Calibration of Remote Sensing Measurements from Surface Observations

Malwina J Luczak

Abstract

This paper presents a range of modelling techniques that may be used to tackle the problem of calibrating remote sensing measurements from surface observations. The problem was presented by the Hong Kong Observatory at the 1st Hong Kong Study Group with Industry at the City University of Hong Kong in July 2002.

1 Introduction

Remote sensing instruments such as radar and satellite are used to measure the distribution of weather elements in space. Typically, satellite data takes the form of measured 2-dimensional distributions of quantities such as surface temperature, cloud top temperature or air humidity. On the other hand, radar can produce detailed precipitation information for large areas from a single location in real time. Here, the data constitutes measurements of the reflectivity of transmitted electromagnetic waves, measured in units termed dBZ.

Roughly speaking, high values of radar reflectance imply heavy rainfall. However, areal and point estimates are often in error by a factor of two or more. Error sources are many and varied, and include the actual measurement of radar reflectivity factor, evaporation and advection of precipitation before reaching the ground, variations in the drop size distribution, and vertical and horizontal air motions. The main focus of this report is calibration (that is, rain gauge adjustment) of radar measurements in order to predict rainfall intensity. Nevertheless, we believe that similar modelling techniques apply to other types of remote measurements.

It is usual to estimate rainfall through a rain gauge network on the surface. Unfortunately, radar data and rain gauge data have very different characteristics, and the precise relationship between them is unclear. The information available via radar is a 3-dimensional near-continuous (that is, on a fine spatial grid of pixels) field of measured reflectance values, with updates available roughly every 6 minutes. In contrast, rain gauge networks provide irregularly distributed discrete measurements on the ground that are updated every 1 to 5 minutes.

Surface measurements are known to give a better estimate of intensity, but the validity of the measurement over a large area is uncertain, particularly in convective rainfall situations. Large networks of rain gauges are sometimes employed to obtain improved estimates, but the design of such networks is a complex issue. A considerable number of rain gauges would be required to obtain a representation of the rainfall field structure, and this would involve prohibitively large expense.

Radar provides measurement over a much wider area and gives a higher resolution representation of the rainfall field structure. However, the quantitative measurements are less accurate than those obtained from rain gauges. It is therefore very natural to try to combine the more qualitative representation of the rainfall field distribution provided by radar with the more quantitative point rainfall measurements made by rain gauges.

* Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WB, UK. e-mail: M.J.Luczak@statslab.cam.ac.uk
Unfortunately, there are various difficulties in analysing the dependence between the two types of data and the actual rainfall. The nature of this dependence is affected in complex ways by the topography of the region, changing winds, and numerous other factors. The raindrop size distribution is rarely known and varies in time and space. A raindrop falling from the sky could experience disturbances, and its size and direction could be altered before it reaches the surface. This is particularly likely to happen during thunderstorms, when vertical and horizontal air motions are of the same magnitude as raindrop velocity. On the other hand, in light rain, when raindrops are very small, many of them could simply fail to hit the ground before they vaporise.

Previous studies have found that, in general, techniques that combine sparse rain gauge records with radar produce smaller measurement errors than either system alone can provide. Thus radar certainly has potential for improving forecasts of severe storms and flash flooding. However, when high accuracy measurements are needed, the usefulness of radar is diminished, as the number of rain gauges required for calibration may itself be sufficient to provide the desired accuracy.

A number of different approaches have previously been tried to calibrate radar data. These range from space-time models of the rainfall field, which incorporate the covariance structure of that field and measurement errors, to simpler formulations based on linear interpolation, kriging or surface fitting - see references in [12]. Many of these methods are fairly effective for long-term forecasts, such as daily rainfall, but fail to make accurate predictions for shorter periods, such as one hour. Our aim here is to suggest some mathematical models and techniques that might be particularly useful for short-term forecasts. The suitability of these models should be assessed using available data. We make no claims to having been thorough and complete: an in-depth study would require far more time than the amount available to us.

