A comparison between single site modeling and multiple site modeling approaches using Kalman filtering

Magda Monteiro and Marco Costa

Citation: AIP Conference Proceedings **1648**, 110003 (2015); doi: 10.1063/1.4912410
View online: http://dx.doi.org/10.1063/1.4912410
View Table of Contents: http://scitation.aip.org/content/aip/proceeding/aipcp/1648?ver=pdfcov
Published by the AIP Publishing

A Comparison between Single Site Modeling and Multiple Site Modeling Approaches using Kalman Filtering

Magda Monteiro[*,†] and Marco Costa[*,**]

[*]*Escola Superior de Tecnologia e Gestão de Águeda, Universidade de Aveiro, 3754–909 Águeda, Portugal*
[†]*Centro de Investigação e Desenvolvimento em Matemática e Aplicações, Universidade de Aveiro, 3810-193 Aveiro, Portugal*
[**]*Centro de Matemática e Aplicações Fundamentais, Universidade de Lisboa, 1649-003 Lisboa, Portugal*

Abstract. This work presents a comparative study between two approaches to calibrate radar rainfall in real time. The weather radar provides continuous measurements in real-time which have errors of either meteorological or instrumental nature. Locally, gauge measurements have a greater performance than radar measurements that can be used to improve radar estimates. One way of doing that is via a state space representation associated to the Kalman filter algorithm. In the single-site modeling approach we use the linear calibration model applied in [1] and [3] while the multivariate state-space model proposed in [6] is used in the multiple site approach. This work aims to discuss and compare these two different state space formulations based on the same data set.

Keywords: Kalman filter, state space model, rainfall estimates, weather radar, calibration
PACS: 02.50Ey, 05.45.Tp

INTRODUCTION

The weather radar provides continuous measurements of precipitation in real-time in a large area, for instance in a circle radar umbrella that can reach 300Km of radius. This kind of observations can not be obtained through telemetered rain gauges networks, even if the network is dense, due to the large space-time variability of precipitation. However, the good spatial coverage of the radar comes along with problems in determining the rainfall intensity since radar measurements have errors associated with the reflectivity measurement errors or reflectivity rainfall rate (Z-R) conversion errors. It is known that locally, rain gauges have a better performance when compared with the radar measurements which have errors of either meteorological or instrumental nature that need to be reduced. In the recent years several approaches have been proposed to minimize radar errors among which is included the merging of radar measurements with gauge observations through a state space representation associated to the Kalman filter algorithm. In [2] it is proposed a single-site modeling approach with a power law model to describe the relationship between gauge and radar data, which needs a linearization procedure with the logarithmic function, while [1] and [3] used a linear calibration model with a multiplicative calibration factor. The radar calibration based on the last two models uses a single site approach combined with interpolation methods. Note that linear calibration model can also be used in a multiple site approach, but the simplicity and parsimony of the single site modeling approach make it more attractive ([4]). The work in [6] presents a multivariate state-space model to assimilate rainfall runoff data from different sources of measurement, which was used in [5] to adjust radar rainfall intensities for forecast outflow from two catchments in Copenhagen. In this work we present a comparison between state-space models based on the same data set. In the single-site modeling approach it was used the linear calibration model combined with an interpolation method while in the multiple site approach it was used the model proposed by [6].

STATE SPACE MODELS AND THE KALMAN FILTER

The Kalman filter algorithm provides a real-time scheme to calibrate radar rainfall estimates based on the rain gauge measurements. It is applied to a class of models that admits a state space representation of the form

$$\boldsymbol{A}_t = \boldsymbol{B}_t \boldsymbol{\beta}_t + \boldsymbol{e}_t \quad (1)$$

$$\boldsymbol{\beta}_t = \boldsymbol{\Phi} \boldsymbol{\beta}_{t-1} + \boldsymbol{\varepsilon}_t. \quad (2)$$

Proceedings of the International Conference on Numerical Analysis and Applied Mathematics 2014 (ICNAAM-2014)
AIP Conf. Proc. 1648, 110003-1–110003-4; doi: 10.1063/1.4912410

