AREA PRECIPITATION MEASUREMENT

Teresa Alpuim
CMAF and DEIO, University of Lisbon, Portugal

Keywords: Radar meteorology; area rainfall estimate; Kalman filter; kriging; cokriging**.**

Contents

Summary

This article examines the area precipitation measurement problem. Due to the large space and time variability of rainfall, the precise evaluation in real time of mean area estimates poses a difficult problem. These estimates are traditionally obtained from a set of rain gauges, located at several points in the area but can be substantially improved when the measurements of a meteorological radar are available. However, while gauges give better quality measurements at isolated points, the radar gives a much more detailed picture of the spatial pattern of precipitation. In view of this, statistical methods are needed to look for the optimal way of merging the measurements from the two devices: radar and rain gauges. After a brief account of the usual deterministic methods of radar calibration, the article describes two important radar calibration techniques: Kalman filter and cokriging. The first one puts the emphasis on the time correlation structure of the precipitation and the second on its spatial correlation pattern. The main theoretical concepts and properties of the two methods are explained as well as their application to the area rainfall estimation problem. Particular attention is given to the construction, estimation and assessment of rainfall measurement statistical models.

1. The Area Precipitation Measurement Problem.

Accurate area rainfall measurements have an increasing importance in all applications of Meteorology. Recent advances in Hydrology show the need for a better precision in real-time measurement of area precipitation for operational river-flow and flash-flood forecasting. However, it is difficult to obtain good measurements of the precipitation

field because of its large spatial and time variability. Rain gauges provide precise point estimates but they are not able to picture the spatial pattern of rainfall. A high-density telemetered gauge network is costly and difficult to maintain but it may fail to give the real spatial structure of precipitation.

An alternative way is the use of weather radar that provides measurements in a thin grid of cells, covering inaccessible areas where gauge installation is not possible. The cells may vary in size from 2kmx2km to 5kmx5km. Nevertheless, although weather radar may outline more accurate rainfall isopleths, their point estimates are poor and cannot approach the accuracy of gauges observations. In view of this, radar rainfall estimates have to be calibrated with the help of a certain number of gauges. The best solution is to use the two types of measurements – radar and gauges – and to look for an optimal way of merging them, taking into account the different and complementary nature of the two sensors.

A first approximation of the rainfall intensities throughout the radar field is obtained with the standard *Z-R* relationship

$$Z = AR^b, \qquad (1)$$

where Z stands for rainfall and R for reflectivity. Typically, $A = 200$ and $b = 1.6$. If rain gauges observations are available at some sites along the measurement area, these can be used to improve the first estimates with the help of some calibration technique. Let us consider $\mathbf{G}_t = \begin{pmatrix} g_t^1 & g_t^2 & \dots & g_t^M \end{pmatrix}^T$ as the vector of the gauges measurements in time t at M different sites, and $\mathbf{r}_t = \begin{pmatrix} r_t^1 & r_t^2 & \dots & r_t^M \end{pmatrix}^T$ the corresponding vector of the radar measurements at the same time and places. A simple correction procedure consists in evaluating a single calibration factor to be applied uniformly over the whole area. Under this approach, the calibration factor or bias at time t, that we will represent by b_t, can be evaluated by the formula

$$b_t = \sum_{i=1}^{M} g_t^i \Big/ \sum_{i=1}^{M} r_t^i$$

or

$$b_t = \frac{1}{M} \sum_{i=1}^{M} \frac{g_t^i}{r_t^i}.$$

The first equation assigns each observation a weight proportional to its value and the second equation gives equal weight to all the sites, independently of its value. However, it has been noticed that the relation g/r may vary considerably from place to place. Hence, a better adjustment may be achieved considering the several factors at each site, $b_t^i = g_t^i / r_t^i$, $i = 1,...,M$, and applying to each radar cell the factor corresponding to the closest calibration gauge. Also, a different calibration factor for each cell may be obtained using an interpolation algorithm, for example, multiquadric interpolation.

