**Hydrology and
Earth System
Sciences
Discussions**

Interactive
Comment

Interactive comment on "A space-time generator for rainfall nowcasting: the PRAISEST model" *by* P. Versace et al.

Anonymous Referee #2

Received and published: 7 May 2008

1 Overview

This paper presents a model for short-term forecasting of gridded rainfall. The model seeks to obtain estimates of rainfall height at a future timestep in terms of an auto-regression of heights at the same pixel from previous timesteps and in terms of a weighted average of heights at neighbouring pixels.

I find this manuscript to be one of the most confusing I have ever read. Firstly, the English grammar is poor and it is inappropriate to expect others to have to make menial corrections almost every sentence. Secondly the mathematical presentation is swollen as the authors have chosen a system of notation that maximises the amount of no-

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

tation and minimises the clarity of what they are saying. I feel that every part of this paper needs to be reworked: the literature review uses far too many references, the mathematical explanation needs to be clearer, the model calibration is hard to follow because of the notational style and there are very few statistics to verify that the model is performing adequately. While the underlying concept has potential, I do not consider this manuscript fit for publication short of extensive revision and radical improvement in the presentation. The manuscript requires 'major revisions' as a minimum. However, I also have some technical reservations about the reproduction of storm advection, the use of average parameter values and the limited number of validation statistics and for these reasons I think the manuscript should be 'rejected'. I can see that a large amount of work has gone into this model, but that the current manuscript is premature and does not rigorously present their work.

2 Major comments

2.1 Model development

I find that the use of subscripts and superscripts is too burdensome and that it is not necessary to constantly remind readers of which is 0+0, ++0, 00+, etc. for each of the parameters. The authors present joint probabilities but in a nowcasting framework it is much more intuitive to present conditional probabilities. For Section 2, I suggest something along the lines:

... ... Let H_i be the rainfall total on the interval $(i-1)\Delta t$ to $i\Delta t$ for a grid cell having area $\Delta x \Delta y$. For some lag in time, v, and for a given pixel, construct the auto-regression of rainfall heights

$$Z_i = \sum_{j=0}^{v-1} \alpha_j H_{i-j} \qquad (1)$$

where α_j are autoregressive coefficients to be estimated for each lag. The v coefficients can be estimated using standard likelihood estimation techniques.

Let W_i be a weighted average of rainfall heights over the four closest neighbouring pixels

$$W_i = \frac{1}{4} \sum_{j=1}^{4} H_i^{(j)} \qquad (2)$$

where the superscript indicates the pixels in the neighbourhood as shown in Figure XYZ. It is plausible to consider larger neighbourhoods or a different weighting procedure to allow for either anisotropy or influence from a larger region, but this is not considered here.

For each pixel the aim is then to predict the rainfall total H_{i+1} at a subsequent timestep from the pair of observed values Z_i, W_i. This gives rainfall triples (H_{i+1}, Z_i, W_i) from which the conditional distribution $f_H(H_{i+1}|Z_i, W_i)$ can be determined to perform the forecast. This method is an explicit scheme since both Z_i and W_i are known. An alternative formulation is to form the triple (H_{i+1}, Z_i, W_{i+1}) where W_{i+1} is a weighted average at the forecast timestep and this results in an implicit scheme. Section 4 outlines a methodology for simulating the implicit scheme. The two schemes are then compared in section XYZ to determine if one performs better than the other. (Note to authors: I did not find discussion on this. If one is not better than the other, then why bother with the implicit scheme as it is more difficult to manipulate?)

Rainfall can be thought of as a binary process of wet and dry pixels where a value is

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

observed if a pixel is wet. For now it suffices to ignore the time-increment subscripts and consider only whether the quantities W, Z and H are respectively wet or dry. The probability of a dry pixel, $P(H = 0)$, results from four separate cases:

P(H=0|Z=0,W=0), P(H=0|Z=z,W=0),
P(H=0|Z=0,W=w) and P(H=0|Z=z,W=w)

where for each case that satisfies the conditions on Z and W the observed proportion of dry pixels can be used as an estimator of the probability. In the event that the pixel is wet, the pdf of rainfall totals can be characterised by one of four separate cases:

f(H=h|Z=h,W=w), f(H=h|Z=z,W=0),
f(H=h|Z=0,W=w) and f(H=h|Z=0,W=0).

Each of these distributions is univariate and will differ due to the conditions of whether Z and/or W are zero or not. These distributions for the rainfall total H are now elaborated upon in detail.