Equation (1) is the measurement equation and relates the observable vector $\boldsymbol{A}_t$ with the unobservable variable $\boldsymbol{\beta}_t$, called the state, while (2) is the transition or state equation. $\boldsymbol{B}_t$ is a known matrix of coefficients and $\boldsymbol{e}_t$ is the measurement error which is a white noise with covariance matrix Σ_e. The state $\boldsymbol{\beta}_t$ is a VAR(1) process where $\boldsymbol{\varepsilon}_t$ is a white noise with covariance matrix Σ_ε. Furthermore it will be considered that the disturbances $\boldsymbol{e}_t$ and $\boldsymbol{\varepsilon}_t$ are uncorrelated and normally distributed.

Assuming that parameters of the state-space model are known, the Kalman filter is an iterative algorithm that produces, at each time t, an estimator of the state vector $\boldsymbol{\beta}_t$, which is the orthogonal projection of the state vector onto the observed variables up to that time. Let $\widehat{\boldsymbol{\beta}}_{t|t-1}$ represent the predictor of $\boldsymbol{\beta}_t$ based on the information up to time $t-1$ and let $\boldsymbol{P}_{t|t-1}$ be its mean square error (MSE). The recursive process needs initial values for the state $\hat{\beta}_{1|0}$ and for its MSE $\hat{P}_{1|0}$.When the parameters of the model are unknown they have to be estimated and plugged in into the Kalman filter recursive equations.

MODELS

Linear Calibration (LC)

The linear calibration model, which was applied by [1] and [3], relates rain gauges and radar measurements through a multiplicative factor of calibration, as follows

$$\begin{aligned} G_t &= \beta_t R_t + e_t & (3) \\ \beta_t &= \mu + \phi(\beta_{t-1} - \mu) + \varepsilon_t. & (4) \end{aligned}$$

G_t is the rain gauge observation in time t, R_t is the radar measurement at the same time and location and β_t is the respective calibration factor. This formulation assumes a local linear relation between radar and rain gauge estimates since it is considered, in each time, that rain gauge measurement is proportional to radar observation added to an error. As this approach is a single-site modeling procedure also needs interpolation methods to calibrate radar estimates in other sites.

Multivariate Modeling (MM)

The authors in [6] proposed a multivariate state-space model based on the notion that "Tomorrow's weather will probably be like today's". At every time, the model is used to predict the rainfall intensity in a grid $m \times n$ having into account the knowledge from the previous time. In this model the prediction of each pixel is made as a weighting of its own previous value with those of its immediate neighbor pixels. For each time t, the adjusted rainfall in the pixel (i,j), $i = 1,\cdots,m$, $j = 1,\cdots,n$, is given by

$$\beta_{i,j,t} = \sum_{r=1}^{m}\sum_{l=1}^{n} \alpha_{r,l}\beta_{r,l,t-1} + \varepsilon_{i,j,t}, \text{ where } \alpha_{r,l} = \left\{ \begin{array}{ll} 0 & (r \neq i-1 \wedge r \neq i+1) \vee (l \neq j-1 \wedge l \neq j+1) \\ a & r = i, l = j \\ \dfrac{1-a}{b} & \text{other cases} \end{array} \right\}, \quad (5)$$

with b the number of neighbors of the pixel (i,j), which can be 8, 3 or 5 if (i,j) is in the middle, in one of the corners or is one of the remaining elements of the frontier in the grid.