The uncertainty of the relationship between radar and gauge estimates – varying randomly both in time and space – makes it necessary to use statistical methods in order to improve the radar calibration process. The more simplistic procedures, as described above, provide quick methods of correcting the area estimates but do not take into account some important features of the system – for example, the gauges measurement errors or the temporal correlation structure of the bias – and do not provide a standard error of the estimate. Hence, better precision may be achieved with the use of more sophisticated statistical procedures. These allow a rigorous and clear mathematical modelling of the problem and produce good results in practice. The most common are those based on the Kalman filter or Cokriging techniques. They will be presented through the next two sections.

2. The Kalman Filter Approach.

This approach formulates a mathematical model for the system incorporating two equations: one that relates radar and gauge observations and another that describes the time variation of the bias. The Kalman filter is an iterative procedure that estimates the bias at each time t, on the basis of these two equations. The advantages of Kalman filter over simply computing *g*/*r* ratios are:

1. taking into account the gauges errors;
2. providing the standard error of the estimated bias;
3. taking into account the temporal correlation structure of the bias;
4. avoiding the instability of the ratio *g*/*r*, when *r* is small.

The method gave good results both in simulated and practical studies and is simple to implement in the computer.

2.1. A State Space Representation of the Problem.

We will consider first the case where the same single factor is applied to the whole radar field. Under such circumstances the equation relating radar and gauges observations, called the measurement equation, becomes

$$\mathbf{G}_t = \mathbf{r}_t b_t + \mathbf{e}_t \qquad (2.a)$$

where $\mathbf{e}_t$ is a sequence of a white noise Mx1 vector of errors with mean zero and variance/covariance matrix $E[\mathbf{e}_t \mathbf{e}_t^T] = \mathbf{\Sigma}$. The equation describing the temporal behaviour of the bias, called the state or transition equation, may be given by an autoregressive, AR(1), process,

$$b_{t+1} = \mu + \phi(b_t - \mu) + \varepsilon_{t+1} \qquad (2.b)$$

where ε_t is a white noise sequence with mean zero and variance τ^2, μ is the mean value of the bias and ϕ the lag-1 autocorrelation. The set of the two equations (2.a) and (2.b) is called the state space representation of the problem. Very commonly, μ is taken to be 0 and $\phi = 1$ which originates a random walk process for the calibration factors sequence.

This model is based on the assumption that the bias is equally likely to increase or decrease over the next time step and avoids the problem of the estimation of the constants μ and ϕ. However, as it produces a non-stationary behaviour for the bias, sometimes better results may be achieved using a stationary autoregressive process with a high correlation parameter ϕ.
Another possibility is to allow a different calibration factor at each calibration gauge location, b_t^i, $1 \le i \le M$, each one described by the equations

$$g_t^i = r_t^i b_t^i + e_t^i$$

and

$$b_{t+1}^i = \mu^i + \phi^i \left(b_t^i - \mu^i \right) + \varepsilon_{t+1}^i .$$

In this case, for each i, e_t^i and ε_t^i are white noise error sequences with mean zero and variances σ_i^2 and τ_i^2, respectively.

Considering $\mathbf{B}_t = \left[b_t^1 \quad b_t^2 \quad \cdots \quad b_t^M \right]^{\mathrm{T}}$ the vector of the calibration factors at each location, a general mathematical formulation for the problem, including both the single and multiple factors formulas, may be given by

$$\mathbf{G}_t = \mathbf{R}_t \mathbf{B}_t + \mathbf{e}_t$$

and

$$\mathbf{B}_{t+1} = \boldsymbol{\mu} + \mathbf{F}\left(\mathbf{B}_t - \boldsymbol{\mu} \right) + \boldsymbol{\varepsilon}_{t+1},$$

where $\mathbf{G}_t$ and $\mathbf{B}_t$ are now random vectors, Mx1 and kx1, respectively. The sequences of error vectors $\mathbf{e}_t$ (Mx1) and $\boldsymbol{\varepsilon}_t$ (kx1) are white noise with variance/covariance matrices $\boldsymbol{\Sigma}$ and $\mathbf{T}$. Other parameters of the model are the kxk matrix $\mathbf{F}$ of coefficients and the kx1 vector $\boldsymbol{\mu}$ of the mean values of the $\mathbf{B}_t$ process. $\mathbf{R}_t$ is a matrix built with the radar observations.