2 Some earlier studies

2.1 Theory of radar measurements

The following theoretical summary is based on the presentation in [11]. The backscattered radar power from precipitation particles is proportional to the summation of the sixth power of particle diameters \(D \delta\) in a unit volume illuminated by the radar beam. Hence the radar reflectivity factor, \(Z\), is defined as

\[
Z = \sum_i N_i D_i^6 = \int_0^\infty N(D)D^6dD,
\]

where \(N_i\) is the number of drops per unit volume of air with diameter \(D_i\), and \(N(D)\) is the number of drops with diameters between \(D\) and \(D + dD\) in a unit volume of air. In the absence of vertical air motions, the rainfall rate, \(R\), is given by

\[
R = \pi \int_0^\infty N(D)D^3V(D)dD,
\]

where \(V(D)\) is the terminal velocity of a raindrop of diameter \(D\), which is often approximated by \(V(D) \approx 1400D^{1/2}\). If one were to substitute a drop size density function in the expressions above, the result would be an expression for the relation between \(Z\) and \(R\). For instance, assuming that the distribution of the drop size is exponential, one obtains a simple formula of the form

\[
Z = AR^b,
\]

where \(A\) and \(b\) are constant parameters. Though the simplicity of (1) is appealing, studies of various data sets have shown that the drop size distribution normally varies over both time and
space. Even taking account of such variations, the assumption that each drop is spherical with
an exponentially distributed radius might not be a very satisfactory fit - see [10] and references
therein. Furthermore, especially during thunderstorms, vertical air motions can be of the same
order of magnitude as particle terminal velocities. All in all, various factors undermine the
validity of the simple relation (1). A thorough discussion of the different elements (including
hardware calibration and changes in the \(Z - R\) relation) that contribute to errors in radar rainfall
measurement can be found in [11]. The same reference also includes a description of a few very
early radar calibration techniques and their performance when applied to real data.

2.2 Calibration of radar measurements

In this section we briefly summarise a number of recent attempts to calibrate rainfall measure-
ments.

In [1], Atlas et al. present a theory for the estimation of total rainfall from an individual
convective storm over its lifetime and the areawide instantaneous rainfall from a multiplicity
of such storms. Their method relies on the existence of a well-behaved probability density
function of rain rate. It involves measuring the volume rainfall using the area-time integral
of the radar echo in excess of a specified threshold over the life of a storm. There are various
difficulties associated with the practical implementation of the method, mostly connected with
the estimation of the values of the relevant parameters and with the calibration of radar data.
However, in [2] the same authors analyse a real data set, and are able to give some evidence in
favour of their technique.

Another technique for rainfall measurement is the window probability matching method
(WPMM). The essence of the method is to use a combination of radar and rain gauge data, and
to match rain gauge intensities to radar reflectivities taken only from small windows centred
about the gauges in time and space. The results are then used to obtain an empirical rela-
tionship between the radar-measured reflectance \(Z\) and rain rate \(R\). In this way (in contrast
to many earlier techniques), the method allows for spatial and temporal variations of the rain
rate throughout the radar domain. One study of this type has been conducted by Rosenfeld,
Wolff and Amitai, and its findings can be found in [9]. The authors show that their technique
can significantly outperform techniques based on the assumption of a time- and space-constant
power-law relationship between \(Z\) and \(R\).

In two companion papers [12] and [13], Wood et al. examine the accuracy of rainfall estimates
obtained from rain gauges and from weather radar using the HYREX network. The HYREX
network consists of 49 0.2 mm tipping bucket rain gauges, located in a grid of 28 \(2 \times 2 \text{ km}^2\)
squares in the catchment of the river Brue in Somerset, England. It has provided a data set
unique in the UK thanks to its densely placed rain gauges and radar stations available for
studying precipitation on both local and global scales in space and time.

