Assimilation of weather radar and binary ubiquitous sensor measurements for quantitative precipitation estimation

David J. Hill

ABSTRACT

Assimilation of data from heterogeneous sensors and sensor networks is critical for achieving accurate measurements of environmental processes at the time and space scales necessary to improve forecasting and decision-making. Owing to different measurement accuracies and types of spatial and/or temporal measurement support of the component sensors, it is often unclear how best to combine these data. This study explores the utility of ubiquitous sensors producing categorical wet/dry rainfall measurements for improving the resolution of areal quantitative precipitation estimates through fusion with weather radar observations. The model developed in this study employs a Markov random field model to compute the probability of rainfall at sub-grid pixels. These likelihoods are used to 'unmix' the cell-averaged rainfall rate measured by the radar. Simulation studies using synthetic and known rainfall fields reveal that the model can improve remotely sensed quantitative rainfall intensity measurements by 40% using networks of ubiquitous sensors with a density of 56 sensors per square kilometer, and for denser networks, the accuracy can increase by as much as 50%.

Key words | Bayesian methods, data assimilation, data fusion, Markov random fields, quantitative rainfall estimation, ubiquitous sensing

David J. Hill
Department of Geography and Environmental
    Studies,
Thompson Rivers University,
Kamloops, BC,
Canada V2C 0C8
E-mail: dahill@tru.ca

INTRODUCTION

Researchers have recently suggested 'smart' urban infrastructure to mitigate social and economic risks posed by extreme weather events and to improve the sustainability of urban centers (Boyle *et al.* 2010; Brown 2010). Smart infrastructure is enabled by real-time sensing, modeling, optimization, and actuation, which permit the infrastructure to reason about the world around it and to autonomously reconfigure in order to meet changing needs. For example, researchers are exploring the use of smart sewer systems that combine real-time measurements of urban rainfall patterns with runoff models and remote actuated control structures to optimize the conveyance capacity of the sewer network during wet weather (Pleau *et al.* 2001; Schutze *et al.* 2004; Colas *et al.* 2005; Pleau *et al.* 2005; Fradet *et al.* 2011).

A significant barrier, however, to the design and deployment of smart infrastructure is the inability of current sensing technology to characterize the environment at sufficient space and time resolutions to support the forecasts necessary to enable autonomous near-real-time decision-making via predictive control algorithms. In the case of urban flooding, extrapolation from a recent study (Berne *et al.* 2004) indicates that the minimum spatial resolution for rainfall observations is on the order of 0.5 km^2. Unfortunately, this observational resolution is not feasible with traditional sensing technologies. Telemetered tipping bucket rain gauges and weather radar are the most popular methods for real-time sensing of rainfall. Rain gauges provide direct measurement of rainfall on the land surface; however, few rain gauge networks exist that are dense

enough to quantify the spatial variability of rainfall (Grimes *et al.* 1999; Sun *et al.* 2000; Xiaoyang *et al.* 2003; Yilmaz *et al.* 2005). Furthermore, rain gauge data contain inaccurate measurements that can lead to significant modeling errors (Steiner *et al.* 1999; Upton & Rahimi 2003). If these data are to be used for real-time management of smart infrastructure, these errors need to be identified and mitigated as they are made, necessitating automated quality control that relies solely on prior observations (Hill 2013). Greater spatial density of precipitation measurements can be achieved using weather radar, a robust, albeit expensive, sensing technology that provides indirect measurements of rainfall on the land surface by measuring the backscattering of microwaves caused by hydrometeors (and other objects) in a finite volume of the atmosphere and relating that quantity to the expected ground-based rainfall rate (Crum & Alberty 1993; Crum *et al.* 1998). Weather radar uses an active sensor that emits microwaves, which are reflected off hydrometeors in an atmospheric volume at some altitude above the land surface. Radar-rainfall estimates are derived from these measurements using a mathematical model that relates the strength of microwave scattering caused by hydrometeors to the expected ground-based rainfall rate (Marshall & Palmer 1948; Smith & Krajewski 1993; Baeck & Smith 1998; Steiner & Smith 1998). However, factors such as radar calibration (Austin 1987), signal attenuation (Bringi *et al.* 2011; Rico-Ramirez 2012), ground clutter (Islam *et al.* 2012), anomalous propagation (Collier 1996; Rico-Ramirez *et al.* 2007), variation in the vertical reflectivity of precipitation, and significant spatiotemporal variation of the relationship between scattering and rainfall rate (Smith & Krajewski 1993; Steiner & Smith 1998) can lead to considerable uncertainty in radar-rainfall estimates. For this reason, radar-rainfall estimates are often adjusted using ground-based rain gauge data to improve their accuracy (Seo & Smith 1991; Smith & Krajewski 1991; Seo *et al.* 1999; Velasco-Forero *et al.* 2009). However, even when combined, the resolution achievable by these methods is still too coarse to support many physics-based models that relate the rainfall process to actionable forecasts, such as flood depths (Seo & Krajewski 2010).

Because of their ubiquity, small, low-cost sensors that are embedded in everyday consumer products have begun to be explored for environmental monitoring (Ferster &

Coops 2013; Hill *et al.* 2014). While these sensors are not specifically designed to monitor the environment, the data they collect can be repurposed using models to provide quantitative measurements of environmental variables at the location of the sensor. For example, sensors embedded in wireless communication devices (Leijnse *et al.* 2007; Zinevich *et al.* 2008) and vehicle-based automatic windshield wipers (Haberlandt & Sester 2010) have been explored for rainfall measurements. Although ubiquitous sensors may produce measurements with lower accuracy than dedicated sensors, it is expected that their pervasiveness in urban environments will result in significantly greater spatial resolution than networks of dedicated environmental sensors. Furthermore, by opportunistically integrating sensors introduced into the environment by private individuals, the improved observational capacity of this new sensing paradigm will be achieved without the large-scale deployment of new dedicated environmental sensors by a centralized agency, such as the United States National Weather Service (NWS).

As computational and communication technology becomes smaller and cheaper, it is expected that these devices will become increasingly ubiquitous and will possess increasing communication capabilities. For example, a survey of new vehicles shows that most automakers are beginning to equip vehicles in both their regular and luxury brands with rain-sensing windshield wipers. Rain-sensing wipers are also sold as an aftermarket upgrade for unequipped vehicles (http://www.raintracker.com). These wipers employ a sensor to determine the presence of rain, and based on the vehicle speed, they actuate and adjust the speed of the wipers. At the same time, most new vehicles are sold with integrated global positioning system (GPS) chips and communication capabilities, which are used for navigation and emergency response, respectively.

This study investigates the use of ubiquitous sensors for quantitative rainfall measurement using a model that would be feasible for automated real-time applications. However, unlike the work of Leijnse *et al.* (2007), Haberlandt & Sester (2010), and Zinevich *et al.* (2008), this work approaches the problem as a field estimation problem that can be achieved through the fusion of measurements from a large network of heterogeneous sensors that includes both ubiquitous and dedicated rainfall sensors.

This approach capitalizes on the spatial resolution of the ubiquitous sensors and the quantitative accuracy of the dedicated sensors, and eliminates the necessity for the precise calibration of individual vehicle-based sensors – a significant logistical challenge. In this study, the ubiquitous sensors are assumed to provide binary measurements (i.e. wet/dry) of the current conditions. This assumption is made for two reasons. First, generating quantitative measurements of rainfall from ubiquitous sensors requires the tedious process of applying and calibrating a relationship that transforms the raw sensor measurement from each type of ubiquitous sensor into a quantitative rainfall estimate, whereas generating binary measurements from these sensors is expected to be significantly easier. Second, there are already a wide variety of sensors deployed in the environment that can be leveraged to provide this type of qualitative rainfall measurement, including rain-sensing windshield wipers, traffic cameras, road weather information systems, and irrigation control systems.

The measurement fusion method developed in this study employs a Markov random field (MRF) to represent the variability of rainfall intensity at scales smaller than the radar resolution. The use of an MRF model for fusing binary (wet/dry) data with weather radar observations has previously been suggested by the author (Hill & Farzan 2012). This preliminary work laid a framework for the fusion process but did not explore how best to set the model parameters, nor did it quantitatively explore the performance of the suggested model. In this work, a new Bayesian solution to the MRF model is proposed that permits the use of noisy measurements from sparse binary sensors to accurately reconstruct a map of rainy versus not-rainy regions. This map is then used to redistribute the spatially averaged rainfall rate measured by the radar to concentrate it only in the rainy regions, leaving the dry regions with a zero rainfall rate. The next section describes this model in detail. The model's ability to reconstruct a synthetic rainfall field is then explored, particularly with respect to its parameters, requirements for ubiquitous sensor density, and field estimation accuracy. Synthetic data are used to provide a ground truth against which to evaluate the model performance and to eliminate spurious errors caused by the stochastic behavior of physical sensor

hardware and telemetry. Finally, conclusions and future work are discussed.