The case of a single calibration factor corresponds, in this formulation, to taking $k = 1$ and $\mathbf{R}_t$ the column vector of radar measurements at the M different gauge locations. The transition equation is, thus, a scalar equation, with $\mathbf{F} = \phi$, μ the mean value of the bias sequence and $\mathbf{T} = \tau^2$. In the multiple factor approach, $\mathbf{R}_t$ is an MxM diagonal matrix which elements are the radar observations, r_t^i, $1 \le i \le M$. The covariances matrices $\boldsymbol{\Sigma}$ and $\mathbf{T}$ are both diagonal with elements σ_i^2 and τ_i^2. The matrix $\mathbf{F}$ is also MxM diagonal, with elements ϕ_i which, for the sake of simplicity, will be considered to be equal.

2.2. The Kalman Filter Algorithm

Once the problem is formulated in a state space representation, the Kalman filter may be applied. It consists of a recursive method of providing, at each time *t*, an estimator for the state vector $\mathbf{B}_t$. This estimator is simply the orthogonal projection of the state vector onto the observed variables $\mathbf{G}_1,...,\mathbf{G}_t$, or, equivalently, it is the minimum mean square error (MSE) linear estimator based on the gauges observations up to time *t*.

Each iteration consists of a two-step procedure. The first is an updating step, where the most recent observation $\mathbf{G}_t$ is incorporated in the estimation of the bias $\mathbf{B}_t$. Let us consider that at time *t*-1, the estimator of $\mathbf{B}_t$ is provided by $\hat{\mathbf{B}}_{t|t-1}$ with mean square error

$$\mathbf{P}_{t|t-1} = E\left[\left(\mathbf{B}_t - \hat{\mathbf{B}}_{t|t-1}\right)\left(\mathbf{B}_t - \hat{\mathbf{B}}_{t|t-1}\right)^{\mathrm{T}}\right].$$

Then, a forecast for the next set of gauges observations is given by

$$\hat{\mathbf{G}}_{t|t-1} = \mathbf{R}_t \hat{\mathbf{B}}_{t|t-1}.$$

It may be easily seen that the MSE of this forecast is:

$$E\left[\left(\mathbf{G}_t - \hat{\mathbf{G}}_{t|t-1}\right)\left(\mathbf{G}_t - \hat{\mathbf{G}}_{t|t-1}\right)^{\mathrm{T}}\right] = \mathbf{R}_t \mathbf{P}_{t|t-1} \mathbf{R}_t^{\mathrm{T}} + \mathbf{\Sigma}.$$

When $\mathbf{G}_t$ is available, the actual error of this forecast may be computed, $\hat{\mathbf{e}}_t = \mathbf{G}_t - \hat{\mathbf{G}}_{t|t-1}$, and a new estimator for $\mathbf{B}_t$ is obtained through a weighted average of the estimator in time *t*-1 and the forecast error it produced for the observation $\mathbf{G}_t$, namely:

$$\hat{\mathbf{B}}_{t|t} = \hat{\mathbf{B}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{G}_t - \hat{\mathbf{G}}_{t|t-1}\right).$$

The matrix $\mathbf{K}_t$ is called the Kalman gain matrix and is computed through the formula

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{R}_t^{\mathrm{T}}\left(\mathbf{R}_t \mathbf{P}_{t|t-1}\mathbf{R}_t^{\mathrm{T}} + \mathbf{\Sigma}\right)^{-1}.$$

Note that, if $\hat{\mathbf{B}}_{t|t-1}$ is the orthogonal projection of $\mathbf{B}_t$ onto the gauges observations up to time *t*-1, the new updated estimator, $\hat{\mathbf{B}}_{t|t}$, is the orthogonal projection onto the same observations and $\mathbf{G}_t$. The MSE of the updated estimator may be obtained through the recursive formula

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{R}_t^{\mathrm{T}}\left(\mathbf{R}_t \mathbf{P}_{t|t-1}\mathbf{R}_t^{\mathrm{T}} + \mathbf{\Sigma}\right)^{-1}\mathbf{R}_t \mathbf{P}_{t|t-1}^{\mathrm{T}}.$$

