Hydrological Sciences Journal

ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage: https://www.tandfonline.com/loi/thsj20

A spatial disaggregation procedure for precipitation

THOMAS SKAUGEN

A spatial disaggregation procedure for precipitation

THOMAS SKAUGEN
*Norwegian Water Resources and Energy Directorate, PO Box 2091, Maj, N-0301 Oslo, Norway**
ths@nve.no

Abstract A disaggregation procedure is presented to render forecast values of precipitation from an atmospheric model with spatial resolution of 11 × 11 km suitable as input for a distributed hydrological model with spatial resolution of 1.1 × 1.1 km. Statistical and morphological properties of the input field, such as spatial mean, variance, correlation structure and intermittency, are respected in the disaggregated field. The adopted approach is a combination of interpolation and simulation. The four nodal points of the atmospheric model grid cell are used both for determining the parameters of the exponential distribution for simulating precipitation values, and in a simple interpolation procedure to determine the spatial location of the precipitation values. A shifted distribution with two parameters is used in the case of full coverage of the grid cell, and a one-parameter distribution with a theoretically derived intermittency parameter is used if intermittency is present. The results are promising with respect to the statistical and morphological properties of the disaggregated field.

Key words spatial disaggregation; precipitation; simulation; exponential distribution; intermittency; distributed models

Une méthode de désagrégation spatiale pour la précipitation

Résumé Une méthode de désagrégation pour la prédiction des valeurs de précipitation est présentée. La méthode repose sur un modèle atmosphérique de résolution spatiale 11 × 11 km, dont les résultats peuvent être utilisés comme données d'entrée pour un modèle hydrologique distribué de résolution spatiale 1.1 × 1.1 km. Les propriétés statistiques et morphologiques du champ de données, comme la moyenne, la variance, la structure de corrélation et l'intermittence spatiale, sont respectées dans le champ de désagrégation. L'approche adoptée est une combinaison d'interpolation et de simulation. Les quatre points nodaux des cellules du réseau du modèle atmosphérique sont utilisés à la fois pour la détermination des paramètres de la distribution exponentielle dans le cadre de la simulation des valeurs de précipitation, et pour une simple procédure d'interpolation pour déterminer la localisation spatiale des valeurs de précipitation. Une distribution à deux paramètres (dont un paramètre de localisation) est utilisée dans le cas où le réseau de cellules est complètement couvert, et une distribution à un paramètre, avec un paramètre d'intermittence théorique, est utilisée si l'intermittence est présente. Les résultats sont encourageants par rapport aux propriétés statistiques et morphologiques du champ désagrégé.

Mots clefs désagrégation spatiale; précipitation; simulation; distribution exponentielle; intermittence; modèles distribués

INTRODUCTION

Meteorological and hydrological processes are currently described on quite different spatial scales. Meteorological operational atmospheric models such as the HIgh Resolution Limited Area Model, HIRLAM (DNMI, 1996), use grid sizes of 11 × 11 km^2 and 50 × 50 km^2, whereas distributed hydrological models ideally should work on the scale of the hydrological response units. Hydrological response units can be described as

* *Also at:* Department of Geophysics, University of Oslo, PO Box, 1022, Blindern, N-0316 Oslo, Norway.

Open for discussion until 1 June 2003

patches in the landscape mosaic having a common climate, land use and pedological, topological and geological conditions controlling their hydrological process dynamics (Flügel, 1995). The spatial scale of the hydrological response units can be determined as the scale at which local variations are smoothened out, which for some Nordic landscapes has been shown to be in the range of 1–2 km^2 (Beldring *et al.*, 1999). The same problem is encountered in studies of possible climate change effects on systems measured and modelled on smaller scales than climate models. This discrepancy in spatial scales thus calls on methodologies for the disaggregation of meteorological data.

Considerable efforts have been made in order to improve the GCM (General Circulation Model) representation of the spatial variability of precipitation. Johnson *et al.* (1993) reported that by implementing the land surface hydrology parameterization of Entekhabi & Eagleson (1989), which introduces subgrid-scale spatial variability of precipitation and soil moisture, runoff generation and hence evaporation were greatly improved in the GCM maintained by NASA (National Aeronautics and Space Administration). Gao & Sorooshian (1994) investigated a fairly simple approach of introducing spatial variability to subgrid precipitation. The approach is based on two standard assumptions: (a) that precipitation processes are homogeneous over the GCM grid (spatial stationarity) and (b) that precipitation intensity within the rainfall area can be represented by an exponential distribution. Although these two assumptions were not found to be suitable for GCM grid-scale applications, they might be reasonable for smaller scales. The above assumptions are also adopted for the present study, although some improvements and alterations were made. The most significant improvement is that it is shown that if assumption (b) holds for individual grid cells, the dry fraction of the grid cell has a theoretical expression. This differs from several studies in which a fixed, precipitation-class related empirical value is assigned to the dry proportion of the grid cell although the non-zero precipitation is exponentially distributed (Dolman & Gregory, 1992; Johnson *et al.*, 1993; Liang *et al.*, 1996; Lammering & Dwyer, 2000). In this study the spatial statistical distribution of the disaggregated precipitation field is modelled as a mixture of exponential distributions, whereas the spatial dependencies (correlations) are maintained by an interpolation procedure.

METHODOLOGY

It is shown that a constant ratio between the spatial extent of precipitation coverages for different intensity thresholds implies that the spatial statistical distribution of precipitation is exponential. The derivation presented links to, and supplements, previous work of Skaugen *et al.* (1996), in which extreme areal precipitation was modelled and simulated by the use of precipitation coverages for different intensity thresholds. The exponential distribution is also a popular choice for describing the spatial distribution of precipitation for disaggregation purposes. This is partly because it only requires one parameter, but also because it is found as a quite suitable representation of precipitation (Eagleson, 1978; Gao & Sorooshian, 1994; Liang *et al.*, 1996; Schaake *et al.*, 1996; Onof *et al.*, 1998). It should be noted here that the proposed model is intended to be used for spatial disaggregation of daily precipitation and that it is assumed that the spatial statistical distribution is exponential for each grid cell and not necessarily for the entire precipitation field.