METHODS

The essence of the fusion method developed in this study is that binary measurements of the rainfall made by ubiquitous sensors can be used to identify the pattern of rainfall within each grid cell of a radar image and that this pattern can be used to concentrate the rainfall into the wet portions of the radar cell, such that the cell average remains the same, but the rainfall rate is zero in the dry portions of the cell. To accomplish this, each grid cell in a radar image is subdivided into smaller pixels at the spatial scale of interest. To facilitate this discussion, the term 'grid cell' will hereafter be used to refer to the coarse-resolution grid-square of the radar image, while 'pixel' will refer to the fine-resolution grid-square at which the rainfall field is reconstructed. Furthermore, when referring to the binary state of the rainfall process, a value of 0 will be used to represent dry conditions, and a value of 1 will be used to represent wet conditions.

Estimation of the binary rainfall pattern at the sub-grid scale will be performed using an MRF model. MRFs are statistical models of a set of spatially correlated random variables that describe the relationship of neighboring variables using a model that relies only on a finite set of N neighbors (i.e. a model that has an N-order Markov property) (Besag 1974; Isham 1981; Ripley 1988). The Markov property permits the spatial field to be subdivided using a relationship that defines whether or not two pixels are neighbors. A random field has the Markov property with respect to a particular neighborhood structure if the pixel value is conditionally independent given its neighborhood. Therefore, the probability that a pixel is wet or dry, given values at all other pixels, depends only on the wet/dry configuration in its neighborhood. Following accepted practice in image reconstruction, wet/dry patterns are modeled using an autologistic model (Besag 1974). In this model, the neighborhood of a pixel is defined to be the set of all pixels immediately horizontally, vertically, and diagonally adjacent, as shown in Figure 1. Pairs of adjacent pixels within the neighborhood are referred to as cliques and can be divided into two categories: Type 1 refers to horizontally
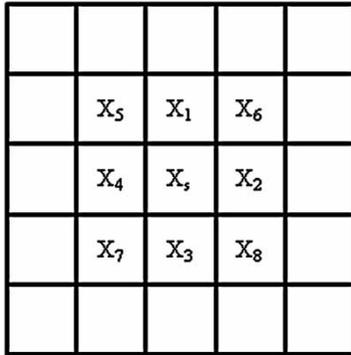
Figure 1 | Neighborhood of pixel X_s.

and vertically adjacent pixels, and Type 2 refers to diagonally adjacent pixels.

Using this notation for neighboring pixels, the autologistic model for the conditional probability of the binary state ($0 =$ dry and $1 =$ wet) of the rainfall process ($X_{i,j}$) can be expressed as

$$P(X_{i,j}|N) =$$

$$\frac{\exp\left\{-\beta_1 \sum_{k=1}^{4} I(X_{i,j} \neq N_k) - \beta_2 \sum_{k=5}^{8} I(X_{i,j} \neq N_k)\right\}}{\sum_{X_{i,j}=0}^{1} \exp\left\{-\beta_1 \sum_{k=1}^{4} I(X_{i,j} \neq N_k) - \beta_2 \sum_{k=5}^{8} I(X_{i,j} \neq N_k)\right\}}$$

$$N = \{X_{i-1,j}, X_{i+1,j}, X_{i,j-1}, X_{i,j+1}, X_{i-1,j-1}, X_{i-1,j+1},$$
$$X_{i+1,j-1}, X_{i+1,j+1}\} \tag{1}$$

where N is the set of pixels in the local neighborhood of pixel (i, j), β_1 and β_2 are parameters, and $I(\cdot)$ is an indicator function operator that returns 1 when the inequality is true and 0 when the inequality is false. The parameters β_1 and β_2 describe the dependance between a pixel and its Type 1 and Type 2 neighbors, respectively. Values close to zero reflect spatial fields where pixels are independent of the value of the pixels in their neighborhood N, whereas when β_1 and β_2 are strictly greater than 0, $X_{i,j}$ is most likely to take the value of the majority neighboring class. By assigning different values to the parameters β_1 and β_2, the model can represent anisotropic conditions. If β_1 is greater than β_2, then similarity in the vertical/horizontal directions (i.e. along the principal axes of the image) is weighted more heavily, whereas if β_1 is less than β_2, then similarity in the diagonal directions is weighted more heavily. Equation (1) can be used to represent either the probability that pixel (i,j) is equal to 0

($P(X_{i,j} = 1|N)$) or 1 ($(P(X_{i,j} = 1|N))$). In this research it was used to calculate the probability that the pixel takes the value 0, and the probability that the pixel takes the value 1 was computed by complementation (i.e. $P(X_{i,j} = 1|N) = 1 - P(X_{i,j} = 0|N)$).

If not all of the values of the neighbors of pixel (i,j) are known, then Equation (1) can be modified to use likelihood weighting to calculate the contributions of the unknown neighbors' values

$$P(X_{i,j} = 0|Z) =$$

$$\frac{\exp\left\{-\beta_1 \sum_{k=1}^{4} P(X_{i,j} \neq \hat{N}_k|Z) - \beta_2 \sum_{k=5}^{8} P(X_{i,j} \neq \hat{N}_k|Z)\right\}}{\sum_{X_{i,j}=0}^{1} \exp\left\{-\beta_1 \sum_{k=1}^{4} P(X_{i,j} \neq \hat{N}_k|Z) - \beta_2 \sum_{k=5}^{8} P(X_{i,j} \neq \hat{N}_k|Z)\right\}}$$
$$\tag{2}$$

where Z is the set of known values of X, and $\hat{N}$ is the set of neighbors of $X_{i,j}$, which may represent estimates of the true value, if the value of that neighbor is not known. In the event that the value of a neighboring pixel N_k is known, then $P(X_{i,j} = N_k|\hat{N}) = 1$, and in the case that all of the values of the pixels in $\hat{N}$ are known, then Equation (2) reduces to Equation (1). Thus, this method accounts for a partial contribution from unknown, but estimated, pixels that is equivalent to the probability that those pixels are different than the value of $X_{i,j}$. For images in which only a few pixel values are known, then, it is likely that there will be pixels for which none of the neighbors are known. In this case, the values of these neighbors must be estimated in turn from their own neighbors. This expansion must continue until at least one neighbor with a known value exists in the neighbor set of the pixel to be estimated. In this way, although Equation (2) represents only a first-order spatial Markov process, the effects of known values of pixels more distant than immediate neighbors can impact the value of a particular pixel. Thus, to solve Equation (2) for images with only a few known pixel values, such as those encountered in this research, the conditional probabilities $P(X_{i,j} \neq \hat{N}_k|N)$ are defined by computing Equation (2) iteratively over all of the pixels in the field for the case of $P(X_{i,j} = 0|N)$ until the conditional probabilities for each pixel in the field stabilized. The complement of this probability $P(X_{i,j} = 1|N)$ was then be computed as $1 - P(X_{i,j} = 0|N)$. When a pixel with unknown value falls

on the boundary of the field (i.e. there are fewer than eight neighboring pixels), the pixels outside the boundaries of the field are treated as pixels with unknown value. This is accomplished by providing a 1-pixel-deep buffer around the field being estimated.

In real networks of ubiquitous sensors, it is expected that the sensors will malfunction, resulting in erroneous and misleading data. Thus, Bayes' Theorem was employed to pose the problem as one of estimating the true random field (X) from an incomplete set of noisy measurements (Y).

$$P(X_{i,j}|Y_{i,j}) = \begin{cases} \dfrac{P(Y_{i,j}|X_{i,j})P(X_{i,j}|\hat{Z})}{P(Y_{i,j})} & \text{if } O_{i,j} = \text{True} \\[2ex] P(X_{i,j}|\hat{Z}) & \text{if } O_{i,j} = \text{False} \end{cases} \quad (3)$$

where O is a matrix of the same size as X that indicates whether a noisy measurement of the true state of pixel (i,j) exists in the set Y, $\hat{Z}$ is an estimate of the values of the observed pixels, $P(X_{i,j}|\hat{Z})$ is the probability that X at location (i,j) takes a specific value given set of estimated pixel values $P(Y_{i,j}|X_{i,j})$, $P(Y_{i,j}|X_{i,j})$ is a probability density table that defines the observation model of the sensors, and $P(Y_{i,j})$ is the normalization constant. $P(X_{i,j}|\hat{Z})$ can be solved using the iterative solution to Equation (2) described above, however, in the first iteration, $\hat{Z}$ is estimated to be equal to Y (the set of noisy observations), whereas in subsequent iterations, $\hat{Z}$ is estimated using the maximum a posteriori estimates of X at the measurement locations, which is calculated using Equation (3).