This completes the updating procedure. The second step is a prediction step that provides a forecast for the next value of the vector of calibration factors,

$$\hat{\mathbf{B}}_{t+1|t} = \mathbf{F}\hat{\mathbf{B}}_{t|t}$$

with mean square error

$$\mathbf{P}_{t+1|t} = \mathbf{F}\mathbf{P}_{t|t}\mathbf{F}^{\mathrm{T}} + \mathbf{T}.$$

From now on, the same steps can be repeated. Also, to start the procedure, initial values for $\hat{\mathbf{B}}_{1|0}$ and its MSE, $\mathbf{P}_{1|0}$, are needed. Clearly, when there are no observations available, the projection of the calibration factor onto a constant is just its mean value. Then, when using a stationary model, that is, $|\phi| < 1$, its MSE corresponds to its variance which is given by $\tau^2/(1-\phi^2)$. If $|\phi| = 1$, the process has no finite variance, and the starting value for the MSE is usually taken as any large constant. Nevertheless, the starting values have not much influence on the behaviour of the estimation procedure, except for the first few iterates.

2.3. Estimating the Model Parameters

A major problem arising in the application of the Kalman filter algorithm is that the variance/covariance matrices $\mathbf{\Sigma}$ and $\mathbf{T}$ must be estimated which involves some difficulties. In many applications it is common to use maximum likelihood (ML) estimation because, when the error terms in both equations are normal, the distribution of $\mathbf{G}_t$ conditional to its past ($\mathbf{G}_1$, $\mathbf{G}_2$,...,$\mathbf{G}_{t-1}$), is Gaussian with mean value $\mathbf{R}_t\hat{\mathbf{B}}_{t|t-1}$ and covariance

$$\mathbf{\Omega}_t = \mathbf{R}_t\mathbf{P}_{t|t-1}\mathbf{R}^{\mathrm{T}}{}_t + \mathbf{\Sigma},$$

for any $t \geq 1$. Hence, the log-likelihood may be computed and maximized numerically with respect to the unknown parameters $\mathbf{\Sigma}$ and $\mathbf{T}$ and, if necessary, also $\mathbf{F}$.

However, precipitation data clearly deviate from the Gaussian curve, showing a highly skewed distribution and including several values equal to zero. Actually, the fact that the error terms are unobservable makes it difficult to adjust a specific distribution to them. Consequently, distribution-free methods are the most appropriate for this problem.

One possible way of constructing such estimators is to obtain linear functions of the error terms by appropriately differencing the observations and equating their empirical and theoretical variances, thus deriving moment estimators. To explain the basic idea of this method, we will assume, to simplify things, that $\boldsymbol{\mu} = \mathbf{0}$. Let us consider, first, the multiple factor approach where, at each location and for any $k \geq 1$,

$$g_{t+k} = r_{t+k}b_{t+k} + e_{t+k}.$$

To simplify notation, as the procedure will be the same for each location i, we will omit the index i, in the multiple factor approach. As the bias follows an AR(1) process, it verifies the equation,

$$b_{t+k} = \phi^k b_t + \phi^{k-1}\varepsilon_{t+1} + \ldots + \phi\varepsilon_{t+k-1} + \varepsilon_{t+k}.$$

Consequently, and supposing that ϕ is known, if we define the quantities

$$\begin{aligned} d_t(k) &= \frac{g_{t+k}}{r_{t+k}} - \phi^k \frac{g_t}{r_t} \\ &= \phi^{k-1}\varepsilon_{t+1} + \ldots + \phi\varepsilon_{t+k-1} + \varepsilon_{t+k} + \frac{e_{t+k}}{r_{t+k}} - \phi^k \frac{e_t}{r_t}, \end{aligned}$$

their variances are given by

$$E\left[d_t^2(k)\right] = \begin{cases} \tau^2 \dfrac{1-\phi^{2k}}{1-\phi^2} + \sigma^2\left(\dfrac{1}{r_{t+k}^2} + \dfrac{\phi^{2k}}{r_t^2}\right) & \text{if } |\phi| < 1; \\ k\tau^2 + \sigma^2\left(\dfrac{1}{r_{t+k}^2} + \dfrac{\phi^{2k}}{r_t^2}\right) & \text{if } |\phi| = 1. \end{cases}$$