In each of these cases the marginal distribution of the positive variables H, Z and W will be assumed to follow a Weibull distribution $F(x) = 1 - exp(-\lambda x^\eta)$ where the parameters η and λ can be estimated via the method of moments. The Weibull distribution is obtained from an exponential distribution with unit-mean $F(\acute{x}) = 1 - exp(-\acute{x})$ via the power transformation $\acute{x} = \lambda x^\eta$. It is important to emphasise that each conditioning case will have different power transformation parameters, as the distribution of one variable differs depending on whether the other variables are positive or zero. After performing a power transformation, to model the joint density of the variables it suffices to consider multivariate exponential distributions (Balakrishnan, 2001). The Moran-Downton multivariate exponential distribution for a vector **x** having p unit mean marginals is written as

$$f(\mathbf{x}) = \theta^{p-1} exp(\theta \Sigma_{i=1}^p \mathbf{x}_i) S_p((\theta - 1)\theta^{(p-1)} \Pi_{i=1}^p \mathbf{x}_i) \tag{3}$$

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

where θ is an association parameter that is the same between all dimensions and where

$$S_p(z) = \sum_{r=0}^{\infty} \frac{z^r}{(r!)^p}. \qquad (4)$$

The association parameter is related to the linear correlation coefficient between any two variables X_i and X_j as $\rho_{i,j} = 1 - 1/\theta$. Therefore the conditional density $f(h|z, w)$ can be written

$$f(h|z, w) = f(h, z, w)/f(z, w) = \theta_{hzw} exp(-\theta_{hzw}h) \frac{S_3((\theta_{hzw} - 1)\theta_{hzw}^2 hwz)}{S_2((\theta_{hzw} - 1)\theta_{hzw}wz)} \qquad (5)$$

where θ_{hzw} denotes that the association parameter is estimated from the observations (h, w, z). As the association parameter is the same in each dimension it is estimated from all of the pairs within this group (h, z), (h, w), (w, z).

The conditional density $f(h|0, w)$ can be written

$$f(h|0, w) = f(h, w|Z = 0)/f(w|Z = 0) = \theta_{hw} exp(-\theta_{hw}h) S_2((\theta_{hw} - 1)\theta_{hw}hw) \qquad (6)$$

where $f(h, w|Z = 0)$ is the density for all pairs (h, w) given the value from the auto-regression Z is zero. The association parameter, θ_{hw} is estimated from the observations that match this condition (i.e. it does not include pairs (h, w) where Z is non-zero). A similar expression for $f(h|z, 0)$ can be developed requiring a separate parameter θ_{hz} that is estimated from the pairs (h, z) when W is observed to be zero. The distribution $f(h|0, 0)$ is simply the univariate exponential distribution and does not require any additional parameters apart from the power transformation parameters.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

To forecast the rainfall total at a given pixel there are v autoregressive coefficients and 4 parameters representing the probability of zero rainfall under different conditioning scenarios. There are 6 power transformation parameters for $f(h|z,w)$, 4 transformation parameters for $f(h|z,0)$, 4 for $f(h|0,w)$ and 2 for $f(h|0,0)$, giving a total of 16 parameters. To correlate the respective joint distributions there are 3 association parameters θ_{hwz}, θ_{hw}, and θ_{hz}. As a result, the model requires 23 global parameters and v autoregressive parameters per pixel. ... ...

Hopefully this gives an alternative perspective on how the text could be made more concise. Please feel free to use this text or portions thereof in your manuscript as is appropriate. I have noted your earlier paper on the PRAISE model and that you have followed the notation on from that, but I feel that it becomes too cumbersome now that the spatial dimension has been added in PRAISEST. Furthermore, you have chosen to present joint densities which becomes quite confusing, when all that reader wants and needs is the necessary information to implement a forecast. As one example, what is the usefulness of an association parameter between Z and W when H_{i+1} is zero? If image i has a region where Z and W are non-zero, all that is required to obtain a zero value for H_{i+1} is the probability of occurrence and not the correlation of Z with W.

I will now proceed to outline other major comments on your paper specific to the various sections, to show more precisely the aspects that are confusing to follow.

2.2   Introduction

- You have about 65 references. This is at least double what you need. For example, line 9 on page 751 you have 8 references for ARMA models; on line 14 you have 16 references for intermittent models and on line 3 of page 752 you have 8 references for meteorological models! In each of these cases you need only 2 or 3 references at most. Please apply similar reduction to all other instances of multiple referencing.

- On line 14 of page 751 you reference quite a few point process models. Have any of these been used for forecasting? You should refocus your entire literature review to solely discuss nowcasting models without a detailed general review of stochastic models (and especially those that are not easily formulated for nowcasting).

2.3 The PRAISEST Model

- It is not clear whether the auto-regression (AR) has different parameters for each grid cell.

- Why isn't a likelihood function chosen to fit the AR(v) model? It is not clear why the gamma-power function is required. Is it to allow for skewness in the rainfall observations?