In the study presented in [13], data from HYREX was subjected to strict quality control to
identify periods when gauges were not functioning properly; data corresponding to such periods
was rejected. Care was also taken to detect more unusual forms of precipitation, such as snow,
hail and freezing rain, since these may be expected to activate the tipping mechanism of a
rain gauge in a different way to regular rainfall. Certain control procedures were also applied
to radar measurements. Having addressed the issue of quality control, the resulting data was
used to estimate measurement accuracy at different time and space scales and as a function of
rainfall intensity. An empirical approach was pursued by the authors which involved dividing
the time into 15 minute intervals and treating each interval separately. For each interval the
"true" rainfall was set to the average rainfall over a large number of gauges and the estimate and
error were calculated. Then the interval length was varied to investigate the effect of the time
scale over which measurements are made. The accuracy of radar measurements with respect to
gauge estimates was also examined without taking account of long-term bias in radar data. The essence of the approach is the assumption that there are so many rain gauges used in calculating the mean rainfall $T$, that it is essentially the same as the unknown true value for the pixel. The value from each single rain gauge or radar pixel, $R$, can be used to define an estimate of the mean square error of $R$ for a chosen time-frame as

$$S^2 = (R - T)^2.$$ 

The procedure then involves averaging values of $S^2$ across time frames or using many time frames to construct scatter plots of $\log S$ against $\log T$. The reader is referred to [13] for details of the methods and results.

The companion paper [12] explores the accuracy of calibrated (that is, rain gauge-adjusted) weather radar data. When comparing rain gauge and radar performance, one tries to use the value for a distant rain gauge as an estimate of rainfall in a given pixel. When radar estimates are concerned, one does not have to rely on the value for a remote pixel, since a value at the point of interest will be available. In view of this fact Wood et al. investigate two main issues. Firstly, at what distance from a 2 km pixel will the rain gauge measurement be outperformed by a coincident radar pixel? Secondly, can accuracy be improved through rain gauge calibration of the coincident radar pixel value? The paper uses different forms of calibration factor to correct the radar image. Two basic categories are long-term (static) factors and short-term (dynamic) factors. The long-term factor involves the calculation of a single corrective parameter from all the data and then applying it in an identical way to all time frames. Two types examined in [12] were the long-term arithmetic mean ratio bias and the geometric mean ratio bias. The use of long-term factors can be improved upon by employing dynamic calibration factors, which are recalculated for each time-frame, for instance every 15 minutes. The basis of the method is to fit a surface to calibration factor values estimated at a number of rain gauge locations and to scale the radar field by the coincident factor values to derive a more accurate calibrated radar field. Wood et al. further investigate a hybrid method which partitions the calibration factor into a spatially-uniform long-term component and a spatially-varying dynamic component. The resulting factor will take on the form of the dynamic calibration at short distances but behave like static calibration over long distances. Overall, this hybrid approach outperforms rain gauge, uncalibrated radar, and statically-calibrated radar estimates for the majority of rain gauges in the catchment. The detailed descriptions of the different factors and their performance can be found in [12].

### 2.3 Spatial-temporal models of rainfall fields

Mathematical modelling of rainfall fields themselves has potential value for use in radar calibration. The work by Wheater et al. [10] presents a number of different modelling techniques and tests their performance using the data from the HYREX experiment discussed above. We note that the models in [10] assume that calibration errors do not seriously bias the spatial and temporal structure of the observed fields. The next challenge would therefore be to combine modelling of rainfall with calibration of radar data. In [10], the authors merely use the data to investigate the influence on rainfall behaviour of various factors such as topography (in particular elevation), large-scale spatial variability, and seasonal effects.

The models used in [10] provide an explicit representation of observables such as the clustering and movement of rain cells. However, they do not attempt to capture the detailed deterministic evolution of the physical processes involved. Instead, this evolution is seen as governed by simple stochastic point processes determined by a small number of parameters with clear physical interpretation. Two main classes of models are spatial-temporal models and multi-site models. The former depict the temporal evolution of rainfall over a continuous spatial region
and are suitable for radar data; the latter do the same at a discrete set of locations in space, and thus are appropriate for studying rain gauge data. Both types are based on a hierarchical structure where rainfall fields occur in a temporal Poisson process, rain bands (storms) occur within each field in a spatial Poisson process, and rain cells occur in each storm clustering in time and space. Typically, all the components in the hierarchy move. Detailed descriptions together with the results of the analyses can be found in [10].

