurable daily rainfall events. After simulating 3000 MCMC samples from the posterior probability distribution for each EDM parameter (μ, p, ϕ) for a given set of observed daily rainfall values $\overrightarrow{y}$ for a particular month and subbasin (as described in the previous appendix and in Section 2.2, and represented by the upper-half of Figure 8), we systematically pass each triplet of parameter values from the MCMC chain to the `rtweedie` package (Dunn and Smyth, 2005, 2008). We then use the `rtweedie` package to simulate 10,000 daily rainfall values from an EDM for each triplet. For example, in Figure 8, a blue vertical line passes through the first value of μ, p, and ϕ in each MCMC chain (represented as μ_1, p_1, and ϕ_1) which are passed to the `rtweedie` package to simulate 10,000 rainfall values for that particular triplet. The 10,000 simulated rainfall amounts using μ_1, p_1, and ϕ_1 are represented by the vector $y_{1,1}, y_{1,2}, \ldots, y_{1,10\,000}$ (highlighted in blue) in the array in the bottom-right of Figure 8. We repeat this procedure for the second triplet of parameter values (highlighted in red in Figure 8) and continue up to the final (i.e. 3,000th) value in each MCMC chain (highlighted in green in Figure 8). The set of simulated, non-zero values from the resulting array (bottom-right corner of Figure 8) constitutes our approximated probability distribution of measurable daily rainfall values for the given month and subbasin.

References

Ancelet S, Etienne M, Benoît H, Parent E. 2010. Modelling spatial zero-inflated continuous data with an exponentially compound Poisson process. *Environmental and Ecological Statistics* **17**(3): 347–376.

Bates B, Charles S, Hughes J. 1998. Stochastic downscaling of numerical climate model simulations. *Environmental Modelling & Software* **13**(3–4): 325–331.

Berry DA. 1996. *Statistics: A Bayesian Perspective*, Duxbury Press: Belmont, CA.

Beven K. 2001. How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences* **5**(1): 1–12.

Bolstad WM. 2004. *Introduction to Bayesian Statistics*, Wiley-Interscience: Hoboken, NJ.

Burton A, Kilsby C, Fowler H, Cowpertwait P, O'Connell P. 2008. RainSim: a spatial-temporal stochastic rainfall modelling system. *Environmental Modelling & Software* **23**(12): 1356–1369.

Cannon A. 2008. Probabilistic multisite precipitation downscaling by an expanded Bernoulli-Gamma density network. *Journal of Hydrometeorology* **9**(6): 1284–1300.

Chapman T. 1998. Stochastic modelling of daily rainfall: the impact of adjoining wet days on the distribution of rainfall amounts. *Environmental Modelling & Software* **13**(3–4): 317–324.

Cowpertwait P. 1994. A generalized point process model for rainfall. *Proceedings: Mathematical and Physical Sciences* **447**(1929): 23–37.

Cowpertwait P. 1995. A generalized spatial-temporal model of rainfall based on a clustered point process. *Proceedings: Mathematical and Physical Sciences* **450**(1938): 163–175.

Croley TE, Hartmann HC. 1985. Resolving Thiessen polygons. *Journal of Hydrology* **76**(3–4): 363–379.

Dunn PK. 2004. Occurrence and quantity of precipitation can be modelled simultaneously. *International Journal of Climatology* **24**: 1231–1239.

Dunn PK, Smyth GK. 2005. Series evaluation of Tweedie exponential dispersion model densities. *Statistics and Computing* **15**(4): 267–280.

Dunn PK, Smyth GK. 2008. Evaluation of Tweedie exponential dispersion model densities by Fourier inversion. *Statistics and Computing* **18**(1): 73–86.

Efron B, Tibshirani RJ. 1993. *An Introduction to the Bootstrap*, Chapman & Hall/CRC: New York.

Elmore K. 2005. Alternatives to the chi-square test for evaluating rank histograms from ensemble forecasts. *Weather and Forecasting* **20**: 789–795.

Ferguson C, Husman AMD, Altavilla N, Deere D, Ashbolt N. 2003. Fate and transport of surface water pathogens in watersheds. *Critical Reviews in Environmental Science and Technology* **33**(3): 299–361.

Fernandes M, Schmidt A, Migon H. 2009. Modelling zero-inflated spatio-temporal processes. *Statistical Modelling* **9**(1): 3.

Fowler H, Blenkinsop S, Tebaldi C. 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology* **27**(12): 1547–1578.

Friederichs P, Hense A. 2007. Statistical downscaling of extreme precipitation events using censored quantile regression. *Monthly Weather Review* **135**(8): 2365–2378.

Furrer E, Katz R. 2007. Generalized linear modeling approach to stochastic weather generators. *Climate Research* **34**(2): 129.

Gabriel K, Neumann J. 1962. A Markov chain model for daily rainfall occurrence at Tel Aviv. *Quarterly Journal of the Royal Meteorological Society* **88**(375): 90–95.

Gelman A, Carlin JB, Stern HS, Rubin DB. 2004. *Bayesian Data Analysis*, Chapman & Hall/CRC: Boca Raton, FL.

Gelman A, Hill J. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press: New York, NY.

Grant SB, Sanders BF, Boehm AB, Redman JA, Kim JH, Mrse RD, Chu AK, Gouldin M, McGee CD, Gardiner NA, Jones BH, Svejkovsky J, Leipzig GV. 2001. Generation of Enterococci bacteria in a coastal saltwater marsh and its impact on surf zone water quality. *Environmental Science & Technology* **35**(12): 2407–2416.