Equation (3) can be viewed as the 2-D equivalent of a hidden Markov model, which blends the information from the set of incomplete noisy measurements and the 2-D spatial MRF model. As in the case of Equation (2), in the case of sparsely observed images, Equation (3) must also be solved iteratively to allow the information contained in the observations to propagate through the image based on the first-order Markov property of the spatial model. This process is similar to the iterated conditional modes (Besag 1983) method for reconstructing images corrupted by noise, except that unlike iterated conditional modes, the method presented in this paper is applicable to incompletely observed, noisy images.

Once the spatial probability distribution of the binary random variable is computed for all pixels in the image,

then the rainfall rate in each pixel can be computed as

$$R_{i,j} = \bar{R}\frac{N \times P(X_{i,j} = 1|Y)}{\sum_{i=1}^{N} P(X_{i,j} = 1|Y)} \quad (4)$$

where X is the binary random variable indicating not-raining/raining conditions in pixel $(i.j)$, $\bar{R}$ is the grid-block-averaged precipitation rate measured by the radar, $P(X_{i,j} = 1|Y)$ is the probability that it is raining in the ith pixel given the observations, and N is the number of pixels in the radar block.

The ubiquitous sensing model for estimating rainfall intensity fields created by combining Equations (2)–(4) will hereafter be referred to as the ubiquitous rainfall sensing (URS) model.

CASE STUDIES

To evaluate the performance of the URS model, experiments were conducted to explore the model's ability to estimate a binary field and to use that binary rainfall field to downscale simulated radar-rainfall estimates. These experiments were performed using both synthetic and real-world rainfall fields.

Synthetic rainfall field

A synthetic rainfall field was used to explore the effects of the model parameters and spatial discretization on the accuracy of the downscaled fields. This field was created using the following analytical function:

$$\begin{aligned} Z(x,y) &= \max\left[\begin{array}{l} 3 \times (1-a)^2 \times \exp\left(-a^2 - (b+1)^2\right) \\ 0, \ -10 \times \left(\dfrac{a}{5} - a^3 - b^5\right) \times \exp\left(-a^2 - b^2\right) \\ -\,{}^1\!/_3 \exp\left(-(a+1)^2 - b^2\right) \end{array} \right] \\ a &= 6 \times \frac{x}{4} - 3 \\ b &= 6 \times \frac{y}{4} - 3 \end{aligned} \quad (5)$$

where x and y are the easting and northing coordinates of the location in the range [0,4]. By using an analytical function, a 'ground-truth' field can be created at any spatial resolution without introducing artifacts from the

interpolation of real-world rain gauge data. This function was used to generate a 4 × 4 km field at 1 m resolution to which a threshold was applied to remove negative values. To simulate the radar-rainfall measurement of the synthetic field, it was divided into 16 1 × 1 km radar blocks, each of which was assigned the average value of the sub-grid pixels. The synthetic field is illustrated in Figure 2. From this figure, it can be seen that the field is composed of a continuous region of rain that covers approximately one-quarter of the total area. This rainy region has three areas of rainfall concentration, one in the north with the peak rainfall intensity, and two that are side by side in the center of the field. A ground-truth binary wet/dry field is created by assigning wet/dry measurements according to whether or not Equation (5) is greater than zero at each pixel centroid (indicating rainy conditions). A simulated radar measurement of the ground-truth rainfall field is created by averaging the 16 1 × 1 km blocks composing the true field. The simulated radar-rainfall field is illustrated in Figure 3.

Real-world rainfall fields

To evaluate the performance of the URS model for reconstructing real-world rainfall fields, data from the United States Department of Agriculture, Agricultural Research Station Goodwin Creek Experimental Watershed were used. The Goodwin Creek Watershed covers a region of
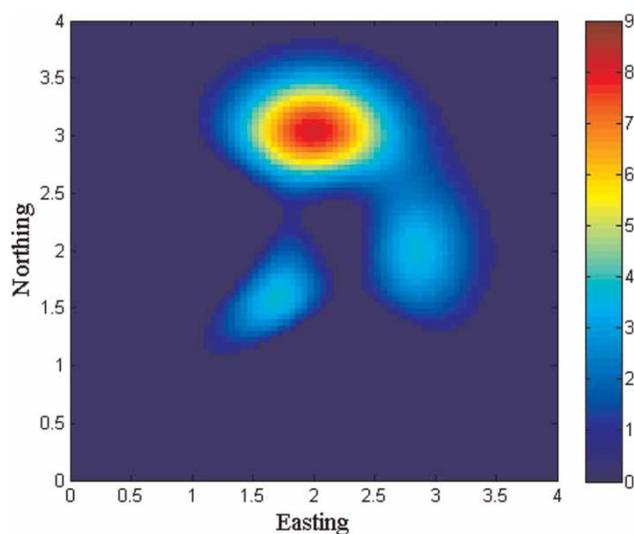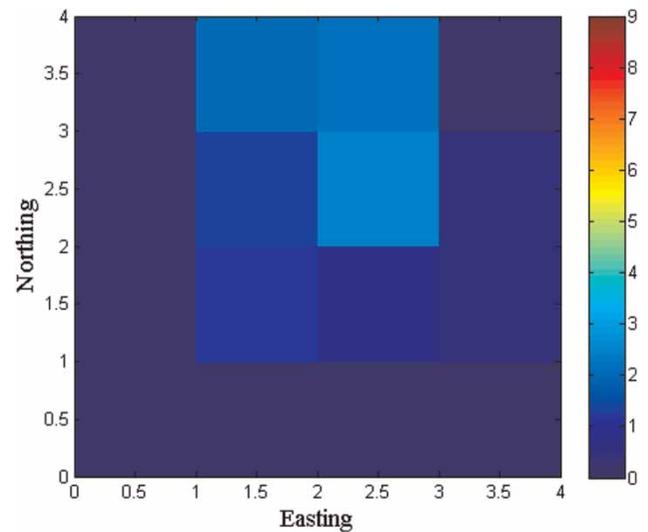

Figure 3 | Simulated radar-rainfall observation of the synthetic ground-truth rainfall field.

approximately 21 km^2 in northern Mississippi and is monitored by a network of 32 weighing precipitation gauges. This rain gauge network was selected because it was the densest rain gauge network for which data could be obtained. Rainfall fields at 10 m resolution were created for a 4 × 4 km region at the southwestern end of the watershed (the densest part of the network) using inverse distance weighted averaging to interpolate 10-min rainfall intensities from the rain gauges. Inverse distance weighting was chosen for interpolating these fields because of its long history of use for this purpose (Garcia et al. 2008), including its use by the NWS River Forecast System for the operational estimation of precipitation over much of the United States (National Weather Service 1999). Radar-rainfall estimates and ubiquitous sensor measurements were then simulated from these rainfall fields. Simulated radar and ubiquitous sensor data were used to focus the analysis on the model behavior without having to detangle errors introduced by the model and those caused by the stochastic behavior of physical sensor hardware and telemetry. To represent radar-rainfall estimates of these fields, they were aggregated into 16 1 × 1 km radar blocks, each of which was assigned the average value of the sub-grid pixels. A ground-truth binary wet/dry field was created by assigning wet conditions to pixels with non-zero rainfall intensities and dry conditions to pixels with zero rainfall.

Figure 2 | Synthetic ground-truth rainfall field.