- It is not clearly indicated that the timestep is hourly (although it is implicit that short timesteps are being used if you are doing nowcasting). It would be helpful at this point to mention the timesteps this model is applicable for. As the timestep is hourly I have a serious reservation that your model makes no mention of storm advection! It is not clear how you handle this and whether a model without advection is appropriate. I am not convinced that the spatial neighbourhood approach is enough to capture, say, a frontal system.

- On page 755, line 1, r_{HZ} has not been defined. Is it the sample autocorrelation?

- On page 755, lines 11 to 15 are not needed. It is a simple weighted average. You can point out that other formulation schemes would be plausible in one sentence.

- Page 256, section heading: 'joint probability density' instead of 'joined probability density'. As discussed previously, presenting the joint density function is not a logical way to discuss a model that is to be used for forecasting.

- Page 256, lines 6 to 9: These lines are ugly. It is better to discuss case-by-case the probability of forecasting a dry pixel and then a wet-pixel rather than have a soup of symbols for the joint case.

- Equation 6 is not needed in a conditional setting. Also, you have too many super- and sub-scripts. So the equation is confusing in the first place. Equations 7 to 13 are not needed either, it is far too formal. I can get the meaning well enough from the (+,0,0) or the (H,0,0), so I do not need more than this.

- Equation 14 would be neater if the infinite summation was presented as a separate function as in Balakrishnan (2001). That way, the reader does not get distracted by the form of the function. This is especially true when it comes to Equation 18.

- Equations 15 to 17 are not needed. A simple statement of the power transform will suffice. It does not need to be repeated for each variable.

- Do not use α and β as transform parameters when you have already used these symbols for the autoregression and weighted average.

- It is not clearly explained where the leading term in Equation 18 comes from if the reader naively attempts to substitute Equation 15, 16 and 17 into Equation 14. This term arises when differentiating the Weibull cdf. Please see prior example of how I would explain this.

- Equations 19 and 20 are not needed. Just state that the method of moments can be used to estimate the parameters. This is straightforward enough for the Weibull distribution.

- Equation 21 seems to be an unnecessary result. Is it not sufficient to calculate the sample correlation for the various cases and then relate these to θ via $\hat{\rho} = 1 - 1/\theta$

- Equation 23 is unnecessarily cluttered. Please see prior example of how I would explain this using the multivariate Moran-Downton equation. Again, repeated use of power transform parameters is annoying, and also the presentation as the joint density is not as useful as the conditional $f(H = h | z = 0, W = w)$.

2.4 Model Calibration

- There are 42 parameters for every cell in the joint density model. Some of the parameters are not needed if you present the model in a conditional setting. I previously mentioned the ZW association parameter when H = 0 is not needed.

- You lump all observations over time to estimate parameters in a given cell. Please explain the rationale for this. Specifically, will the parameters be constant over time? Season to season? Storm to storm? Even within a storm it is possible that the parameters will evolve (growth and decay). What is the value of using average parameters?

2.5 Generation Algorithms

- Page 761, line 25: You mention that the cumulative distribution function $F(H_{i+1} \leq h_{i+1} | W_i = w_i, Z_i = z_i)$ which is used for simulation. This is my justification for demanding you use a conditional approach for each of the cases. You have laboured a joint density in the previous section, but then you do not use it explicitly.

- Page 761, line 24: 'The usual Monte Carlo techniques'? Which ones? This is one point that I would have liked more information on. The conditional cumulative density is not easily sampled from. Do you numerically evaluate it and integrate

it so that you can then use an inverse cdf to get a random sample? Do you use Markov Chain Monte Carlo? Do you use important sampling? The answer to this would be a practical help to implementing your model.

- What value is the implicit scheme? It is certainly harder to sample with this method than an explicit scheme. If there is no reason, then you should not present it. Does it preserve the spatial structure better in the H_{i+1} timestep than using an explicit scheme?

- If you use the explicit scheme, does this mean that the forecast value in each pixel is conditionally independent of it's neighbours? You obtain, Z and W for a pixel and then you sample a random number H_{i+1} for that pixel. The neighbouring pixel, while having similar Z and W, will have H_{i+1} independently sampled. Please explain.

2.6 Model Validation

- Page 764, line 26: You need to explain the convolution operations. I am not sure why they are needed.

- The number of statistics chosen to validate your model is insufficient. How does the model perform spatially? How well does it at reproducing autocorrelations and spatial correlations? How well does it reproduce overall dry ratios, wet-to-dry and dry-to-wet ratios for two subsequent images?

- From Figures 11 and 12 it does not seem that you correctly reproduce the portion of dry pixels.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

2.7 References

- Is it reasonable to have Italian references? These will not be able to be read by a majority of the HESS readership.

3 Minor comments

It is too much work to point out the minor grammatical comments. They appear in almost every sentence. Most universities provide trained professionals who are able to proof-read documents and assist those who use English as a second language. Please have your manuscript proof-read by a proficient English speaking person before considering resubmission.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 5, 749, 2008.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper