Journal of Hydrology (2007) **346**, 33–50

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jhydrol

Stochastic modelling of rainfall from satellite data

Chee-Kiat Teo [1], David I.F. Grimes *

TAMSAT, Department of Meteorology, University of Reading, P.O. Box 243, Earley Gate, Reading RG6 6BB, UK

Received 4 January 2007; received in revised form 15 June 2007; accepted 20 August 2007

KEYWORDS
Geostatistics;
Satellite;
Rainfall;
Uncertainty;
Sequential simulation;
Ensemble

Summary Satellite-based rainfall monitoring is widely used for climatological studies because of its full global coverage but it is also of great importance for operational purposes especially in areas such as Africa where there is a lack of ground-based rainfall data. Satellite rainfall estimates have enormous potential benefits as input to hydrological and agricultural models because of their real time availability, low cost and full spatial coverage. One issue that needs to be addressed is the uncertainty on these estimates. This is particularly important in assessing the likely errors on the output from non-linear models (rainfall-runoff or crop yield) which make use of the rainfall estimates, aggregated over an area, as input. Correct assessment of the uncertainty on the rainfall is non-trivial as it must take account of

- the difference in spatial support of the satellite information and independent data used for calibration
- uncertainties on the independent calibration data
- the non-Gaussian distribution of rainfall amount
- the spatial intermittency of rainfall
- the spatial correlation of the rainfall field

This paper describes a method for estimating the uncertainty on satellite-based rainfall values taking account of these factors. The method involves firstly a stochastic calibration which completely describes the probability of rainfall occurrence and the pdf of rainfall amount for a given satellite value, and secondly the generation of ensemble of rainfall fields based on the stochastic calibration but with the correct spatial correlation structure within each ensemble member. This is achieved by the use of geostatistical sequential

* Corresponding author. Tel.: +44 0 118 3786693; fax: +44 0 118 3788905.
E-mail addresses: CKTeo@ntu.edu.sg (C.-K. Teo), d.i.f.grimes@reading.ac.uk (D.I.F. Grimes).
[1] Present address. Temasek Laboratories@NTU, Research TechnoPlaza, 8th and 9th Storey, Border X Block, 50 Nanyang Drive, Singapore 637553, Singapore. Tel.: +65 63162990; fax: +65 67900215.

doi:10.1016/j.jhydrol.2007.08.014

simulation. The ensemble generated in this way may be used to estimate uncertainty at larger spatial scales. A case study of daily rainfall monitoring in the Gambia, west Africa for the purpose of crop yield forecasting is presented to illustrate the method.

Introduction

Algorithms to calculate rainfall from satellite images have been available since the 1960's (e.g. Lethbridge, 1967) and have been in routine use for more than 25 years. Satellite-based rainfall monitoring has become an invaluable source of information for investigating the global rainfall climate particularly over the ocean and in sparsely populated regions of the land surface where few ground-based measurements exist (Kidd, 2001). However, in many parts of the world satellite based monitoring also has important operational applications in drought monitoring, crop forecasting and flood warning which are only just beginning to be exploited. These applications are particularly relevant in Africa where ground-based raingauge networks are often inadequate both in terms of spatial coverage and timeliness of data (Washington et al., 2006; Hughes, 2006).

Research into the feasibility of using satellite-based rainfall estimates (SRFE) as input to hydrological or crop yield forecasting models is therefore of great practical concern in Africa and the results are also applicable in many other parts of the world. The use of SRFE as input to crop yield forecast models has been reported by Thornton et al. (1997), Reynolds et al. (2000), Verdin and Klaver (2002) and Senay and Verdin (2004), while the feasibility of using SRFE for water resource and flood prediction has been demonstrated by Tsintikidis et al. (1999), Andersen et al. (2002), Grimes and Diop (2003) and Sandholt et al. (2003).

An important question in the use of SRFE in these applications is the level of uncertainty ascribed to the estimates and the propagation of this uncertainty to other applications. A full assessment of the uncertainties is the main focus of this paper and is non-trivial for the following reasons

- rainfall accumulations typically have a non-Gaussian distribution
- comparison of SRFE with gauge data is complicated by the different spatial support of raingauge observations and satellite information
- raingauge data used for validation at the appropriate spatial scale will also have uncertainties which should be taken into account in assessment of SRFE uncertainty
- Treatment of uncertainties when aggregated to larger spatial scales must allow for the spatial correlation of the rainfall,

Although many studies (e.g., Laurent et al., 1998; Todd et al., 1999; Sorooshian et al., 2000) have made an assessment of the errors associated with different rainfall algorithms under specific conditions, only a few have taken account of the contribution to the estimated uncertainty of the data used for verification (for example Ali et al., 2005; Thorne et al., 2001) or the non-Gaussian distribution of rainfall accumulations. Stochastic methods of synthesizing rainfall fields from satellite data have been used to investigate uncertainties in satellite estimates but most (for example Bell, 1987; Bell et al., 1990; Astin, 1997; Steiner et al., 2003 Gebremichael and Krajewski, 2004) have focussed on quantifying the effects of limited satellite overpasses. There has been relatively little investigation of the uncertainty inherent in the stochastic nature of the relationship between the satellite proxy and the true rainfall. Work in this area includes Bellerby and Sun (2005), Hossain and Anagnostou (2006), and those reported in Fiorucci et al. (2001).

In terms of downstream applications, several studies show the importance of quantifying the SRFE uncertainties. For example, Guetter et al. (1996) and Tsonis et al. (1996) in a study of the Des Moines river basin in the United States showed differences in hydrological model output depending on whether rainfall input came from gauge data or SRFE, while Tsintikidis et al. (1999) have reported similar results for the Blue Nile. Sensitivity and non-linearity of crop models with respect to rainfall inputs have been discussed by Nonhebel (1994a,b), Hansen and Jones (2000), Heinemann et al. (2002) and Sultan et al. (2005).

The method of SRFE uncertainty estimation presented here involves a full statistical comparison of SRFE and daily raingauge data at satellite pixel scale. The statistical relationship serves as a kernel for generating an ensemble of possible rainfall fields consistent with the initial satellite-derived field. Full account is taken of spatial intermittency, spatial correlation and the non-Gaussian distribution of rainfall amount. The approach is similar to that of Bellerby and Sun (2005) in that it makes use of geostatistical techniques to generate an ensemble of realisations of the rainfall field which have the correct spatial correlation structure. It differs from their approach in that it explicitly allows for the spatial intermittency of the rainfall and uncertainty in the gauge data used for calibration and validation. The effect of intermittency can be important especially when considering rainfall at a daily or shorter time step. The ensemble approach also facilitates quantification of the uncertainty associated with large area averages obtained by upscaling the original pixel values. This is very useful for input to a lumped hydrological or agricultural model covering a catchment or crop growing district.

The methodology for uncertainty estimation is applicable to any SRFE algorithm. However, as a case study we have investigated the application to a TIR-based algorithm in west Africa. The specific motivation for this work stems from the need to assess the robustness of crop yield forecasts driven by SRFE for groundnut production in the Gambia in west Africa. This paper is concerned with the estimation of the uncertainty in the daily areal mean rainfall amounts for crop growing districts; the results of the crop yield forecasting are published elsewhere (Grimes and Teo, submitted for publication). The remainder of the paper is structured as follows: Section ''Methodology'' describes the rainfall simulation algorithm, Section ''Case study: simulation of daily