Furthermore, the adjusted rainfall is related with the radar observations and the observations of k rain gauges through a linear function. In matrix notation we have

$$\begin{aligned} \boldsymbol{Y}_t &= C\boldsymbol{\beta}_t + \boldsymbol{e}_t. & (6) \\ \boldsymbol{\beta}_t &= A\boldsymbol{\beta}_{t-1} + \boldsymbol{\varepsilon}_t & (7) \end{aligned}$$

At each time t, $\boldsymbol{\beta}_t$ is a vector state, with size $mn \times 1$, containing, by traversing the grid lines, the adjusted rainfall depth for all pixels in the rain plane, that is the adjusted rainfall of the (i,j) cell of the grid is the $(i-1)n + j$ element of $\boldsymbol{\beta}_t$. A is a $mn \times mn$ weighting matrix performing the spatial average defined in (5) and $\boldsymbol{\varepsilon}_t$ is a $mn \times 1$ vector errors with covariance matrix $\Sigma_1 = \sigma_\varepsilon^2 I_{(mn)^2}$, corresponding to the absence of correlation between the states. States and measurements are related through (6), where $\boldsymbol{Y}_t$ is the observation vector with size $(mn+k) \times 1$ containing the non-adjusted measures from all radar pixels and also k rain gauge measurements. C is a matrix of zeros and ones, with dimension $(mn+k) \times mn$, relating states and observations on the same pixel, for all pixels in the grid. The vector of observations errors, $\boldsymbol{e}_t$, has length $(mn+k) \times 1$ with covariance matrix $\Sigma_2 = diag(\sigma_R^2, \cdots, \sigma_R^2, \sigma_G^2, \cdots, \sigma_G^2)$, which means constant variances σ_R^2 and σ_G^2 for radar and rain gauges observations respectively and no correlation between observation positions. In order to obtain radar calibrated estimates through the Kalman filter approach, we need to have parameters estimates in both models. The adopted methodology for parameter estimation was the maximum likelihood method.

COMPARATIVE STUDY

It is used a data set of 6 stratiform storms between January of 1999 and November of 2000 (in a total of 58 hourly precipitation estimates) in a 10×14 Km^2 area, in a grid of 5 by 7, located around 40 Km north of Lisbon at a distance from 31 to 44 Km from the weather radar in *Cruz do Leão*. This area has five rain gauges: Merceana (Mr), Meca (M), Olhalvo (O), Penedos (P) and Abrigada (A) and it has the highest gauge density under the radar umbrella ($\sim$ 1 gauge/28Km^2). In order to compare the calibration performances of the two models, following the work of [1], we considered the Olhalvo and Abrigada gauges in the modeling procedure to calibrate the remaining gauges. Since we have few storms, it was not possible to separate the estimation procedure from the evaluation procedure, which would be the most appropriate procedure, thus all the storms were used in both phases.

For each model it was implemented, in R software, the Kalman filter equations enabling the maximum likelihood estimation of the parameters through an optimization routine, which estimates can be found in table 1. Note that in MM model the estimated variance associated to the gauges measurements is superior to the estimated variance associated with radar measurements which goes against the knowledge of better local performance of gauge measurements. In LC model, the estimated variance associated to Olhalvo gauge is superior to the estimated variance associated to the calibration factor which is not very common in this kind of models.

TABLE 1. Maximum likelihood parameter estimates

MM model				LC model				
a	σ_ε^2	σ_R^2	σ_G^2		μ	ϕ	σ_ε^2	σ_e^2
0.2353	0.4842	0.3486	1.9540	Abrigada	1.8947	0.3116	1.5028	0.3341
				Olhalvo	1.4131	0.9221	0.0632	1.2141

After parameter estimation, the Kalman filter was applied to predict the state $\boldsymbol{\beta}_t$ at each hour t. As it is considered a real-time procedure, the filtered prediction $\widehat{\boldsymbol{\beta}}_{t|t}$ was used to predict $\boldsymbol{\beta}_t$. In the MM model the application of the Kalman filter allows obtaining the calibrated radar for each cell of the grid whereas the LC model needs interpolating the calibration factors, $\widehat{\beta}_{t|t}$, of Abrigada and Olhalvo to other locations of the grid. In this case it is considered the inverse square distance method to calibrate the radar estimates. Note that for LC model the radar calibration is obtained by multiplying the radar estimate R_t by the filtered calibration factor $\widehat{\beta}_{t|t}^{(LC)}$.

Figure 1 presents, for each of the six storms, an illustration of the radar measurements, the gauge measurements and the calibrated radar for each model for Meca's location and Merceana's location.

The models performance assessment is done according to the empirical square root of the mean square error of point prediction using the six storms that we have. It is compared the gauges rainfall estimates G_t with the calibrated radar cell measurement $\widehat{R}_t^{(l)}$, with $l =$ LC and MM, at the locations of Penedos ($i = 1$), Meca ($i = 2$) and Merceana ($i = 3$). The empirical square root of the mean square error RMSE$^{(l)}$ with model l is computed by

$$\text{RMSE}^{(l)} = \sqrt{\frac{1}{3 \times 58} \sum_{i=1}^{3} \sum_{t=1}^{58} \left(G_t^i - \widehat{R}_t^{(l),i} \right)^2},$$

which assumes the value 1.0973 in MM model and the value of 0.9210 in the LC model. In order to analyse the impact of the calibration procedures we used the pre-calibration RMSE of the 3 rain gauges used in the calibration procedure as the reference value. For the data set under consideration this value was 1.3235. This means that there was about 17%

of reduction of RMSE in MM model whereas the RMSE in LC model presented a decrease of about 30%, indicating that LC model performs better than the model MM in terms of calibrating the radar. In MM model, the variance of the model predictions σ_ε^2 is generally estimated smaller than that of the observations σ_G^2, when computing the adjusted rainfall values, the Kalman filter will therefore show a tendency to smooth the observed values not allowing the desired adjustment.

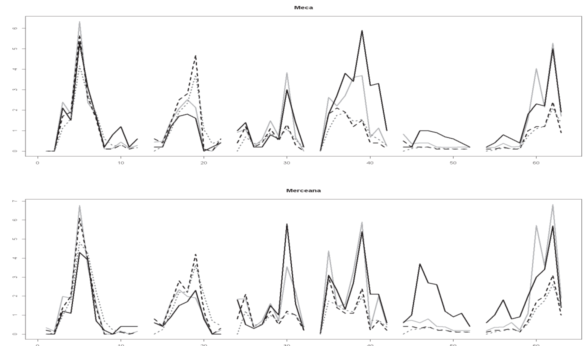

FIGURE 1. Radar measurements (dashed), gauge measurements (black line) and the calibrated radar (MM - dotted; LC - light gray) in Meca and Merceana

ACKNOWLEDGMENTS

Marco Costa was partially supported by Portuguese Foundation for Science and Technology ("FCT–Fundação para a Ciência e a Tecnologia"), PEst OE/ MAT/ UI0209/ 2011. Magda Monteiro was partially supported by Portuguese funds through the CIDMA - Center for Research and Development in Mathematics and Applications, and the Portuguese Foundation for Science and Technology ("FCT–Fundação para a Ciência e a Tecnologia"), within project PEst-OE/MAT/UI4106/2014

REFERENCES

1. T. Alpuim and S. Barbosa, *The Kalman filter in the estimation of area precipitation*, *Environmetrics*, **10**, 377 – 394 (1999).
2. M. Brown, P. Diggle, M. Lord and P. Young, *Space-Time Calibration of Radar Rainfall Data. J. R. Stat. Soc., Series C (Applied Statistics)*, **50**, 221 – 241 (2001).
3. M. Costa and T. Alpuim, *Parameter estimation of state space models for univariate observations. J. Stat. Plan. Inference*, **140**, 1889 – 1902 (2010).
4. M. Costa and T. Alpuim, *Adjustment of state space models in view of area rainfall estimation. Environmetrics*, **22**, 530 – 540 (2011).
5. R. Löe, P.S. Mikkelsen, M.R. Rasmussen and H. Madsen, *State-space adjustment of radar rainfall and skill score evaluation of stochastic volume forecasts in urban drainage systems Water Science & Technology*, **68**, 584 – 590 (2013).
6. M. Grum, P. Harremöes and J J. Linde, *Assimilating a multitude of rainfall and runoff data using a stochastic state space modeling approach*. In: Proceedings of the 9th International Conference on Urban Drainage, pp. 8 – 11 September 2002, Portland, Oregon, USA.