MRF parameter values and spatial discretization

The MRF model (Equation (2)) used to estimate the binary wet/dry map requires that the spatial domain be discretized into sub-grid pixels that are each assigned a probability of experiencing wet weather. Thus, as the pixel size decreases, the method can more accurately resolve boundaries between wet and dry regions. If each observed pixel is monitored by one ubiquitous sensor, then as the pixel size decreases, the number of ubiquitous sensors must increase to observe the same percentage of the spatial domain. In addition, as the pixel size decreases, the number of pixels inside each radar grid block increases, and thus the computational complexity of solving the MRF increases. Furthermore, decreasing the pixel size for a given coverage level increases the number of unobserved pixels separating observed pixels. Since the MRF model used in this study only considers immediately adjacent pixels, this will result in the MRF relying more heavily on the autologistic neighborhood model to interpolate between observed locations. The autologistic parameters β_1 and β_2 control the strength of the similarity between neighboring pixels. The parameter β_1 defines the strength of the similarity between vertical/horizontal neighbors, and the parameter β_2 defines the strength of the similarity between diagonal neighbors. Since it is difficult to quantify anisotropy *a priori* (i.e. before observing a rainfall event), as is necessary for real-time applications, these parameters will be given the same value, which will hereafter be referred to as β, with no subscript. The larger the magnitude of this parameter, the more likely neighboring pixels will have the same value. Thus, as the pixel size decreases, the model may require a larger value of β in order for the influence of observed pixel values to propagate to other parts of the image.

To explore the sensitivity of the URS model to the pixel size, coverage level, and β value, the synthetic field was generated using pixel sizes of 25, 50, 100, 250, and 500 m. A maximum pixel size of 500 m was selected because this is the minimum resolution suggested by Berne *et al.* (2004) for capturing rainfall variability relevant to urban flooding. These fields were used to simulate binary ubiquitous sensor measurements by randomly sampling the ground-truth binary rainfall field to produce a map with levels of observational coverage ranging from 10 to 80%, in 5% increments. One-hundred realizations of the simulated ubiquitous sensor observations were made at each domain coverage ratio, in order to average out the effects of favorable/unfavorable sensor configurations. The URS model was then applied to the cell-averaged field and the binary ubiquitous sensor observations using β values that ranged between 0.2 and 2, in increments of 0.2. To focus the results on the effects of the autologistic model parameters and the discretization, no erroneous measurements were present in the simulated ubiquitous sensor measurements, and the probability that a ubiquitous sensor measurement accurately represented the wet/dry state of the pixel was set to 100% (i.e. in Equation (3), $P(Y = 0|X = 1) = P(Y = 1|X = 0) = 0.0$ and $P(Y = 1|X = 1) = P(Y = 0|X = 0) = 1.0$). Thus, there are no misleading/erroneous measurements expected from the ubiquitous sensors and no errors present in their measurements. For each combination of coverage, β value, and discretization, the mean squared error (MSE) of the URS model-based rainfall field estimation was compared to the MSE of the cell-averaged field.

The ratios of the URS MSE to the cell-averaged MSE for these combinations of parameter values are illustrated in Figure 4. This figure shows that for the smallest (25 m) pixel size, the MRF-based reconstruction error is between 50 and 78% of the cell-averaged error. Thus, the MRF-based method improves the field estimation error, relative to the cell-averaging method, by 22–50%. For this pixel size, an MSE reduction of 50% was achieved for most of the combinations of β and observational coverage. Only when the observational coverage was very small (10%) or when the β parameter conditioned unobserved pixels very weakly ($\beta < 0.3$) or very strongly ($\beta > 0.8$) on neighboring values did the MRF-based method fail to achieve a 50% reduction in MSE over the cell-averaged representation. However, even in these cases, the MRF-based reconstruction had an MSE of at least 22% less than the MSE of the cell-averaged representation. As the pixel size was increased, the maximum MSE reduction remained around 50%, except in the case of the largest pixel size (500 m). However, the region of parameter space where this reduction was achieved became smaller. For β-values between 0.4 and 0.6, as the pixel size increases, the observational coverage must also increase to achieve the maximal MSE reduction of 50%.
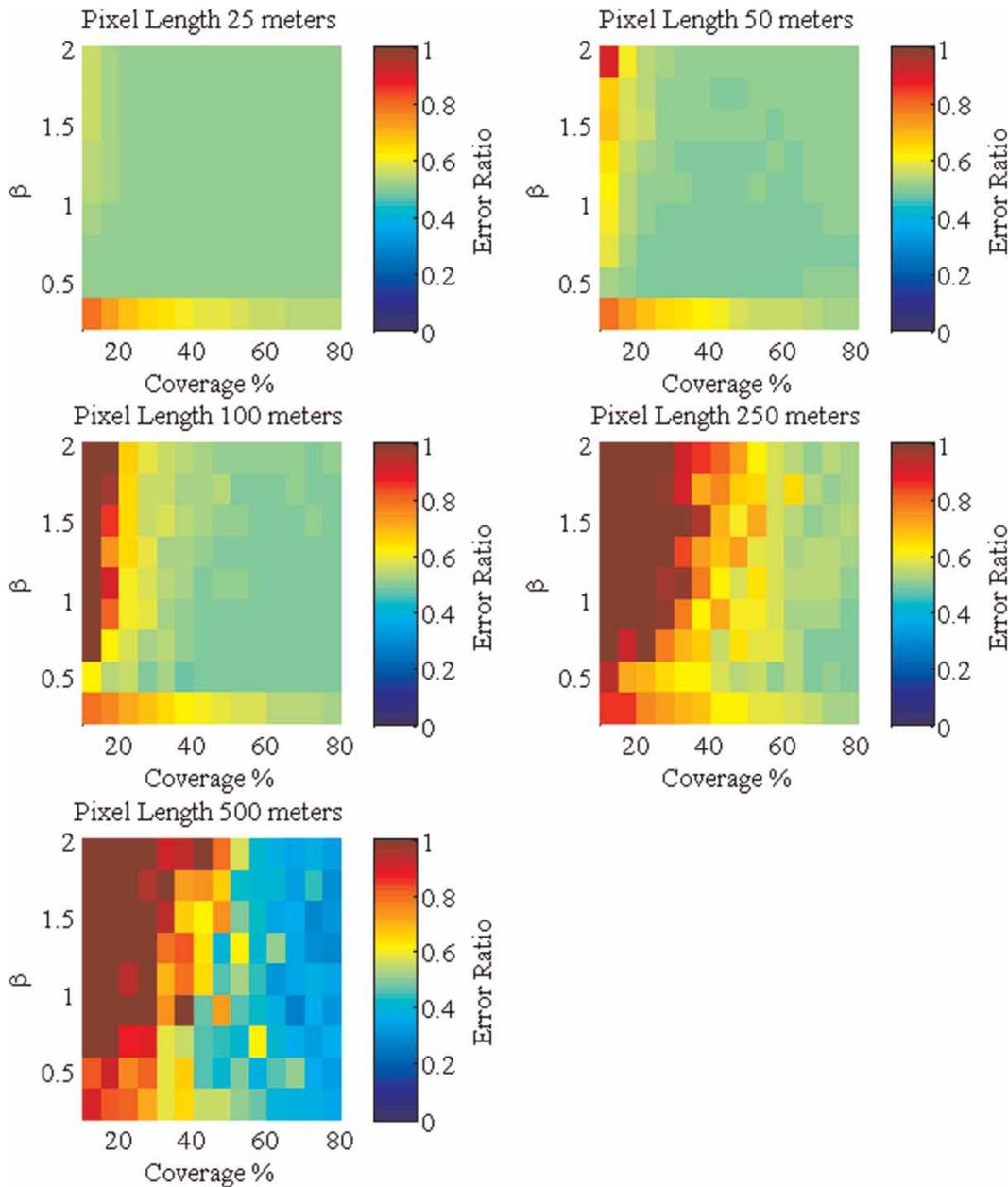
Figure 4 | Sensitivity of URS model to pixel size, coverage of binary observations, and β parameter. The error ratio illustrated is the MSE of the MRF reconstruction divided by the MSE of the block-averaged reconstructions.