In the context of rainfall modelling, two other papers are worth mentioning, namely [4, 5]. There the authors investigate evidence for long-term changes in precipitation in the Galway Bay region of Western Ireland. The paper studies rainfall occurrence and rainfall amounts by fitting Generalised Linear Models.

3 Space-time models of radar and rain gauge data

In this section we discuss the results of Brown et al. contained in [3]. There the authors combine several modern statistical techniques in a novel way, with very promising results. They build an empirical space-time model to describe the relationship between radar reflectance and rainfall intensity. The modelling strategy reflects the fact that the gauge data are spatially sparse but temporally dense. First, a time series model is fitted to the data at each individual site. The time-space covariance structure of the minimum mean-square error predictors of the dynamic regression coefficients is then examined and used to formulate an integrated space-time model for the entire region. The model has been tested on data from a weather radar station at Hameldon Hill, Lancashire, England. The radar reflectance value at a gauge site is assumed to be the value at the nearest pixel centre. Also, reflectance values have been pre-processed to remove known anomalies.

Space-time models have recently been used in various environmental monitoring applications. It is common to have a discrete-time formulation, with a random variable $Y_{i,t}$ defined at unit times $t$ and a finite set of locations $i$ in some index set $I$. The time series $Y_t$ at the $|I|$ locations can be linked via a spatially correlated set of variables $Z_{it}$ with a covariance structure derived from an underlying continuous spatial process $Z(x)$ realised independently at each time $t$.

Other classes of models are so-called separable models, where a space-time process $Y(x, t)$ is decomposed as

$$Y(x, t) = M(x, t) + S(x, t)U(x, t).$$

Here $M(x, t) = F(x) + \eta(t)$, where $F(x)$ is a purely spatial process, called the mean-field; similarly $S(x, t) = H(x)\kappa(t)$, where $H(x)$ is the spatial field; also, $\eta(t), \kappa(t)$ are both white noise. Finally, $U(x, t)$ models residual space-time variation.

Another approach, very similar in spirit to that discussed in [3] is to represent $Y(x, t)$ as

$$Y(x, t) = S(x, t) + \epsilon(x, t),$$

where $S(x, t)$ incorporates space-time dependence whereas $\epsilon(x, t)$ is spatially and temporally uncorrelated.

In the future it would be useful to try a few different models of this form for the radar calibration problem and test their performance on real data. In order to improve on previous studies, the choice of representation should take account, in a more sophisticated way than has so far been used, of the physical processes involved.

3.1 Single-site model

Consider a single site where we record both gauge measurements and radar rainfall estimates at equally spaced time intervals through a rainfall event. The model assumes that the relationship
between gauge and radar reflectance measurements can be described by a power law,

\[ G_t = \alpha R_t^\beta \epsilon_t, \]

where \( G_t \) is the gauge measurement at time \( t \), \( R_t \) is the radar measurement and \( \epsilon_t \) is a multiplicative error term. Although previous research has shown that the spatially and temporally invariant power law is not a good fit, there is hope for improvement through introducing a suitable form of dependence upon time and space into the relationship.