Let the areal precipitation over a grid cell A be formulated, as:

$$E(z)_A = \frac{1}{A}\int_A z(x)\mathrm{d}x \tag{1}$$

where $z(x)$ is the accumulated precipitation at a point x. If the fractional area (FA), a_τ, of the spatial coverage is calculated for discrete intensity thresholds, termed τ, $\tau = 0, 1, ..., T$, then equation (1) can be approximated by:

$$m(z)_A = \Delta\tau\sum_{\tau=0}^{T} a_\tau \tag{2}$$

where $\Delta\tau$ is the interval separating the discrete steps τ and $\Delta\tau = 1$ for the time being. The fractional area, a_τ, with precipitation more than τ mm is:

$$a_\tau = \frac{1}{A}\int_A I_\tau(x)\mathrm{d}x \qquad 0 \le a_\tau \le 1 \tag{3}$$

where I_t is the indicator function, $I_\tau(x) = \begin{cases} 1 \text{ if } z(x) \ge \tau \\ 0 \text{ otherwise} \end{cases}$

The following property of FAs is observed:

$$a_\tau > a_{\tau+1} > a_{\tau+2} \ldots > a_T \tag{4}$$

which simply states that, for any given field, the FA for a given threshold is larger than the FA for higher thresholds, where T is a maximum intensity when a_T approaches zero. The reduction of the FAs as the precipitation intensity increases is termed:

$$h_\tau = a_\tau / a_{\tau-1}, \qquad 0 < h_\tau < 1 \qquad \tau = 1, 2, ..., T \tag{5}$$

In principle, all the FAs can be expressed in terms of h. The FA for threshold $\tau + 1$ is $a_\tau h_{\tau+1}$, the consecutive, $a_{\tau+2}$, is $a_\tau h_{\tau+1} h_{\tau+2}$, etc., so, by combining equations (4) and (5), the expression for the mean areal precipitation can be written as:

$$m(z)_A \approx a_0\Delta\tau(\sum_{\tau=1}^{T}\prod_{j=1}^{\tau} h_j) \tag{6}$$

In Skaugen *et al.*(1996), FAs and their ratios h, were investigated for extreme rainfall events over catchments of different sizes in southern Norway. The mean value and the standard deviation of the ratio h, were found to be independent of thresholds and catchment size. Also, the correlation between the h for consecutive thresholds τ, was found to be weak. On the basis of these observations, the assumption is made that h is a stochastic variable, independent and identically distributed (iid) across the thresholds with mean $\bar{h}$. If, further, a_0 is equal to 1, i.e. it rains everywhere within the grid cell, one can write equation (6) as:

$$m(z)_A \approx \Delta\tau(\sum_{\tau=1}^{T}\bar{h}^{\tau}) \tag{7}$$

If one considers a regular partition of $(0, T)$ so that $(T - 0)/n = \Delta\tau$, then equation (7) is recognized as equivalent to a Riemann sum and the integral of a continuous function is defined by:

$$m(z)_A \approx \lim_{\Delta\tau \to 0} \left[\Delta\tau\left(\sum_{i=1}^{n} \overline{h}^{\tau_i^*}\right) \right] = \int_0^T \overline{h}^{z} \mathrm{d}z \tag{8}$$

where τ_i^* is in $[\tau_{i-1}, \tau_i]$. If T in equation (8) approaches infinity, the integrand in equation (8) can be regarded as the complementary cumulative distribution function (cdf) of z:

$$F(z) = 1 - \overline{h}^{z} \qquad 0 < z < \infty \tag{9}$$

The probability density function (pdf) is then:

$$f(z) = -\log(\overline{h})\overline{h}^{z} \tag{10}$$

Writing $-\log(\overline{h}) = \lambda$ and hence $\overline{h} = \mathrm{e}^{-\lambda}$ and substituting these expressions into equation (10), the familiar exponential distribution $f(z) = \lambda \mathrm{e}^{-\lambda z}$ with moments $E(z) = 1/\lambda$ and $\mathrm{var}(z) = 1/\lambda^2$ is obtained. When a description of fractional areas with constant ratio $\overline{h}$, is appropriate, then the spatial statistical distribution of precipitation is an exponential distribution with parameter $\lambda = -\log(\overline{h})$.

At this point, it is convenient to describe the pattern of the fractional areas for precipitation seen from a grid-cell point of view. There are two possible outcomes of a precipitation event over the grid cell. The grid cell might be fully covered with precipitation, for which there is a minimum positive precipitation intensity in the grid cell and the fractional areas for that intensity and less, are equal to one. This minimum intensity is denoted b, i.e. $a_k = 1$ for all $k\Delta\tau \leq b$. The second outcome is that there is an intermittent field, where only a fraction of the grid cell is covered by precipitation. The spatial distribution of precipitation can, for both these cases, be described by the exponential distribution with, for each case, the introduction of an additional parameter.

Case of complete coverage of a grid cell

The positive, minimum intensity b will serve as a location parameter of the exponential distribution, so the pdf is:

$$f(z) = \lambda \mathrm{e}^{-\lambda(z-b)} \qquad b < z < \infty \tag{11}$$

with moments:

$$E(z) = b + \frac{1}{\lambda} = b + \frac{-1}{\log(\overline{h})} \tag{12}$$

$$\mathrm{var}(z) = \frac{1}{\lambda^2} = \frac{1}{\log^2(\overline{h})} \tag{13}$$

Case of partial coverage of a grid cell

The point of departure for this case is the *a priori* knowledge of the unconditional mean and variance (moments including zeros), which are derived from the nodal points of the grid cells. From this, one can estimate the conditional mean and variance (for the positive precipitation values) and the dry fraction of the grid cell.

Let z and z' denote precipitation including and not including zeros, respectively. Then one obtains the moments:

$$E(z) = \frac{n-m}{n}0 + \frac{m}{n}E(z') = pE(z') \tag{14}$$

where $p = m/n$ is the fraction of the grid cell of positive precipitation and similarly:

$$E(z^2) = \frac{n-m}{n}0 + \frac{m}{n}E(z'^2) = pE(z'^2) \tag{15}$$

and the variance:

$$\text{var}(z) = E(z^2) - E(z)^2 \tag{16}$$

One can substitute equation (15) into equation (16):

$$\text{var}(z) = pE(z'^2) - E(z)^2 \tag{17}$$

and if the spatial distribution of z' is assumed exponential with parameter λ, then $E(z'^2)$ can be expressed in terms of $E(z)$ by using the fact that for the exponential distribution, $E(z'^2) = 2E(z')^2$, and by means of equations (14) and (17), one obtains:

$$\text{var}(z) = \frac{2}{p}E(z)^2 - E(z)^2 \tag{18}$$

which gives the fraction p, of positive precipitation within a grid cell as:

$$p = \frac{2}{\frac{\text{var}(z)}{E(z)^2} + 1} \tag{19}$$

It is appropriate here to discuss the ratio between the unconditional variance and the square of the unconditional mean in equation (19) and to relate this ratio to the spatial structure of rainfall. For a non-shifted exponential distribution, the spatial standard deviation is equal to the spatial mean and the ratio is equal to one. From equation (19) this corresponds to $p = 1$, i.e. complete coverage. When this ratio is higher than one, one can again see from equation (19) that the values of p are in the interval [0,1], i.e. the field is intermittent. When the ratio is less than one, precipitation cannot be exponentially distributed unless there is a minimum intensity, b, implying a shifted exponential distribution with moments shown in equations (12) and (13).

A DISAGGREGATION SCHEME

The theoretical platform presented above provides the necessary tools to establish a disaggregation scheme for precipitation. The use of an exponential distribution as the spatial statistical distribution of precipitation has been justified. In cases where a grid cell is only partly covered by precipitation, the exponential distribution allows a theoretical expression for the wet fraction, p, of the grid cell. Also, in cases where a minimum intensity is higher than zero within the grid cell, an exponential distribution is used with a location parameter b. In the following, a general description of the disaggregation procedure of *one* grid cell is first presented. Then, a more detailed procedure is outlined, which describes how a field of precipitation forecast from the HIRLAM model (11×11 km^2) is disaggregated into a field of 1.1×1.1 km^2.

General description of disaggregation procedure

The precipitation field of the HIRLAM model consists of N grid cells (11×11 km), while the resulting disaggregated field consists of $N \cdot I$ pixels (1.1×1.1 km), where I is the partitioning of the grid cell, the number of pixels. Let the interpolated values of pixels in a grid cell be expressed as:

$$\boldsymbol{Y} = \boldsymbol{AX}$$

where $\boldsymbol{Y}$ is a $(I \times 1)$ vector, $\boldsymbol{A}$ is a $(I \times J)$ matrix of weights derived by the chosen interpolation method, and $\boldsymbol{X}$ is a $(J \times 1)$ vector, of the input values; J is the number of input values (i.e. the four HIRLAM nodal values). The vector $\boldsymbol{Y}$ is obtained simply by partitioning the grid cell into pixels and assigning each pixel with an interpolated value. The components of the weight matrix $\boldsymbol{A}$, are derived from the chosen interpolation method. Inverse distance technique was used here, so that a fixed set of weights could be assigned for each pixel. Let $\boldsymbol{V}$ be a $(I \times 1)$ vector of ordered simulated values, drawn from an exponential distribution (simulated with parameters determined from $\boldsymbol{X}$) and sorted into ascending order. Then, if $\boldsymbol{rankY}$ is a vector of the ranks of the interpolated input $\boldsymbol{Y}$, the disaggregated field $\boldsymbol{Z}$, can be expressed as:

$$\boldsymbol{Z}[i] = \boldsymbol{V}[\boldsymbol{rankY}[i]] \text{ for all } i = 1, ..., I \tag{20}$$

where the brackets [] indicate the components of the vector and also the spatial location (e.g. $i = 1$ is the lower left pixel of the grid cell). The result can be described as a field with exponentially distributed values of non-zero precipitation, with unconditional mean and variance identical to that of the input values (the four HIRLAM nodal values), and with spatial dependencies inherited from the interpolated field $\boldsymbol{Y}$.

The general procedure described above is local in that the interpolation of the pixel values and the simulation of precipitation values (and intermittency fractions) are carried out according to the nodal values of an actual grid cell. To disaggregate an entire HIRLAM field, one could repeat this procedure according to the number of grid cells. However, this scheme provides a local maximum for each grid cell, which gives a very peaked and irregular image of the spatial event (Skaugen, 2001). One could also consider a strictly global approach, in which the simulation of precipitation values, intermittency fractions and the interpolation was carried out based on global estimates of the parameters. This was not investigated, because it could readily be determined from the global statistical parameters and fraction of zeros of several events that the exponential distribution did not apply as a global spatial statistical distribution. This leaves a kind of in-between solution where the simulation of precipitation values is done locally and the distribution of precipitation values in space is done globally according to the global rank of locally interpolated pixel values.

Detailed description of disaggregation procedure

1. Each of I pixels in the grid cell is assigned a value interpolated from the nodal values (the four corner values). For simplicity, the inverse distance method (of power 2) has been used in this study. This procedure is repeated for the N grid cells. The pixels are, from the interpolated values, assigned a rank (1 to $N \cdot I$) in order to determine the relative magnitude of precipitation and their location within the interpolated precipitation field. The value of $\boldsymbol{rankY}$ of equation (20) is thus determined, but with $N \cdot I$ components.
2. The unconditional spatial mean and spatial variance of the grid cell are estimated from the (four) nodal values of the grid cell.
3. The ratio between the unconditional variance and the squared unconditional mean (see equation (19)) is evaluated. If this ratio is higher than one, the precipitation field is intermittent with fractional coverage p determined from equation (19) and positive precipitation within the grid cell exponentially distributed, $f(z; \lambda)$. If the ratio is less than one, the grid cell is fully covered with minimum intensity b, and exponentially distributed precipitation, $f(z; \lambda, b)$.
4. In the case of intermittency for grid cell n, $p_n \cdot I$ values are simulated from $f(z; \lambda)$ and $(1 - p_n)I$ are assigned the value zero. In the case of full coverage, I values are simulated from $f(z; \lambda, b)$.
5. Points 2. to 4. are repeated for every grid cell 1, ..., N.
6. The $N \cdot I$ simulated values are then ordered and the $I \cdot N$-dimensional vector $\boldsymbol{V}$ is obtained. The disaggregated field is thus determined as:

$$\boldsymbol{Z}[i] = \boldsymbol{V}[\boldsymbol{rankY}[i]] \text{ for all } i = 1, ..., I \cdot N \tag{21}$$

In the case of intermittency, the $\sum_{n=1}^{N}(1 - p_n)I$ lowest ranked pixels are assigned the value zero.

CASE STUDY AND DISCUSSION

The simulation procedure was carried out for 15 months of daily HIRLAM forecasts for the Gaula region (5082 km^2) in central Norway, south of the city Trondheim (see map in Fig. 1). From the 15 months, 10 events are selected because they represent a variety of situations that have to be adequately represented by the disaggregation procedure. Both heavy frontal events and events with intermittency and cell structure are included. The disaggregation scheme should be able to reflect the following statistical properties of the precipitation fields: spatial mean, spatial variance, intermittency and spatial correlation structure. Figures 2 and 3 show the HIRLAM fields and the disaggregated fields of three events and Table 1 provides the statistics of the HIRLAM fields and disaggregated fields for the 10 events. The values for the disaggregated fields are the means of 10 repeated disaggregation simulations. The global means are near perfectly reproduced for all the events, and one can observe a slight increase in the global standard deviation of the disaggregated field of the order of 5–9%. The standard errors of the means of the simulations are small, indicating a stable simulation procedure. Not surprisingly, the largest standard errors are found for

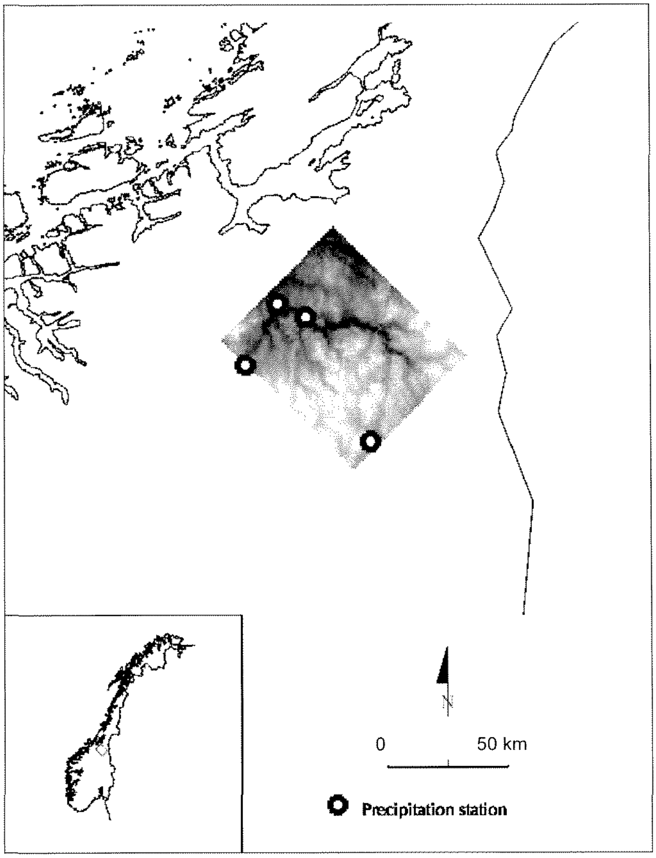

Fig. 1 Study area in central Norway. The location of four precipitation gauges is marked.

the maximum values. An interesting aspect of the usefulness of disaggregation can be observed in Table 1. Comparing the two events of 1 December 1999 (991201) and 6 August 2000 (000806), the mean and the maximum values of the HIRLAM field are largest for the latter event, while the maximum value of the disaggregated field is largest for the event of 991201. Apparently the configuration of the HIRLAM values (the gradients) of the event of 991201 is such that an enhanced local variability is detected by the disaggregation procedure, and thus very high values are simulated. The hydrological importance of revealing the potential magnitude and location of maximum values of the precipitation fields is evident, especially for flood forecasting purposes. With a distributed hydrological model, the timing and the magnitude of flash floods can be better forecast with disaggregated precipitation fields. The disaggregated fields also reveal a point, regarding the representativity of precipitation gauges. One can clearly see how all the precipitation gauges are situated in areas where relatively modest precipitation intensity is forecast. Given that the HIRLAM model captures and reproduces realistic spatial patterns of precipitation, it is obvious that e.g. estimates of areal precipitation based on these four precipitation gauges will be seriously underestimated.

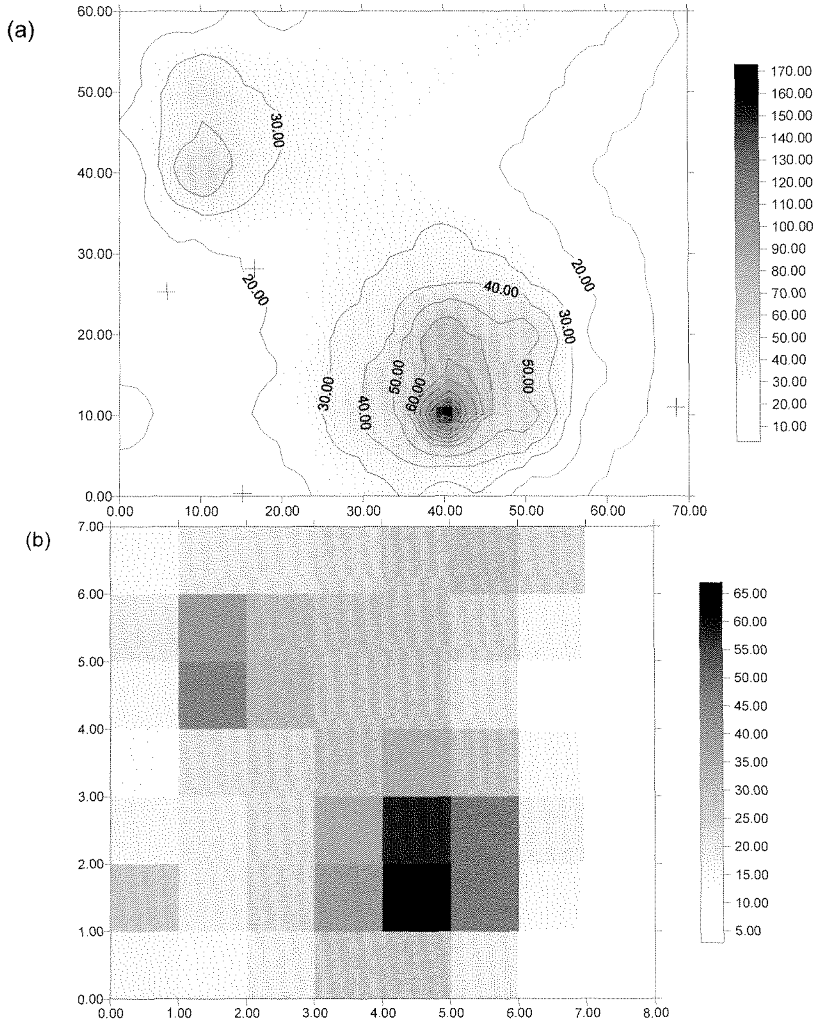

Fig. 2 (a) HIRLAM field and (b) disaggregated field of 29 September 1999 (+: location of precipitation gauge).

Comparisons against observed values

The actual magnitude of the maximum values of the disaggregated field is difficult to assess, because no observations of the maxima in real rainfall fields are available. Also, when validating the disaggregation procedure against observations, it is difficult to distinguish between errors due to the disaggregation procedure and errors due to the atmospheric model HIRLAM. However, one can analyse the disaggregation results by constructing quantile–quantile plots (q–q plot) of the time series of disaggregated rainfall at a pixel, a HIRLAM grid cell and observed precipitation at a rainfall gauge. Here, the quantiles of the observed rainfall are plotted against the corresponding quantiles of the disaggregated rainfall in a pixel located near the rainfall gauge and the quantiles of the HIRLAM grid cell in which the rainfall gauge is located (see e.g.