Also of note is the sharp decrease in error ratio that occurs when β increases from 0.4 to 0.6, whereas for a given pixel size, the error ratio appears to increase more smoothly as β increases from 0.6 to 2.2. This feature is most easily seen in the case of 50-m pixels but is present in all six pixel sizes evaluated. This behavior is likely due to the use of a spatial autologistic model. In such a model, if pixel (i,j) is

surrounded by eight neighbors of identical value v, then $P(X_{i,j} = v|N) = 0.5$ for $\beta = 0$. As β increases, this probability increases asymptotically towards unity according to the logistic function, where $P(X_{i,j} = v|N)$ increases rapidly with increasing β until about $\beta = 0.4$, at which point each additional increase in β results in less and less of an increase in $P(X_{i,j} = v|N)$. Thus, the model will be very sensitive to

values of $\beta \leq 0.4$. For values of β greater than 0.6, $P(X_{i,j} = v|N)$ will be greater than 0.99; thus, the region $0.4 \leq \beta \leq 0.6$ provides the 'sweet spot' for this parameter, where it is neither too sensitive nor too insensitive to changes in its value. Although these numeric results consider only the case where the unknown pixel is completely surrounded by equivalently valued neighbors, they provide the intuition that for values of $\beta \leq 0.4$, small changes in the value of β will result in significant changes to the model output, whereas for $\beta \geq 0.6$, the value of the unknown pixel will be driven more by the diversity of the values of the neighboring pixels than by the value of β. Finally, Figure 4 suggests that pixel lengths of 500 m provide the greatest reduction in error when coupled with spatial coverage greater than 70%. This behavior is most likely caused by the need for fewer steps of the first-order Markov model to propagate the influence of an observation to other regions of the image. For example, in the case of 50-m pixels, the Markov model would have to be applied sequentially 20 times to propagate the influence of a single measurement to a radius of 1,000 m, whereas the model would only need to be propagated twice if the pixel size were 500 m. This result may suggest that a pixel size of 500 m would be preferable to smaller pixel sizes; however, at this pixel size, a large percentage of the field must be observed to achieve this error reduction. These results suggest that if a radar cell is subdivided into a large number of small pixels (25–50 m), the method is relatively insensitive to the parameter settings, and if a β-value between 0.4 and 0.6 is used, then the maximal improvement in field reconstruction accuracy can be realized for any level of observational coverage. Thus, for the remainder of the experiments, a discretization of 50 m and a β-value of 0.5 was used.

Ubiquitous sensor malfunction

In real networks of ubiquitous sensors it is expected that the sensors will malfunction and report erroneous and misleading measurements. To determine how resilient the URS model is to erroneous ubiquitous sensor measurements, the ability of the MRF to reconstruct the binary wet/dry rainfall field using a partial set of noisy observations was explored. Noisy ubiquitous sensor

measurements were simulated by sampling the ground-truth binary wet/dry field to produce a map with a specific observation density (as previously described). This map was then corrupted with noise from a simulated sensor malfunction. In this simulation, when a sensor malfunctioned, it produced a measurement with a value opposite that of the ground-truth field. One-hundred realizations of ubiquitous sensor observations were generated for observational coverage values of 10, 20, 30, 40, and 50% and sensor malfunction rates of 2, 5, 10, 15, and 20%. For each combination of observational coverage and sensor malfunction rate, the false positive (where the MRF model predicts rain when the pixel is dry) and false negative (where the MRF predicts dry when the pixel is wet) rates were computed for the binary wet/dry field reconstruction using two versions of the URS model. The first version (URS-noerr) does not account for the possibility of misleading ubiquitous sensor measurements, although they are expected to occur. In this case, the following values are used in Equation (3): $P(Y = 0|X = 1) = P(Y = 1|X = 0) = 0.0$ and $P(Y = 1|X = 1) = p(Y = 0|X = 0) = 1.0$. In the second version (URS-5%err), the URS model is specified to expect a 5% error rate in the ubiquitous sensor measurements; thus, $P(Y = 0|X = 1) = P(Y = 1|X = 0) = 0.05$ and $P(Y = 1|X = 1) = P(Y = 0|X = 0) = 0.95$.

Figure 5 illustrates the false-positive/false-negative rates for the URS-noerr (left column) and URS-5%err (right column). This figure shows that for low sensor malfunction rates, the accuracy of the URS-noerr model improves with increasing observational coverage, whereas for high sensor malfunction rates, the accuracy of the URS-noerr model declines with increasing observational coverage. This result is due to the increasing number of misleading observations in the more highly observed cases, which mislead the URS-noerr because it is so strongly conditioned on the observations. This result implies that the sweet spot for estimating a field from incomplete noisy images using the URS-noerr depends on balancing the observational coverage with the sensor accuracy, which would be challenging in real-world scenarios. On the other hand, the URS-5%err is well behaved: the error decreases with increasing coverage and increases with increasing sensor error rates. It is important to note that this behavior was achieved without tuning the
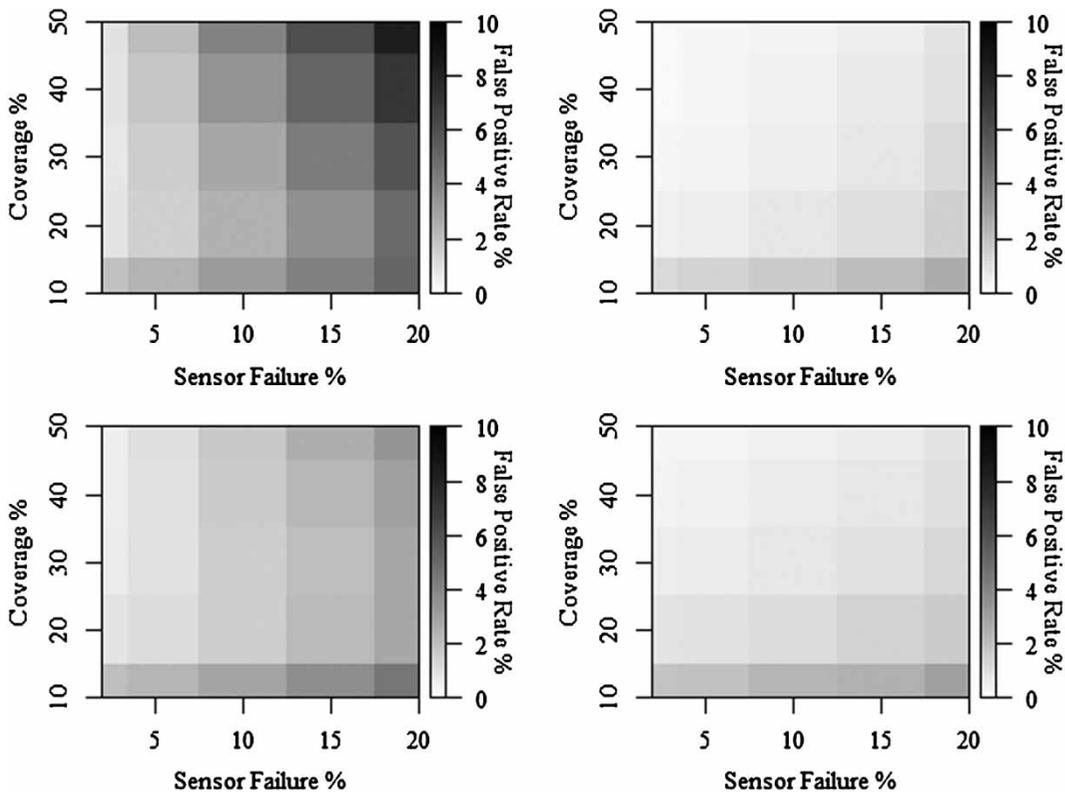
Figure 5 │ Expected false-positive (top row) and false-negative (bottom row) rates of URS-noerr (left column) and URS-5%err (right column) for estimating the ground-truth binary wet/dry field as a function of the level of observational coverage and rate of sensor malfunction.

error rate in the URS-5%err model to the actual error rate in the observations (the modeled error rate is always 5%), indicating that the method is insensitive to this parameter and that the performance of the model could be enhanced if information were available to suggest an appropriate expected error rate to use for the model parameter.

Synthetic rainfall field reconstruction accuracy

The previous two experiments have shown that the URS model is relatively insensitive to the model parameterization and that with a 50-m pixel size, the following parameter values provide good overall performance: $\beta = 0.5$, $P(Y = 0|X = 1) = P(Y = 1|X = 0) = 0.05$, and $P(Y = 1|X = 1) = P(Y = 0|X = 0) = 0.95$. Using these parameters, the impact of ubiquitous sensor coverage on the URS model's accuracy for reconstructing the ground-truth rainfall field was explored. The true binary image was randomly sampled to generate 100 realizations of the ubiquitous sensor measurements for observational coverage levels

that varied from 2 to 95%. The URS model was then used to estimate the true field from these observations. The accuracy of the estimated fields is quantified in terms of their MSE. Figure 6 illustrates the relationship between URS-based estimation accuracy and level of observational coverage and compares the URS MSE values to the MSE incurred by using the 1×1 km cell-averaged field to represent the ground-truth field at 50 m resolution. The MSE of the cell-averaged estimate is 1.3. As can be seen in this figure, there is a steep decline in the URS error when the observational coverage is low, which transitions to a shallow decline at a coverage level of around 15%. For observational coverage levels smaller than 8%, the URS model MSE is greater than the cell-averaged MSE; however, at a coverage level of 14%, when the point of diminishing returns is reached, the URS model MSE is 0.79, representing a 40% reduction in MSE over the cell-averaged estimate. Given the pixel size of 50 m, 14% coverage within a 1-km radar grid cell corresponds to only 56 observed pixels, requiring a ubiquitous sensor density of only 56 per km^2. Although the point of diminishing
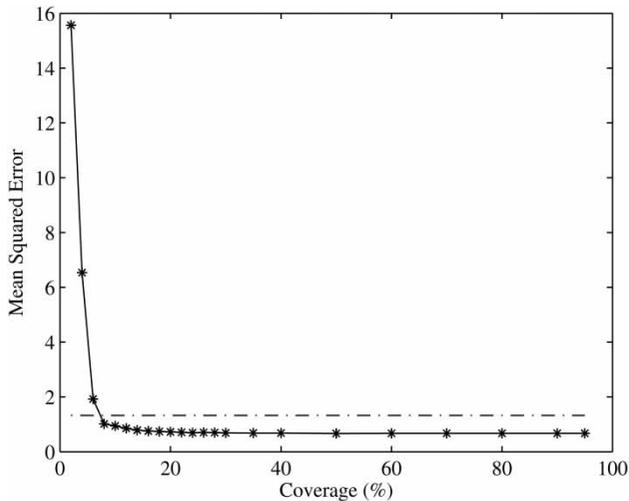
Figure 6 | MSE of the URS model for reconstructing the synthetic rainfall field for different levels of observational coverage. The horizontal dashed line indicates the MSE incurred by estimating the rainfall field with the simulated radar-rainfall field.

Figure 7 | Rainfall field estimated by the URS model using ubiquitous sensors measuring the wet/dry conditions in 15% of the 50 m pixels. The MSE of this estimate is 0.66, 50% less than the MSE incurred by estimating the rainfall field with the simulated radar-rainfall field.

returns is reached around 15% coverage, the MSE of the URS model continues to decline steadily with increasing observational coverage. For example, with 95% coverage (380 sensors per km^2), the URS model exhibits an MSE of 0.67, a 49% reduction in error over the cell-averaged representation. This behavior is desirable in a ubiquitous sensor network because the deployment of ubiquitous sensors is driven by individual user choices which cannot be controlled. The results of this experiment suggest that the observational accuracy of the URS model will always increase as more ubiquitous sensors enter the environment and thus that the population of ubiquitous sensors need not be managed to achieve observational optimality. Furthermore, the ubiquitous sensor density to achieve a meaningful increase in accuracy is quite modest, with a 40% increase in accuracy being achieved with only 56 sensors per km^2.

Figure 7 illustrates the improvement in quantitative rainfall estimation that can be achieved using the URS model. A visual inspection of Figure 3 reveals that the simulated radar-rainfall measurement significantly smears the ground-truth rainfall field, resulting in a complete loss of the shape of the rainy region and an under/over-prediction of the rainfall intensity in regions of high/low-intensity rainfall. The smearing is caused by spatial (cell) averaging the ground-truth rainfall field by the radar. In comparison, the rainfall field, shown in Figure 7, estimated by combining
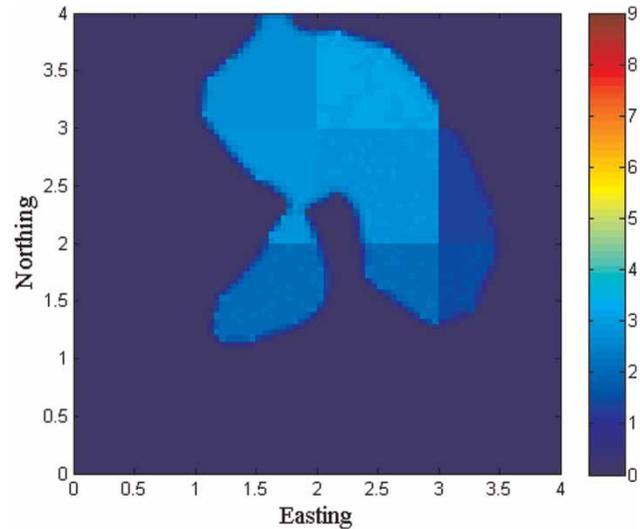
the radar and ubiquitous sensor measurements using the URS model, captures the shape of the wet area and mitigates some of the under/over-estimation in the high/low-intensity regions of the rainfall field. The MSE of the URS model estimate is 0.66: 50% less than the MSE incurred by estimating the rainfall field with the simulated radar-rainfall field. The reason that this accuracy improvement is greater than suggested by Figure 6 is because the data plotted in Figure 6 is averaged over many realizations of ubiquitous sensor locations, whereas the values stated here reflect the accuracy associated with a single realization of the ubiquitous sensor locations that happens to be a bit better than average. This comparison suggests that even though the ubiquitous sensors are only providing very simple categorical wet/dry measurements, the spatial information contained within these measurements can significantly improve the rainfall field estimation accuracy. In the case of the synthetic rainfall field explored in this study, this improvement can be as much as a 40% improvement over the cell-averaged radar observations alone.

Real-world rainfall fields

Using an analytical function to represent the underlying rainfall field permitted convenient ground-truth

comparisons; however, real-world rainfall fields may differ in geometry from the field created by Equation (5). To explore how the URS model performs on real-world rainfall fields, data from the Goodwin Creek Experimental Watershed was used. One-hundred rainfall fields were created from these data by interpolating the 10-min rainfall intensity measured by the Goodwin Creek rain gauges during 100 randomly selected rainfall events from the period of January 2006 through December 2007.

These rainfall fields are used to explore the accuracy of the URS model compared to the simulated radar-rainfall estimate. This analysis revealed that a little less than a 30% increase in field reconstruction accuracy is achieved using the URS model instead of the radar-rainfall estimate alone, and that the point of diminishing returns occurs at 15% observational coverage, yielding an increase in accuracy of around 30%. This result is very similar to what was observed using the synthetic rainfall field.

Figure 8 compares the MSE of the URS model with 15% coverage to the MSE of the corresponding simulated radar-rainfall estimate for the 100 randomly selected rainfall events. Each point on this figure represents one of the rainfall events, and the MSE of the URS model is shown on the vertical axis, whereas the MSE of the simulated radar-rainfall estimate is shown on the horizontal axis. Thus, points falling above a one-to-one line indicate cases where the URS model performs worse than the radar-rainfall estimate,
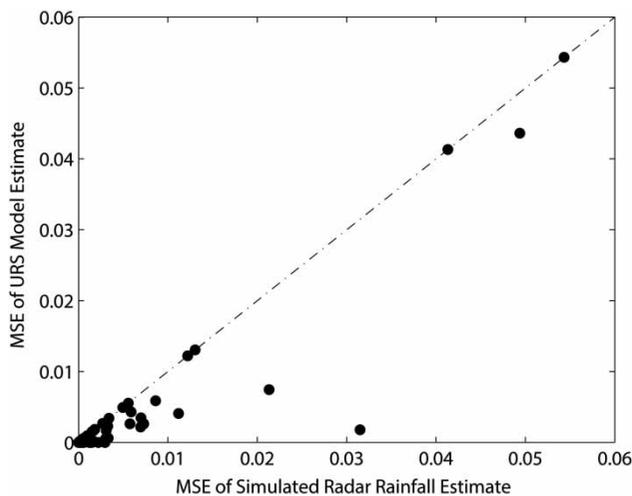
while points falling below the one-to-one line indicate cases where the URS model performs better. As can be seen in this figure, there are no points above the one-to-one line, indicating that the URS model never performs worse than the radar-rainfall estimate; however, most of the time, the URS model provides some amount of benefit (i.e. decreased MSE) over the radar-rainfall estimate alone. In many of these cases, the improvement is small, resulting in a cluster of points near the origin; however, in two cases, the improvement was significant, resulting in points at (0.022, 0.008) and (0.031, 0.002), reflecting MSE reductions of 63 and 94%, respectively. In these two cases, the rainfall was generated by storms of small spatial extent, resulting in rainfall boundaries falling within the simulated radar cells.