Now, if $\overline{D}(k)$ represents the time average of the *N-k* values of $d_t^2(k)$, then

$$D(k) = E(\overline{D}(k)) = E\left(\frac{1}{N-k}\sum_{t=1}^{N-k} d_t^2(k)\right) = \tau^2 \Psi(k) + \sigma^2 H(k)$$

where

$$\Psi(k) = \begin{cases} \dfrac{1-\phi^{2k}}{1-\phi^2} & \text{if } |\phi|<1; \\ k & \text{if } |\phi|=1. \end{cases}$$

and

$$H(k) = \frac{1}{N-k}\sum_{t=1}^{N-k} \frac{1}{r_{t+k}^2} + \frac{\phi^{2k}}{r_t^2}.$$

If there are values of *t* for which either r_t or r_{t+k} are zero, the differences $d_t(k)$ cannot be computed and the averages $\overline{D}(k)$ and $H(k)$ will be taken only over the smaller number of terms such that both radar observations are not zero. Finally, considering any pair of indices *k* and l a system of the form

$$\begin{cases} \overline{D}(k) = D(k); \\ \overline{D}(\ell) = D(\ell); \end{cases}$$

can be constructed and the solution will provide estimators for τ^2 and σ^2, namely,

$$\begin{cases} \hat{\tau}^2 = \dfrac{H(\ell)\bar{D}(k) - H(k)\bar{D}(\ell)}{\Psi(k)H(\ell) - \Psi(\ell)H(k)}; \\ \hat{\sigma}^2 = \dfrac{\Psi(k)\bar{D}(\ell) - \Psi(\ell)\bar{D}(k)}{\Psi(k)H(\ell) - \Psi(\ell)H(k)}. \end{cases}$$

These estimators are unbiased and consistent. Ideally, the best choice for the indices should be $k = 1$ and $\ell = 2$ because they use more observations and originate more efficient estimates. However, a problem that may occur when using this method is that some pairs of indices may originate negative values for the variances. If so, another pair of indices should be tried. In general, fixing one of the indices in a small value, say 1, and increasing the other index will result on estimators belonging to the space of parameters.

The same method can be used to estimate the variances and covariances in the single factor model. To simplify things, let us suppose that the matrix $\mathbf{\Sigma}$ is diagonal, that is, that the gauges errors are not correlated which is something very likely to occur in practice. Then we have to estimate τ^2 and σ_i^2, for $i = 1,...,M$. The same type of differences are used at each site i:

$$d_t^i(k) = \frac{g_{t+k}^i}{r_{t+k}^i} - \phi^k \frac{g_t^i}{r_t^i} = \phi^{k-1}\varepsilon_{t+1} + ... + \varepsilon_{t+k} + \frac{e_{t+k}^i}{r_{t+k}^i} - \phi^k \frac{e_t^i}{r_t^i}$$

Again, if $\overline{D}^i(k)$ represents the time average of the elements $\left[d_t^i(k)\right]^2$, its mean value is given by

$$D^i(k) = E\left[\overline{D}^i(k)\right] = \Psi(k)\tau^2 + \sigma_i^2 H^i(k),$$

where $\Psi(k)$ and $H^i(k)$ are defined as previously. For fixed $k \geq 1$ there are M equations of the form

$$\overline{D}^i(k) = D^i(k),$$

so one more is needed to estimate the M+1 parameters. Another equation of this type, for different ℓ and any location i_0 may be added, but as the choice of the location is somewhat arbitrary a more efficient solution is to average all the equations for ℓ over all the sites to get

$$\overline{\overline{D}}(\ell) = \Psi(\ell)\tau^2 + \frac{1}{M}\sum_{i=1}^{M} \sigma_i^2 H^i(\ell).$$