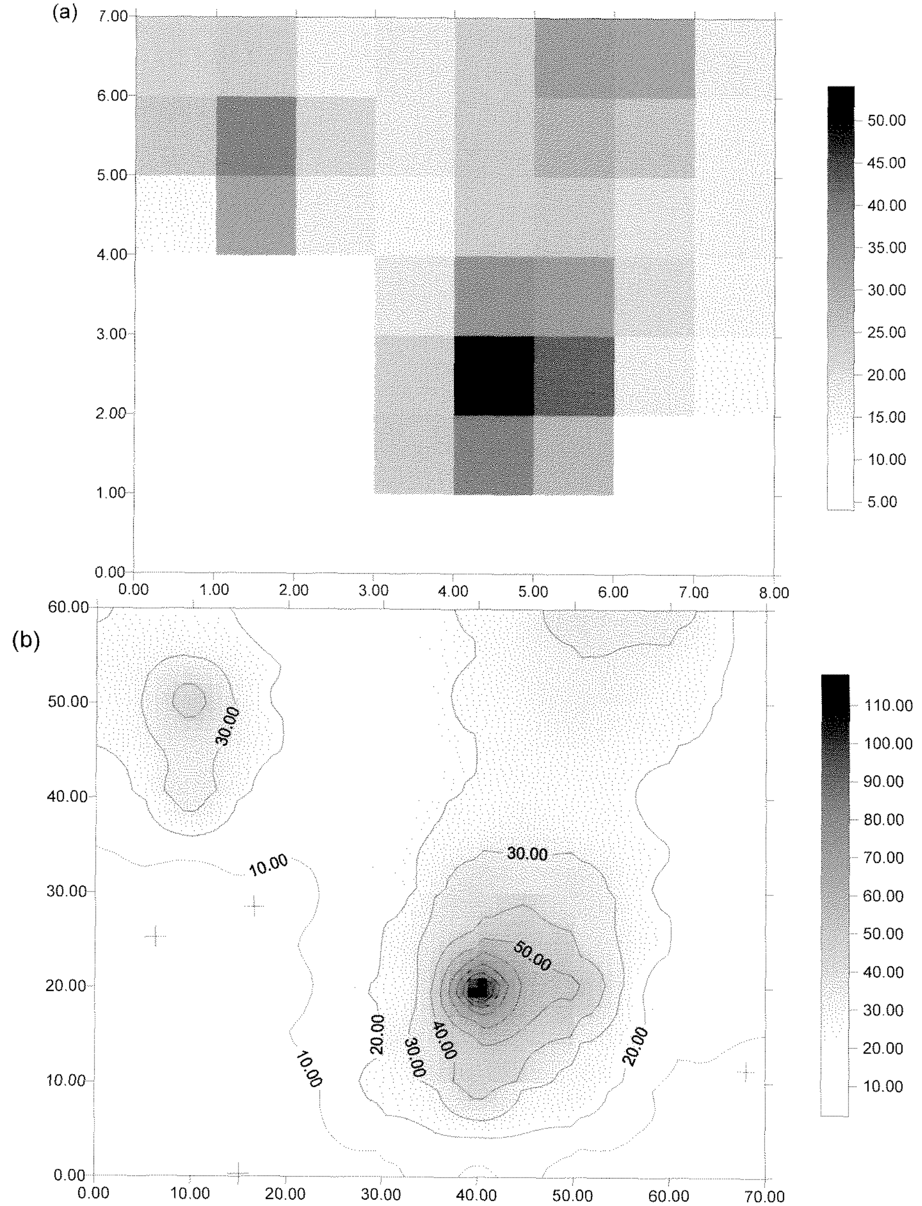

Fig. 3 (a) HIRLAM field and (b) disaggregated field of 10 November 2000 (+: location of precipitation gauge).

Fig. 2 for the locations). If the temporal statistical distribution of either the HIRLAM values or the disaggregated results is equal to that of the observed values, the q–q plot shows the points located on the diagonal line. Figure 4 shows q–q plots for four rainfall gauges located within the study area. With one exception (Fig. 4(c)), the disaggregated values are in better correspondence with the observed than the HIRLAM values. Figure 4 also illustrates another point concerning both the HIRLAM values and

Table 1 Global statistical parameters of precipitation from HIRLAM (HA) and the disaggregation scheme (DA). The values for the disaggregated scheme are the mean of 10 disaggregation runs. Numbers in brackets refer to standard errors of the mean. Dates are expressed as yymmdd.

Date	Source	Mean	Standard dev.	Fraction of zeros	Max.	Min.
990929	HA	24.86	13.39	0	67	3
	DA	24.76 (0.087)	14.21 (0.073)	0	143.6 (5.77)	3.16 (0.02)
991025	HA	0.76	1.94	0.05	8.8	0
	DA	0.76 (0.006)	2.08 (0.027)	0.11	28.2 (1.51)	0
991118	HA	2.31	5.30	0.023	27.57	0
	DA	2.31 (0.018)	5.62 (0.076)	0.025	67.9 (4.93)	0
991201	HA	14.35	9.65	0	42.48	0.59
	DA	14.33 (0.039)	10.14 (0.083)	0	110.5 (9.09)	0.53 (0.006)
000428	HA	0.23	0.40	0.20	1.81	0
	DA	0.23 (0.001)	0.42 (0.004)	0.15	4.8 (0.24)	0
000806	HA	19.61	11.20	0	53.13	4.96
	DA	19.56 (0.025)	11.80 (0.05)	0	104.6 (4.19)	2.48 (0.016)
000815	HA	31.17	13.61	0	62.64	8.86
	DA	31.10 (0.036)	14.24 (0.063)	0	145.7 (5.80)	8.77 (0.013)
000919	HA	0.42	0.64	0.26	2.76	0
	DA	0.42 (0.002)	0.68 (0.007)	0.21	7.2 (0.42)	0
000925	HA	0.041	0.046	0.25	0.17	0
	DA	0.041 (0.000)	0.051 (0.000)	0.14	0.50 (0.014)	0
001110	HA	0.64	1.28	0.035	5.49	0
	DA	0.64 (0.004)	1.35 (0.008)	0.061	16.3 (0.733)	0

the disaggregated values. The frequency of zeros is too low and the frequency of small values is too high compared to the observed series. This is expected, because, after all, the HIRLAM values and the disaggregated values represent areal averages. However, it is believed that the atmospheric model HIRLAM has a tendency to be reluctant to forecast zero rainfall in intermittent fields (K. H. Midtbø, DNMI, personal communication). This property is carried over to the disaggregation procedure as the local variability has been estimated using grid-cell values (averages) from the HIRLAM model as point observations. This problem is basically a scaling issue, which involves the rather complex, region-specific interactions of spatial scale, spatial variability and frequency of precipitation. The variance-reducing effect due to areal averaging is well known (Rodriguez-Iturbe & Meija, 1974), but as such relationships have to be investigated and determined at specific sites with very detailed observations, it has not been taken specifically into account in this study. From equation (19) it can be seen that the fraction of dry pixels, $(1-p)$, would increase with increased variance of the nonconditional rainfall, which would be expected for a disaggregated field.

Spatial correlation

Figure 5 shows the spatial correlation structure of the HIRLAM fields and the disaggregated fields for the same events as in Figs 2 and 3. The correlation points plotted constitute the upper right quadrant of the correlogram. The general correlation structure of the HIRLAM model is well reproduced by that of the disaggregated field for both events. It is interesting to note that the correlation structure is quite different