DISCUSSION

Few studies have explored the utility of ubiquitous sensors for environmental monitoring. Of these studies, most focus on how to calibrate the ubiquitous sensors for measurement of the phenomenon of interest. This study, however, explores how a spatial statistical model can be used to fuse measurements from ubiquitous sensors and dedicated environmental sensors to create accurate quantitative estimates of rainfall. The fusion method combines a Bayesian solution to an autologistic MRF with a scaling model that concentrates rainfall into rainy regions. The model assumes that a cell-averaged rainfall estimate is available from a weather radar installation; however, this estimate could also come from a satellite or other remote-sensing technology that provides a spatially averaged rainfall estimate over a broad region. The ubiquitous sensors explored in this study provide only binary categorical measurements of either wet or dry. This type of measurement could be provided by a number of ubiquitous sensors already being deployed in the environment, for example as components in rain-sensing windshield wipers, traffic cameras, road weather information systems, or irrigation control systems. To achieve the integration of ubiquitous sensors with heterogeneous measurement capabilities, future work will be necessary to explore how the URS model could be extended to consider ubiquitous sensors with a larger number of possible categorical measurements (e.g. no rain/light rain/

Figure 8 │ Comparison of the MSE of the block-averaged simulated radar-rainfall estimate of the 100 randomly selected Goodwin Creek rainfall fields to the SSE of the URS model with 15% observational coverage by the ubiquitous sensors.

moderate rain/heavy rain) or ubiquitous sensors with different measurement reliabilities. Likely this could be accomplished by modifying Equation (2) to generalize to more than two values of the categorical value X and modifying Equation (3) to include different observation models $P(Y_{i,j}|X_{i,j})$.

Categorical wet/dry measurements alone are insufficient to provide quantitative rainfall estimates; however, they do provide important spatial information on the extent of the rainy region. By combining the quantitative accuracy of the remote-sensing method with the spatial information contained in the ubiquitous sensor measurements, a quantitative rainfall field estimate can be produced that is resilient to malfunction of the ubiquitous sensors and provides an accuracy increase of 40% relative to the remotely sensed rainfall estimate alone. This accuracy increase can be achieved with a ubiquitous sensor density of only 56 sensors per km^2. To put this in perspective, many urban residential areas in North America have one-quarter acre lots. Thus, this level of observational resolution could be achieved if ubiquitous sensors were associated with only 6% of the houses in an urban residential area, for example as part of a smart irrigation system. If more sensors were available, however, the URS model would produce an even more accurate rainfall field estimate.

The URS model treats sensor malfunction very generally, by including a term that quantifies the probabilities that a sensor measurement is an accurate or inaccurate representation of the 'true' state of the rainfall process in a given pixel (Equation (3)). Clearly, this model of sensor malfunction is appropriate for ubiquitous sensors that sometimes report binary observations that are inconsistent with the true state of the rainfall field at that location (e.g. that report wet when it is in fact not raining in that pixel). However, this model can also account for random positioning errors that are to be expected from a GPS device. Positioning errors for civilian devices using the NAVSTAR GPS system are expected to be within the 10–30 m range (Gomerasca 2009; Konecny 2014). Using a pixel size of 50 m, this means that measurements from ubiquitous sensors near the pixel boundary may be erroneously located in a neighboring pixel. If this pixel shares the same wet/dry condition as is measured by the mislocated ubiquitous sensor, then this positioning error will not negatively

impact the model. On the other hand, if this pixel does not share the same wet/dry condition as is measured by the mislocated ubiquitous sensor, then this positioning error will appear as if the sensor is simply malfunctioning by reporting an observation inconsistent with the true state, and thus will not impact the URS model performance differently than already described.

Experiments using the synthetic field showed that as the number of ubiquitous sensors increased, the accuracy of the quantitative precipitation field estimate monotonically increased, albeit at a diminishing rate. While the time to compute the field increased as the number of ubiquitous sensors increased, this increase was insignificant, and the compute time remained very small. For example, on an IBM ThinkCentre with a 3 GHz Intel Pentium G3220 processor and 4 GB RAM, it took 0.17 s to run the URS model with 50 m pixels and 15% ubiquitous sensor coverage; increasing the coverage to 95% increased the computing time to 0.20 s. These behaviors negate the need to carefully manage the number of sensors participating in the observation network and permit sensor networks created by allowing individual users to choose to connect their devices to the observation system. This opt-in strategy is already prevalent in other Web-based systems hosting user-generated content. Furthermore, since the compute time scales with the square of the domain size, it is conceivable that this model could scale up to moderate-sized ($\sim$100 km^2) urban areas, such as Vancouver, and still be computed in less time than the 5-min data frequency of a weather radar. For regions larger than this, parallel processing could be used to reduce computation times.

In addition to increasing the field estimation accuracy, as the number of ubiquitous sensors increases, the URS model becomes more resilient to ubiquitous sensor malfunction, and the speed of computing the binary wet/dry field increases, because there are fewer unobserved locations that need to be estimated. The model presented here assumes a single ubiquitous sensor per pixel, and thus a strategy, such as majority voting, would need to be employed to resolve the situation of multiple sensors per pixel. This could serve to further reduce the impact of erroneous sensors, since it is unlikely that the majority of sensors in a pixel would malfunction at the same time. However, for very large numbers of sensors, such as one might encounter

in dense urban areas, it might become necessary to limit the number of sensors participating in the URS model to improve computational efficiency.

Finally, the results presented in this paper were achieved using simulated ubiquitous sensors located randomly throughout the study region and simulated radar-rainfall estimates that are considered to be error-free. In practice, however, the ubiquitous sensors may not be randomly distributed, but rather clustered according to development patterns, and radar measurements are certainly subject to error (Rico-Ramirez 2012; Bringi *et al.* 2011; Islam *et al.* 2012). If clustering of the ubiquitous sensors is present, then it is expected that this will reduce the level of quantitative rainfall field estimation accuracy achievable by the URS model and may change the location of the point of diminishing returns in the accuracy-versus-coverage curve. However, it is also expected that the URS model, using a reasonable number of ubiquitous sensors, would still produce a superior spatial estimate of rainfall than would a radar estimate alone. Furthermore, although errors in the radar-rainfall estimate will likely negatively impact the level of quantitative rainfall field estimation accuracy achievable by the URS model, it is expected that the model presented here could be integrated with the radar-rainfall estimation and bias-correction process to minimize the impact of radar-rainfall estimation errors and possibly increase the overall field estimation accuracy. Previous work by the author (Hill 2013) has shown that integration of rain gauge data with radar-rainfall estimates may hold potential for real-time correction of radar-rainfall estimates, even when the rain gauge data streams contain measurement errors. In this work, a temporal dynamics model that accounted for the difference in the temporal correlation structure between dry and wet periods was fused with the radar-rainfall and rain gauge data. Such a temporal dynamics model could also be integrated into the URS model to leverage temporal patterns in the rainfall fields.

CONCLUSION

Small, low-cost sensors that are embedded in everyday consumer products have the potential to provide high-resolution environmental observations that can be used to enable smart urban infrastructure. This study explores how a statistical model can be used to fuse noisy measurements from ubiquitous sensors providing categorical wet/dry measurements to improve remotely sensed rainfall intensity measurements by as much as 40%. An exploration of the sensitivity of this URS model to its parameters suggests that there are a wide range of possible parameter combinations that will provide good results even when the ubiquitous sensors malfunction frequently (up to 20% of the time). These results suggest that even with a rough calibration, the URS model will provide a good estimate of an arbitrary rainfall field. Furthermore, the URS model's accuracy increases as the sensor density increases. Thus, as more and more embedded sensing devices are deployed in the urban environment, a fusion model, such as the URS model presented here, will produce more and more accurate estimates of spatial phenomena.

Although many issues of privacy and connectivity must be addressed before ubiquitous sensing of the environment will become practical (Hill *et al.* 2014), this initial study illustrates the immense potential that can be realized by leveraging ubiquitous sensing for environmental monitoring. In addition, future effort should be directed towards issues of statistical modeling of non-stationary spatiotemporal processes and the real-time fusion of asynchronous measurements from heterogeneous sensors with these models.

REFERENCES

Austin, P. M. 1987 Relation between measured radar reflectivity and surface rainfall. *Mon. Weather Rev.* **115**, 1053–1070.

Baeck, M. L. & Smith, J. A. 1998 Rainfall estimation by the WSR-88D for heavy rainfall events. *Weather Forecast.* **13**, 416–436.

Berne, A., Delrieu, G., Creutin, J.-D. & Obled, C. 2004 Temporal and spatial resolution of rainfall measurements required for urban hydrology. *J. Hydrol.* **299**, 166–179.

Besag, J. 1983 On the statistical analysis of dirty pictures. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **48** (3), 259–302.