Green J. 1964. A model for rainfall occurrence. *Journal of the Royal Statistical Society. Series B (Methodological)* **26**(2): 345–353.

Gronewold AD, Borsuk M. 2009. A software tool for translating deterministic model results into probabilistic assessments of water quality standard compliance. *Environmental Modelling & Software* **24**(10): 1257–1262.

Gronewold AD, Clites A, Hunter T, Stow C. 2011a. An appraisal of the Great Lakes advanced hydrologic prediction system. *Journal of Great Lakes Research* **37**: 577–583.

Gronewold AD, Myers L, Swall JL, Noble RT. 2011b. Addressing uncertainty in fecal indicator bacteria dark inactivation rates. *Water Research* **45**: 652–664.

Gronewold AD, Qian SS, Wolpert RL, Reckhow KH. 2009. Calibrating and validating bacterial water quality models: a Bayesian approach. *Water Research* **42**: 2688–2698.

Gronewold AD, Wolpert RL. 2008. Modeling the relationship between most probable number (MPN) and colony-forming unit (CFU) estimates of fecal coliform concentration. *Water Research* **42**(13): 3327–3334.

Hasan M, Dunn P. 2010. A simple Poisson-gamma model for modelling rainfall occurrence and amount simultaneously. *Agricultural and Forest Meteorology* **150**(10): 1319–1330.

Hasan M, Dunn P. 2011a. Two Tweedie distributions that are near-optimal for modelling monthly rainfall in Australia. *International Journal of Climatology* **31**(9): 1389–1397.

Hasan M, Dunn P. 2011b. Understanding the effect of climatology on monthly rainfall amounts in Australia using Tweedie GLMs. *International Journal of Climatology*. DOI: 10.1002/joc.2332.

Haylock M, Cawley G, Harpham C, Wilby R, Goodess C. 2006. Downscaling heavy precipitation over the United Kingdom: a comparison of dynamical and statistical methods and their future scenarios. *International Journal of Climatology* **26**(10): 1397–1415.

Holman K, Gronewold AD, Notaro M, Zarrin A. 2012. Improving historical precipitation estimates over the Lake Superior basin. *Geophysical Research Letters* **39**(3): L03405.

Hosseini R, Le N, Zidek J. 2011. Selecting a binary Markov model for a precipitation process. *Environmental and Ecological Statistics*. **18**(4): 795–820, DOI: 1–2610.1007/s10651-010-0169-1.

Ihaka R, Gentleman R. 1996. R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5**(3): 299–314.

Lofgren BM, Quinn FH, Clites AH, Assel RA, Eberhardt AJ, Luukkonen CL. 2002. Evaluation of potential impacts on Great Lakes water resources based on climate scenarios of two GCMs. *Journal of Great Lakes Research* **28**(4): 537–554.

Milly PC, Betancourt J, Falkenmark M, Hirsch RM, Kundzewicz ZW, Lettenmaier DP, Stouffer RJ. 2008. Stationarity is dead: whither water management? *Science* **319**(5863): 573–574.

Onof C, Wheater H. 1993. Modelling of British rainfall using a random parameter Bartlett-Lewis rectangular pulse model. *Journal of Hydrology* **149**(1–4): 67–95.

Onof C, Wheater H. 1994. Improvements to the modelling of British rainfall using a modified random parameter Bartlett-Lewis rectangular pulse model. *Journal of Hydrology* **157**(1–4): 177–195.

Press SJ. 2003. *Subjective and Objective Bayesian Statistics: Principles, Models, and Applications*, Wiley-Interscience: Hoboken, NJ.

Quiroz R, Yarlequ C, Posadas A, Mares V, Immerzeel WW. 2011. Improving daily rainfall estimation from ndvi using a wavelet transform. *Environmental Modelling & Software* **26**(2): 201–209.

Raftery A, Gneiting T, Balabdaoui F, Polakowski M. 2005. Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review* **133**(5): 1155–1174.

Reckhow KH. 1999. Water quality prediction and probability network models. *Canadian Journal of Fisheries and Aquatic Sciences* **56**(7): 1150–1158.

Reckhow KH, Chapra SC. 1983. Confirmation of water quality models. *Ecological Modelling* **20**(2–3): 113–133.

Ribatet M, Sauquet E, Grésillon J, Ouarda T. 2007. A regional Bayesian POT model for flood frequency analysis. *Stochastic Environmental Research and Risk Assessment* **21**(4): 327–339.

Stern R, Coe R. 1984. A model fitting analysis of daily rainfall data. *Journal of the Royal Statistical Society. Series A (General)* **147**(1): 1–34.

Stow CA, Jolliff J, Jr. Doney DJM, Allen SC, Friedrichs JI, Rose MA, Wallhead KA. P 2009. Skill assessment for coupled biological/physical models of marine systems. *Journal of Marine Systems* **76**(1–2): 4–15.

Timbal B, Fernandez E, Li Z. 2009. Generalization of a statistical downscaling model to provide local climate change projections for Australia. *Environmental Modelling & Software* **24**(3): 341–358.

Trefry C, Watkins D, Jr. Johnson D. 2005. Regional rainfall frequency analysis for the State of Michigan. *Journal of Hydrologic Engineering* **10**: 437.

Velghe T, Troch P, De Troch F, Van de Velde J. 1994. Evaluation of cluster-based rectangular pulses point process models for rainfall. *Water Resources Research* **30**(10): 2847–2858.

Wagener T, Wheater HS. 2006. Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. *Journal of Hydrology* **320**(1–2): 132–154.

Wright E. 1933. On the coefficients of power series having essential singularities. *Journal of the London Mathematical Society* **8**: 71–79.