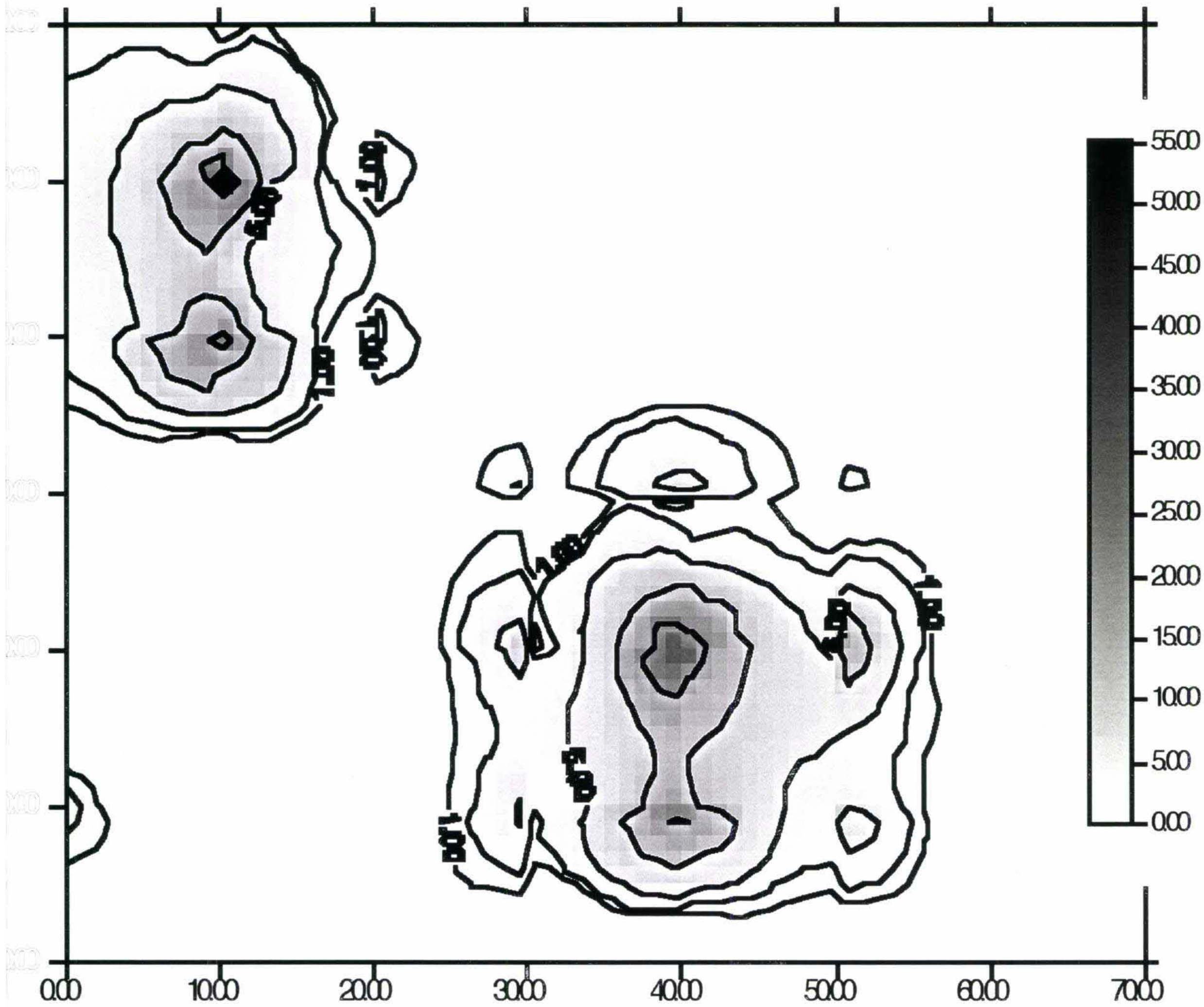
NVE

Disaggregation of precipitation

Thomas Skaugen

2
2001

OPPDRAGSRAPPORT



Disaggregation of precipitation

Norges vassdrags- og energidirektorat

2001

Oppdragsrapport nr 2

Disaggregation of precipitation

Oppdragsgiver:	SINTEF BM
Redaktør:	
Forfatter:	Thomas Skaugen

Trykk:	NVEs hustrykkeri
Opplag:	30
Forsideillustr.:	Eksempel på disaggregering av nedbør.

Sammendrag: A method for disaggregation of forecasted precipitation is developed and demonstrated.