Besag, J. E. 1974 Spatial interaction and the statistical analysis of lattice systems. *J. R. Stat. Soc. Ser. B* **36**, 162–236.

Boyle, C., Mudd, G., Mihelcic, J., Anastas, P., Collins, T., Culligan, P., Edwards, M., Gabe, J., Gallagher, P., Handy, S., Kao, J.-J., Krumdieck, S., Lyles, L. D., Mason, I., McDowall, R., Pearce, A., Riedy, C., Russell, J., Schnorr, J. L., Trotz, M., Venables, R., Zimmerman, J. B., Fuchs, V., Miller, S., Page, S. & Reeder-Emery, K. 2010 Delivering sustainable infrastructure that supports the urban built environment. *Environ. Sci. Technol.* **44**, 4836–4840.

Bringi, V. N., Rico-Ramirez, M. A. & Thurai, M. 2011 Rainfall estimation with an operational polarimetric C-Band radar in the United Kingdom: comparison with a gauge network and error analysis. *J. Hydrometeorol.* **12** (5), 935–954.

Brown, C. 2010 The end of reliability. *J. Water Resour. Plan. Manage.* **136** (2), 143–145.

Colas, H., Robitaille, L., Laverdiere, C. A. & Lessard, D. 2005 Applications of real time control for CSO and SSO abatement: Lessons learned from 6 years of operation in Quebec City. In: *Impacts of Global Climate Change Proceedings of the American Society of Civil Engineers World Water and Environmental Resources Congress*. May 15–19, 2005, Anchorage, Alaska, USA.

Collier, C. G. 1996 *Applications of Weather Radar Systems*. John Wiley and Sons, London.

Crum, T. D. & Alberty, R. L. 1993 The WSR-88D and the WSR-88D operational support facility. *Bull. Am. Meteorol. Soc.* **74**, 1669–1687.

Crum, T. D., Saffle, R. E. & Wilson, J. W. 1998 An update on the NEXRAD program and future WSR-88D support to operations. *Weather Forecast.* **13**, 253–262.

Ferster, C. J. & Coops, N. C. 2013 A review of earth observation using mobile personal communication devices. *Comput. Geosci.* **51**, 339–349.

Fradet, O., Pleau, M. & Marcoux, C. 2011 Reducing CSOs and giving the river back to the public: innovative combined sewer overflow control and riverbanks restoration on the St. Charles River in Quebec City. *Water Sci. Technol.* **63** (2), 331–338.

Garcia, M., Peters-Lidard, C. D. & Goodrich, D. C. 2008 Spatial interpolation of precipitation in a dense gauge network for monsoon storm events in the southwestern United States. *Water Resour. Res.* **44**, W05S13.

Grimes, D. I. F., Pardo-Igúzquiza, E. & Bonifacio, R. 1999 Optimal areal rainfall estimation using raingauges and satellite data. *J. Hydrol.* **222**, 93–108.

Gomerasca, M. A. 2009 *Basics of Geomatics*. Springer, New York.

Haberlandt, U. & Sester, M. 2010 Areal rainfall estimation using moving cars as rain gauges – a modeling study. *Hydrol. Earth Syst. Sci.* **14**, 1139–1151.

Hill, D. J. 2013 Automated Bayesian quality control of streaming rain gauge data. *Environ. Model. Software* **40**, 289–301.

Hill, D. J. & Farzan, F. 2012 Downscaling radar-rainfall data via ubiquitous sensing and data fusion. In: *Proceedings of the American Society of Civil Engineers Environmental & Water Resources Institute World Water & Environmental Resources Congress*. Albuquerque, NM, May 20–24.

Hill, D., Kerkez, B., Rasekh, A., Ostfeld, A., Minsker, B. & Banks, M. 2014 Sensing and cyberinfrastructure for smarter water management: the promise and challenge of ubiquity. *J. Water Resour. Plann. Manage.* **140** (7), 01814002.

Isham, V. 1981 An introduction to spatial point processes and Markov random fields. *Int. Stat. Rev.* **49**, 21–43.

Islam, T., Rico-Ramirez, M. A., Han, D. & Srivastava, P. K. 2012 Artificial intelligence techniques for clutter identification with polarimetric radar signatures. *Atmos. Res.* **109–110**, 95–113.

Konecny, G. 2014 *Geoinformation: Remote Sensing, Photogrammetry and Geographic Information Systems*, 2nd edn. CRC Press, Boca Raton.

Leijnse, H., Uijlenhoet, R. & Strickler, J. N. M. 2007 Rainfall measurements using radio links from cellular communication networks. *Water Resour. Res.* **43**, W03201.

Marshall, J. S. & Palmer, W. M. 1948 The distribution of raindrops with size. *J. Meteorol.* **5**, 15–166.

National Weather Service 1999 *National Weather Service River Forecast System (NWSRFS) User's Manual*. Office of Hydrology, National Weather Service, NOAA, Silver Spring, MD.

Pleau, M., Pelletier, G., Colas, H., Lavallee, P. & Bonin, R. 2001 Global predictive real-time control of the Quebec urban community's westerly sewer network. *Water Sci. Technol.* **43** (7), 123–130.

Pleau, M., Colas, H., Lavellee, P., Pelletier, G. & Bonin, R. 2005 Global optimal real-time control of the Quebec urban drainage system. *Environ. Model. Software* **20** (4), 401–413.

Rico-Ramirez, M. A. 2012 Adaptive attenuation correction techniques for C-band polarimetric weather radars. *IEEE Trans. Geosci. Remote Sens.* **50** (12), 5061–5071.

Rico-Ramirez, M. A., Cluckie, I. D., Shepherd, G. & Pallot, A. 2007 A high-resolution radar experiment on the island of Jersey. *Meteorol. Appl.* **14**, 117–129.

Ripley, B. D. 1988 *Statistical Inference for Spatial Processes*. Cambridge University Press, Cambridge.

Schutze, M., Campisano, A., Colas, H., Schilling, W. & Vanrolleghem, P. 2004 Real time control of urban wastewater systems – where do we stand today? *J. Hydrol.* **229** (3–4), 335–348.

Seo, B.-C. & Krajewski, W. F. 2010 Scale dependence of radar uncertainty: Initial evaluation of NEXRAD's new super-resolution data for hydrologic applications. *J. Hydrometeorol.* **11**, 1191–1198.

Seo, D.-J. & Smith, J. A. 1991 Rainfall estimation using raingages and radar – a Bayesian approach: 1. Derivation of estimators. *Stoch. Hydrol. Hydraul.* **5** (1), 17–21.

Seo, D.-J., Breidenbach, J. P. & Johnson, E. R. 1999 Real-time estimation of mean field bias in radar rainfall data. *J. Hydrol.* **223**, 131–147.

Smith, J. A. & Krajewski, W. F. 1991 Estimation of the mean field bias of radar rainfall estimates. *J. Appl. Meteorol.* **30**, 397–412.

Smith, J. A. & Krajewski, W. F. 1993 A modeling study of rainfall rate-reflectivity relationships. *Water Resour. Res.* **29**, 2505–2514.

Steiner, M. & Smith, J. A. 1998 Convective versus stratiform rainfall: An ice-mircophysical and kinematic conceptual model. *Atmos. Res.* **47–48**, 317–326.

Steiner, M., Smith, J. A., Burges, S. J., Alonso, C. V. & Darden, R. W. 1999 Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. *Water Resour. Res.* **36** (8), 2487–2503.

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

Upton, G. J. G. & Rahimi, A. R. 2003 On-line detection of errors in tipping-bucket raingauges. *J. Hydrol.* **278**, 197–212.

Velasco-Forero, C. A., Sempre-Torres, D., Cassiraga, E. F. & Gómez-Hernández, J. J. 2009 A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. *Adv. Water Resour.* **32**, 986–1002. doi:10.1016/j.advwatres.2008.10.004.

Xiaoyang, L., Jietai, M., Yuanjing, Z. & Jiren, L. 2003 Runoff simulation using radar and raingauge data. *Adv. Atmos. Sci.* **20** (2), 213–218.

Yilmaz, K., Hogue, T. S., Jsu, K.-L., Sorooshian, S., Gupta, H. & Wagener, T. 2005 Intercomparison of raingauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting. *J. Hydrometeorol.* **6** (4), 497–517.

Zinevich, A., Alpert, P. & Messer, H. 2008 Estimation of rainfall fields using commercial microwave communication networks of variable density. *Adv. Water Resour.* **31** (11), 1470–1480.