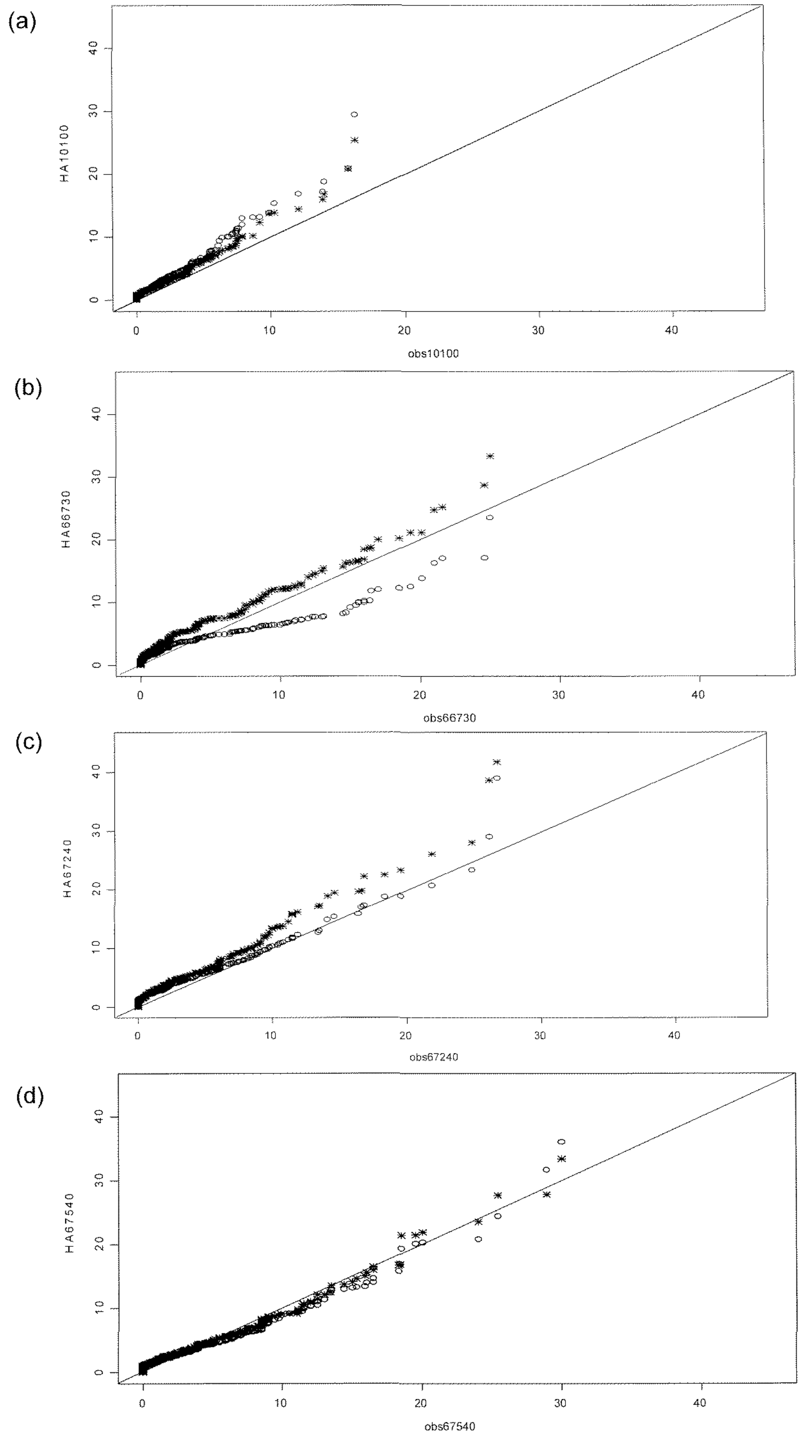

Fig. 4 Quantile–quantile plots of time series of disaggregated *vs* observed (*) and HIRLAM *vs* observed (o) for the rainfall stations (a) 10100, (b) 66730, (c) 67240 and (d) 67540.

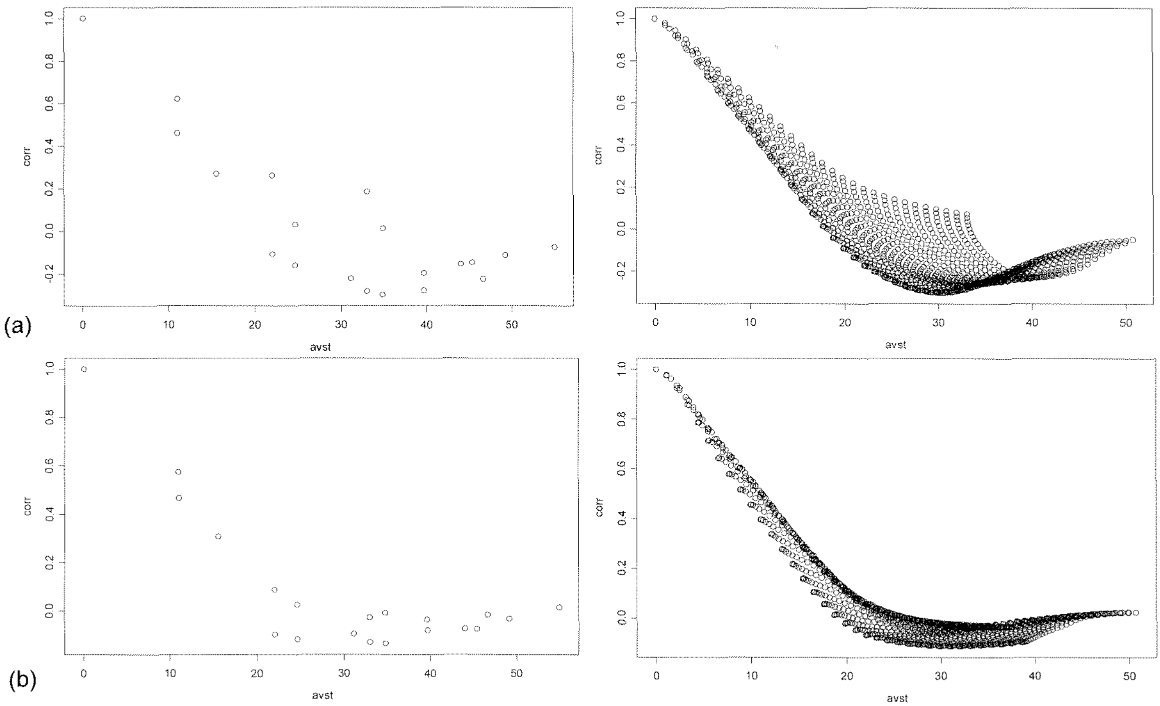

Fig. 5 Spatial correlation structure of HIRLAM field (left) and disaggregated field (right) (a) on 29 September 1999 and (b) on 10 November 2000. The x-axis represents distance in km.