Emneord: Precipitation, disaggreagation, distributed models

Norges vassdrags- og energidirektorat
Middelthuns gate 29
Postboks 5091 Majorstua
0301 OSLO

Telefon: 22 95 95 95
Telefaks: 22 95 90 00
Internett: www.nve.no

Januar 2001

Innhold

Forord

Prosjektet 'Kobling av hydrologiske og meteorologiske modeller' er et samarbeidsprosjekt mellom SINTEF, DNMI og NVE. Prosjektet er finansiert av ENFO og Norsk Forskningsråd og er planlagt gjennomført i perioden 2000-2002. Hovedmålet for prosjektet er å forbedre hydrologiske og meteorologiske prognoser. Dette skal utprøves ved at den meteorologiske modellen HIRLAM oppdateres med data fra den hydrologiske modellen LANDPINE. Prosjektleder er Trond Rinde, SINTEF BM.

Denne rapporten inngår som statusrapport for delprosjekt 5 av 6 delprosjekter som har vært gjennomført i år 2000. Delprosjekt 1 og 2 er gjennomført av SINTEF BM. Delprosjekt 4 er gjennomført av DNMI. Delprosjekt 3, 5 og 6 er gjennomført av NVE. Prosjektmedarbeidere på NVE har vært Thomas Skaugen, Liss M. Andreassen, Elin Langsholt og Hans-Christian Udnæs.

I NVE's delprosjekter er det benyttet HIRLAM-data fra DNMI. Trond Rinde har vært behjelpelig med tilrettelegging av disse dataene.

Oslo, januar 2001

Kjell Repp
avdelingsdirektør

Sverre Husebye
seksjonssjef

Sammendrag

Abstract

This study is part of a larger project with the aim of coupling atmospheric and hydrological models. The problem at hand is to make forecasted values of precipitation from the atmospheric model (HIRLAM) suitable as input for a distributed hydrological model. The spatial resolution of the atmospheric model is 10 × 10 km and the spatial resolution of the hydrological model is 1 × 1 km. Certain statistical and morphological properties of the input field has to be respected of the disaggregated precipitation field. These are spatial mean, spatial variance, spatial correlation structure and intermittency. The adopted approach is a combination of interpolation and simulation. Each *grid cell* of 10 × 10 km of the atmospheric model is divided into 10 × 10 *pixels* each of length 1 km. The four nodal points of the grid cell are used in a simple interpolation procedure (inverse distance weights) to assign ranks to the pixels according to the interpolated values. The four nodal points of the grid cell are then used for estimating the two parameters of a shifted exponential distribution (or one parameter if intermittency is present) which is assumed an appropriate distribution for precipitation within the pixels. Hundred (10 × 10) values are then simulated with the exponential distribution and sorted in ascending order. The pixels of different rank are assigned the corresponding ordered simulated precipitation value. This procedure is performed for the whole forecasted precipitation field of interest. A theoretical expression for the fraction of non-zero precipitation pixels in a grid cell, based on an assumption of exponentially distributed non-zero precipitation, has been developed as $p = 2(Var(z) / E(z)^2 + 1)^{-1}$, where p is the fraction of the grid cell of positive precipitation and $Var(z)$ and $E(z)$ are the variance and the mean of precipitation within the grid cell, including the zeros. The results are promising with respect to the statistical and morphological properties of the disaggregated field. Some discontinuities can be seen, as the pixels of each grid cell are, to a certain degree, simulated independently of the neighbouring grid cells.

1. Introduction

This study is a contribution to the project "Coupling of meteorological and hydrological models". The purpose of this project is to mutually improve the performances of the meteorological and hydrological models. It is believed that by introducing an improved representation of the land surface from hydrological models to the atmospheric models, meteorological forecasts and climate change scenarios can be improved.

Meteorological and hydrological processes are currently described on quite different spatial scales. Meteorological operational atmospheric models like the High Resolution Limited Area Model (HIRLAM) works on grid sizes of 10 × 10 km and 50 × 50 km, while distributed hydrological models ideally should work on the scale of the hydrological response units where the variation of process dynamics within one type of hydrological response unit is small compared to the dynamics in a different type of unit. Hydrological response units can be described as 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 when local variations are smoothened out, which for the Nordic landscapes has been shown to be in the range of 1-2 km^2 (Beldring et al., 1999). The discrepancy in spatial scales of the meteorological and hydrological data sets thus calls on methodologies for the disaggregation of meteorological data.