Taking logarithms of (2), we obtain a linear relationship, with a special treatment of zero values. The zero value for radar reflectance is thought to give a very reliable indication that rain is not currently falling, that is \( R_t = 0 \) implies that \( G_t = 0 \) with high probability. Zero values may therefore be treated as missing data. The parameters \( \alpha \) and \( \beta \) may vary stochastically over time. We write down the linear observation equation

\[ Y_t = A_t + B_t u_t + Z_t, \]

where \( Y_t = \log(G_t) \), \( u_t = \log(R_t) \), and \( Z_t \) is an uncorrelated Gaussian noise with mean 0 and variance \( \sigma^2_z \). The stochastic processes \( A_t \) and \( B_t \) describe the evolution of the parameters \( \alpha \) and \( \beta \) above via the state equations

\[ A_t - \mu_A = \phi_A(A_{t-1} - \mu_A) + \eta_t, \]
\[ B_t - \mu_B = \phi_B(B_{t-1} - \mu_B) + \epsilon_t, \]

where \( \eta_t \) and \( \epsilon_t \) are uncorrelated Gaussian stochastic processes with zero means and respective variances \( \sigma^2_A \) and \( \sigma^2_B \), and \( \phi_A \) and \( \phi_B \) are parameters each lying in the range 0-1. The equations (4) and (5) define standard first-order autoregressive processes. Since both \( \phi_A \) and \( \phi_B \) each have modulus strictly less than 1, these processes are said to be stationary. This means that \( A_t \) and \( B_t \) hover around the constant means \( \mu_A \) and \( \mu_B \) respectively. Given the availability of a set of observations up to and including \( A_T \) and \( B_T \), the optimal predictor \( l \) steps ahead is the expected value of \( A_{T+l} \) and \( B_{T+l} \) conditional on the information available at time \( T \). This predictor is optimal in the sense that it has minimum square error - see [6] for a theoretical exposition of time series.

The time series in equations (4) and (5) are in state space form [6]. Therefore, the Kalman filter [6] can be applied to them, leading to efficient algorithms for prediction and smoothing. Furthermore, the Kalman filter enables the maximum likelihood estimation of the unknown parameters \( \sigma^2, \mu_A, \mu_B, \phi_A, \phi_B, \sigma^2_A \) and \( \sigma^2_B \).

The results reported in [3] appear to suggest that the assumption of a first-order autoregressive vector process is reasonable (errors are uncorrelated). The prediction intervals on the \( B \)-parameter remain wide, so that it is rather poorly identified. The authors' interpretation of this is that the predictive ability of the algorithm is not very sensitive to the power law parameter, and in subsequent analysis this parameter is treated as constant, denoted by \( \beta \).

3.2 Spatial Analysis

Brown et al. go on to suggest an integrated space-time model to describe the relationship between gauge measurements and radar reflectance values on a 2 km x 2 km pixel grid. The reflectance values are given at each of the \( N \) pixels whereas the gauge values are given only at \( n \) sites, where \( n \ll N \). Formally, the model is defined on the set of \( N \) pixels but the measurements \( Y_{it} \) at non-gauge sites are treated as missing data.
The Kalman filter will output maximum likelihood estimates of model parameters. Hence we can construct the minimum mean-square error predictors of the process $A_{it}$ at non-gauge sites and predict the unobserved values of $Y_{it}$ as

$$Y_{it} = 0, \quad \text{if} \quad R_{it} = 0,$$

$$Y_{it} = \hat{A}_{it} + \hat{\beta} u_{it}, \quad \text{if} \quad R_{it} > 0.$$

Since radar fields are almost continuous in space, the model should be constructed so that its parameters may be interpreted as parameters of a continuous space-time process $A(x, t)$. That is, one should be able to extrapolate the covariance structure to continuous space-time.

A complete spatial model can be formed by combining the single-site models into a first-order vector autoregressive model:

$$A_t - \mu = \Phi(A_{t-1} - \mu) + H \eta_t.$$  \hspace{1cm} (6)

The simplest form of this, with identical values of the parameters at all sites and excluding all spatial interaction, would be a model with $\Phi = \phi I$, and $H = \sigma^2 A I$. To build in spatial interaction, at least one of $\Phi$ and $H$ would have to be non-diagonal.

In the above model, the covariance structure is as follows:

$$\Gamma_k = \text{cov}(A_t, A_{t-k}),$$

$$\Gamma_k = \Phi^k \Gamma_0,$$

$$\Gamma_0 = \phi \Gamma_0 \phi' + HH^T.$$

The corresponding numerical values, for given data, will be output by the Kalman filter.