This originates the estimators

$$\hat{\tau}^2 = \frac{\overline{\overline{D}}(\ell) - \frac{1}{M}\sum_{i=1}^{M} D^i(k)H^i(\ell)/H^i(k)}{\Psi(\ell) - \Psi(k)\frac{1}{M}\sum_{i=1}^{M} 1/H^i(k)}$$

and

$$\hat{\sigma}_i^2 = \frac{\overline{D}^i(k) - \Psi(k)\hat{\tau}^2}{H^i(k)} \qquad \text{for } i = 1,...,k.$$

Finally, values have to be assigned to the constants μ and ϕ. As $E(g_t/r_t) = \mu$, this parameter may be estimated by simply averaging over time the ratios g_t/r_t. Note that when $\mu \neq 0$, the differences $d_t(k)$ are given by

$$d_t(k) = \left(\frac{g_{t+k}}{r_{t+k}} - \mu\right) - \phi^k\left(\frac{g_t}{r_t} - \mu\right).$$

The estimation of ϕ is more problematic. A possible suggestion is to estimate ϕ – and even μ – through the minimization of the sum of squares of $d_t(1)$. However, this has to be done with caution once the least squares method will be applied to observations with different unknown variances. In fact, this is a problem deserving further investigation. Of course, there remains the possibility of trying several values and choose the one producing the best practical result in the sense it is explained in section 2.4.

-
-
-

Bibliography

Ahnert P.R., Krajewski and W.F. and Johnson E.R. (1986). Kalman filter estimation of radar rainfall field bias. 23 *rd Conference on Radar Meteorology* 3, JP33-JP37. [This introduces the use of Kalman filter in the radar calibration process.]

Alpuim T. (1999). Noise variance estimators in state space models based on the method of moments. *Annales de l' I.S.U.P.* XXXXIII, fasc. 2-3, 3-23. [This suggests the non-parametric estimating method of the state space parameters in the Kalman filter approach and derives some of its properties.]

Alpuim T. and Barbosa S. (1999). The Kalman filter in the estimation of area precipitation. *Environmetrics* 10, 377-394. [This explains how to estimate the parameters of the radar/raingauge state space model and presents a case study with real data where the single and multiple factors approaches are compared.]

Collinge V.K. (1991) Weather radar calibration in real-time: prospects for improvement. *Hydrological Applications of Weather Radar*, 25-41. [Includes a comprehensive overview of radar calibration methods and problems.]

Cressie N. (1993). *Statistics for Spatial Data*, Wiley and Sons, New York, USA. [This book is a complete and detailed description of all spatial statistics methods including kriging and cokriging.]

Hamilton J. D. (1994). *Time Series Analysis*, Princeton University Press, New Jersey, USA. [This is a detailed and rigorous approach to time series theory with a very good account about the Kalman filter.]

Hardy, R.L. (1971). Multiquadric equations of topography and other irregular surfaces. *J. Geophys. Res*. 76(8), 1905-1915. [This introduces the method of multiquadric interpolation.]

Krajewski W.F. (1987). Cokriging Radar-Rainfall and Rain Gauge Data. *J. Geophys. Res*. 92(d8), 9571-9580. [Describes the cokriging approach and presents a controlled simulation experiment.]

Krajewski W.F. and Ahnert P.R. (1986). Near Real-Time Tests of a Multivariate Analysis System. 23 *rd Conference on Radar Meteorology* 3, JP66-JP69. [Introduces the cokriging methodology applied to radar calibration and describes an operational experience with an hydrological application.]

Biographical Sketch

Teresa Alpuim is Professor at the Department of Statistics and Operations Research, University of Lisbon. Her main research and teaching interests focus on the application of Statistics to the Natural and Environmental Sciences. Also, from the methodological point of view she works in time series, Kalman filter, spatial statistics and linear models. Since 1993, she has collaborated with the group of meteorologists in charge of the installation and maintenance of the weather radar's network. She belongs to the Center for Mathematics and Applications of the University of Lisbon where she directs a research project on Environmental Statistics. This unit studies statistical methods to be applied to many different areas like meteorology and climate, marine biology, water quality monitoring, epidemiology and public health.