Quite an effort has been put forward 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 parameterisation of Entekhabi and Eagleson (1989), which introduces subgrid scale spatial variability of precipitation and soil moisture, runoff generation and hence evaporation were greatly improved. A fairly simple approach of introducing spatial variability to subgrid precipitation was tested by Gao and Sorooshian (1994). The approach is based on two standard assumptions: 1) precipitation processes are homogeneous over the GCM grid (spatial stationarity) and 2) precipitation intensity within the rainfall area can be represented by an exponential distribution. Although these two assumptions where not found to be suitable for GCM grid scale applications, they might be reasonable for smaller scales. The above assumptions have also been adopted for the present study, although some improvements/alterations have been made. Most important is that we show that if assumption 2) holds, the dry fraction of the grid-cell has a theoretical expression. This differs from several studies that, although the non-zero precipitation is exponentially distributed, the dry proportion of the grid cell is a fixed, precipitation-class related empirical value (Dolman and Gregory, 1992, Johnson et al. 1993, Lammering and Dwyer, 2000). The adopted approach in this study is that we would like to have control of the spatial distribution (an exponential distribution) of the disaggregated field and at the same time introduce spatial dependencies (correlation) by interpolation procedures. Let us express the interpolated field as:

$$Y = AX\,,$$

where Y is the interpolated field, an (n×1) vector, A, is a matrix (n×m) of weights derived by the chosen interpolation method and X is input values, the aggregated field, an (m×1) vector.

Let V be a vector of ordered exponential simulated values (simulated with parameters determined from X) sorted into ascending order. Then if $rankY$ is a vector of the ranks of the interpolated input Y, the disaggregated field Z, can be expressed as:

$$Z[i] = V[rankY[i]]\text{, for all } i = 1,..n$$

where the brackets [] indicate the position within the vector and hence also the spatial location. The result is an exponentially distributed field with spatial dependencies inherited from the interpolated field Y.

The next section describes an approach of the description of the spatial distribution through the analysis of fractional coverages of precipitation for different intensities. It is shown that such a description implies an exponential distribution for precipitation. It is further suggested that the exponential distribution is suitable for partial coverages and for minimum intensities higher than zero. Section three describes the disaggregation procedure, while section four presents and discusses the results. Conclusions are found in section five.

2. Methodology

Let us formulate the areal precipitation over a grid cell A, as:

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

where $z(x)$ is the accumulated precipitation at a point x. If we calculate the coverage for each mm τ, by the indicator function $I_\tau(x) = 1_{z(x)\geq\tau}$, then (1) can be approximated by:

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

where $\Delta\tau = 1$ for the time being and a_τ, the fractional area (FA) with precipitation more than τ mm is:

$$a_\tau = \frac{1}{A}\int_A I_\tau(x)dx \qquad 0 \leq a_\tau \leq 1\ . \tag{3}$$

T is a maximum intensity where a_T is approaching zero.

The following property of the 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. The reduction of the FAs as the discharge intensities increase is termed:

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

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 the following expression for the mean areal precipitation can by (4) and (5) be written as:

$$m(z)_A = a_0 \Delta\tau \left(\sum_{\tau=0}^{T} \prod_{j=0}^{\tau} h_j \right) \tag{6}$$

If we let a_0 be equal to 1, i.e. it rains everywhere within the grid cell and *h* is assumed to be a stochastic variable, independent and identically distributed (iid) with mean $\bar{h}$, we can write (6) as:

$$m(z)_A = \Delta\tau \left(\sum_{\tau=0}^{T} \bar{h}^{\tau} \right) \tag{7}$$

If we measure the FAs in fractions or multiples of *mm* (i.e. using a varying threshold interval, $\Delta\tau$), we would find that the ratio *h*, is obviously dependent on the variations of $\Delta\tau$. When $\Delta\tau$ approaches zero, we see from (5) that a_τ approaches $a_{\tau-1}$, and *h* approaches one. Let us introduce a set of thresholds *k* with intervals $\Delta\tau$ in such a way that $\tau = k\Delta\tau,\ k = 0,1,2,..,K$, and K such that $T = \Delta\tau K$. If we in (7) substitute τ for $k\Delta\tau$ and T for $\Delta\tau K$, and adjust the limits of the summation from $k = 0,1,...,K$, we get:

$$m(z)_A = \Delta\tau \left(\sum_{k=0}^{K} \bar{h}^{k\Delta\tau} \right) \tag{8}$$

As *h* belongs to *(0,1)*, the left-hand side of (6) is recognised as a convergent power series which sum is:

$$m(z)_A = \Delta\tau / (1 - \bar{h}^{\Delta\tau}) \tag{9}$$

For limiting conditions ($\Delta\tau \to 0$), (7) can be written as:

$$m(z)_A = \lim_{\Delta\tau \to 0} \left[\Delta\tau / (1 - \bar{h}^{\Delta\tau}) \right] = \int_0^\infty \bar{h}^z dz = -\frac{1}{Log(\bar{h})} \tag{10}$$