from event to event, varying in scatter according to the type of precipitation event. This should be taken into consideration when assessing different interpolation methods based on the spatial correlation structure (such as various forms of kriging). This is consistent with the findings of Skaugen (1997), where different correlation structures for kriging interpolation were used based upon a classification of the precipitation process into small-scale (showers) and large-scale (frontal precipitation).

CONCLUSIONS AND PROJECTIONS FOR FUTURE RESEARCH

A disaggregation scheme is put forward which adopts the spatial dependencies derived from interpolation and gives the global statistical distribution of precipitation as a mixture of exponential distributions. The exponential distribution has been found to represent both events in which intermittence in grid cells occurs, and events in which the minimum intensity in the grid cell is higher than zero. The disaggregated fields are promising when compared to the aggregated fields. The spatial mean and spatial correlation structure are very similar to those of the HIRLAM field, while an expected increase in the spatial standard deviation is found for the disaggregated field. A theoretical expression is derived for the fraction of dry area, given that the non-zero precipitation is exponentially distributed. For such intermittent rainfall fields, the disaggregation procedure reproduces the fraction of zero rainfall of the HIRLAM field. However, this intermittence is considered to be too low to represent real rainfall fields. Future research is projected to investigate how different estimation procedures of local variability can improve the estimation of zero rainfall areas. Another projected topic for future study is to use the disaggregation scheme in order to simulate the spatial distribution of snow, snowmelt and snow coverage and their temporal evolutions through a winter season. It is hoped that this will provide a better understanding of how snow should be modelled in rainfall–runoff models.

Acknowledgements The help in preparation of the HIRLAM data of the Norwegian Meteorological Institute and some of the figures by Hans Christian Udnæs of the Norwegian Water Resources and Energy Administration is greatly appreciated. The thoughtful comments of Dr Christian Onof of Imperial College, London have been very helpful. The author would like to thank the Norwegian Research Council and the Norwegian Electricity Industry Association for financial support of this project.

REFERENCES

Beldring, S., Gottschalk, L. T., Seibert, J. & Tallaksen, L. M. (1999) Distribution of soil moisture and groundwater levels at patch and catchment scales. *Agric. For. Met.* **98–99**, 305–324.

DNMI (The Norwegian Meteorological Institute) (1996) HIRLAM documentation manual, HIRLAM project. The Norwegian Meteorological Institute, Oslo, Norway.

Dolman, A. & Gregory, D. (1992) The parameterization of rainfall interception in GCMs. *Quart. J. Roy. Met. Soc.* **118**, 455–467.

Eagleson, P. S. (1978) Climate, soil and vegetation. The distribution of annual precipitation derived from observed storm sequences. *Wat. Resour. Res.* **14**(5), 713–721.

Entekhabi, D. & Eagleson, P. (1989) Land surface hydrology parameterization for atmospheric general circulation models including subgrid scale spatial variability. *J. Climate* **2**, 816–831.

Flügel, W. A. (1995) Delineating hydrological response units by geographical information system analysis for regional hydrological modelling using PRMS/MMS in the drainage basin of the river Bröl, Germany. In: *Scale Issues in Hydrological Modelling* (ed. by J. D. Kalma & M. Sivapalan), 181–194. Wiley, Chichester, West Sussex, UK.

Gao, X. & Sorooshian, S. (1994) A stochastic precipitation disaggregation scheme for GCM applications. *J. Climate* 7(2), 238–247.

Johnson, K., Entekhabi, D. & Eagleson, P. (1993) The implementation and validation of improved land-surface hydrology in an atmospheric circulation model. *J. Climate* **6**, 1009–1026.

Lammering, B. & Dwyer, I. (2000) Improvement of water balance in land surface schemes by random cascade disaggregation of rainfall. *Int. J. Clim.* **20**, 681–695.

Liang, X., Lettenmaier, D. P. & Wood, E. F. (1996) One-dimensional statistical dynamic representation of subgrid spatial variability of precipitation in the two-layer variable infiltration capacity model. *J. Geophys. Res.* **101**(D16), 21403–21422.

Onof, C., Mackay, N. G., Oh, L. & Wheater, H. S. (1998) An improved rainfall disaggregation technique for GCMs. *J. Geophys. Res.* **103**(D16), 19 577–19 586.

Rodriguez-Iturbe, I. & Meija, J. M. (1974) On the transformation of point rainfall to areal rainfall. *Wat. Resour. Res.* **10**(4), 729–735.

Schaake, J. C., Koren, V. I., Duan, Q-Y., Mitchell, K. & Chen, F. (1996) Simple water balance model for estimating runoff at different spatial and temporal scales. *J. Geophys. Res.* **101**(D3), 7 461–7 475.

Skaugen, T. (1997) Classification of rainfall into small- and large-scale events by statistical pattern recognition. *J. Hydrol.* **200**, 40–57.

Skaugen, T. (2001) A simple disaggregation scheme for precipitation. *Geophys. Res. Abstracts*, vol. 3.

Skaugen, T., Creutin, J-D. & Gottschalk, L. T. (1996) Reconstruction and frequency estimates of extreme daily precipitation. *J. Geophys. Res.* **101**(D21), 26 287–26 295.

Received 16 November 2001; accepted 16 July 2002