Again, the reader is referred to [3] for suggestions as to possible choices of the matrix $H$ and $\phi$. The model becomes computationally easier to handle for the choices that make it separable. In a separable model, the correlation function

$$\rho(x, t) = \text{corr}\{A(x_0 + x, t_0 + t), A(x_0, t_0)\},$$

factorises in the form $\rho(x, t) = \rho_1(x)\rho_2(t)$. However, it has been shown - see [7] - that a stationary Gaussian process $Y(x, t)$ has separable covariance structure if and only if the conditional expectation of $Y(x, t + 1)$ given the values of $Y(u, t)$ for all locations $u$ in a set including $x$ is equal to the conditional expectation given only the value of $Y(x, t)$. Thus separability can substantially restrict the allowed structure.

Brown et al. [3] analysed both separable and non-separable versions. Their computational results show the rainfall data to be well-approximated by the separable model. Furthermore, various diagnostic checks they conducted suggest no gross violations of the overall fit. The predictions were reasonably accurate, though only up to a certain time, and the level of accuracy was significantly lower at sites away from calibration gauges.

The aim of the reference [3] was to calibrate space-time radar reflectance values against tipping bucket gauge measurements, without use of any other explanatory variables. Within this framework their approach was to develop the simplest possible space-time model consistent with the data. One could improve on their findings by trying to combine statistical methods with greater understanding of the physical situation. It might be useful to include prevailing wind directions and topographic information as additional explanatory variables in the basic observation equation (6). Some of this extra information could also be included in the covariance structure. The current formulation absorbs almost all the physics into the latent stochastic process $A(x, t)$. Further, it might be fruitful to build a model of the raindrop size distribution as a function of the rain rate, and then use it to determine the functional dependence between radar reflectivity and rain rate.
4 Other methods of estimating rainfall intensity

A research group in the Engineering Department at the University of Essex has embarked on a project concerning the verification of the theory that the difference in the attenuation experienced by a pair of microwave links (operating at different frequencies along the same path) can, for certain frequency combinations, provide accurate path-averaged estimates of rain rate. A detailed introduction to the experiment is given in Holt et al. [8]. There are several pairs of microwave "links" (link=a microwave transmitter and receiver) and one wants to measure the attenuation (that is, signal loss) on each link, sampling every second. Using appropriate attenuation difference/rain rate relationships - see [8] - one can estimate path-averaged rainfall.

In the experiment, the longest link is 23.3 km, and the shortest is 8.9 km. Various frequency bands have been investigated, and for each pair of frequencies the specific attenuation difference has been calculated as a function of rain rate.

Additionally, data are available from a network of tipping-bucket rain gauges, and from these Holt et al. have calculated comparable path-averaged rainfall estimates. Subsequently, it becomes a matter of determining how accurate the link estimates are. There is a small amount of radar data as another rainfall estimate to compare with. So far, the research has experienced various problems; the main ones are as follows:

- The links are subject to non-rain induced fluctuations (due to, for instance, atmospheric effects, equipment effects, etc.). These fluctuations make it difficult to identify the baseline attenuation level from which to estimate rainfall.

- Clearly, there is no exact measure of rainfall: links, radar and gauges all provide estimates of varying quality. In particular, the measurement of rainfall in urban areas or in steep-sided valleys can be problematic, as it is difficult to find good locations for rain gauges, and radar cannot sense near to the ground. So in effect, as with all the other studies, one is always comparing different estimates without really knowing which one, if any, are actually correct.

Despite the difficulties, initial data tests have been quite encouraging, and further work is underway. The group are hopeful that the experiment, in combination with rain gauge and radar data can bring about improvements in storm water management and flood warning in the future. Detailed descriptions of the experiments already performed, and interpretations of the results can be found in [8].

Acknowledgement I would like to thank all other people who have worked on this problem for interesting discussions. I would also like to thank Alistair Fitt for a careful reading of the manuscript.

References