The integrand in (10) can be regarded as the complementary cumulative distribution function of z:

$$1 - F(z) = \bar{h}^{z}, \qquad 0 < z < \infty \tag{11}$$

and the cumulative distribution function (CDF) as:

$$F(z) = 1 - \bar{h}^{z}, \qquad 0 < z < \infty \tag{12}$$

The probability density function (PDF) is then:

$$f(z) = -Log(\bar{h})\bar{h}^{z} \tag{13}$$

If we write $-Log(\bar{h}) = \lambda$ and substitute in (13) then we get the exponential distribution:

$$f(z) = \lambda e^{-\lambda z}$$

with moments:

$$E(z) = \frac{1}{\lambda} = \frac{-1}{Log(\bar{h})}$$

$$Var(z) = \frac{1}{\lambda^2} = \frac{1}{Log^2(\bar{h})}$$

If a description of fractional areas with a constant ratio is assumed appropriate, the spatial distribution of precipitation is thus an exponential distribution with parameter $\lambda = -Log(\bar{h})$. The reason for this quite elaborate derivation is that it links up to previous work of Skaugen et al. (1996) and Skaugen (1997), where extreme areal precipitation was simulated by the use of fractional areas. The exponential distribution is also a popular choice for describing the spatial distribution of precipitation for disaggregation purposes, partly because it only requires one parameter, but also because it is found as a quite suitable representation of precipitation (Onof et al. 1998; Schaake et al. 1996; Gao and Sorooshian, 1994; Eagleson , 1978).

At this point, it is convenient to describe the pattern of the fractional areas for precipitation seen from a grid-cell, or catchment 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 where the fractional areas are equal to one. This minimum intensity is denoted *b,* i.e. $a_k = 1_{k\Delta\tau \le b}$. The second outcome is that we have 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, and we will have the PDF:

$$f(z) = \lambda e^{-\lambda(b-z)} \qquad b < z < \infty \tag{14}$$

with moments:

$$E(z) = b + \frac{1}{\lambda} = b + \frac{-1}{Log(\bar{h})} \tag{15}$$

$$Var(z) = \frac{1}{\lambda^2} = \frac{1}{Log^2(\bar{h})} \tag{16}$$

Case of partial coverage of a grid-cell

The point of departure for this case is the apriori knowledge of the unconditional mean and variance (moments including zeros) which we derive from the nodal points of the grid cells. From this we want to estimate the conditional mean and variance, i.e for the positive precipitation values, and we want an estimate of the fraction of the grid-cell which is dry.

Let z and z' denote precipitation including and not including zeros respectively. Then we have the moments:

$$E(z) = \frac{n-m}{n}0 + \frac{m}{n}E(z') = pE(z'), \tag{17}$$

where *p* is the fraction of the grid-cell of positive precipitation and similarly:

$$E(z^2) = \frac{n-m}{n}0 + \frac{m}{n}E(z'^2) = pE(z'^2) \tag{18}$$

and the variance,

$$Var(z) = E(z^2) - E(z)^2 \tag{19}$$

We can substitute (18) into (19):

$$Var(z) = pE(z'^2) - E(z)^2 \tag{20}$$

and if we assume that the distribution of z' is 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 (17), we get:

$$Var(z) = \frac{2}{pE(z)^2} - E(z)^2 \tag{21}$$

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

$$p = \frac{2}{\dfrac{Var(z)}{E(z)^2} + 1} \tag{22}$$

It is appropriate here to discuss the relation between the location parameter b, (in 14 and 15) and p from (22). From (15) we see that b is zero when the spatial standard deviation is equal to the spatial mean. From (22) this corresponds to $p = 1$, i.e complete coverage. For negative values of b (i.e. the spatial standard deviation is higher than the spatial mean) we get, from (22) p-values less than one, i.e. only a fraction of the grid cell is covered by precipitation. Finally, when b is positive (i.e. the spatial standard deviation is less than the spatial mean), we have complete coverage with a minimum intensity equal to b, and p takes on values higher than one.

3. A disaggregation scheme

The theoretical platform presented in section 2 provides us with the necessary tools to establish a disaggregation scheme for precipitation. In the following we will present how a precipitation forecast from the HIRLAM model (10×10 km) is disaggregated into 1×1 km.

The precipitation field from the HIRLAM model consists of *grid cells* (10×10 km), while the resulting disaggregated field consists of *pixels* (1×1 km). The following procedure is repeated for every grid cell.

1) The four corner values of the grid cell (nodal values) are used to estimate by a simple interpolation technique the order of the pixels. For simplicity, the inverse distance (power 2) method has been used in this study. For the hundred pixels in the grid cell, each pixel is assigned a value interpolated from the nodal values (the four corner values). The pixels are, from the interpolated values, assigned a rank (1 to 100) in order to determine which pixel is higher than the other.
2) From the nodal values of the grid cell the spatial mean, the spatial variance and of the grid cell, are estimated.
3) From (15), a value of b is calculated. Dependent on the whether b is less, equal to or higher than zero, it is decided if the precipitation field is intermittent with fractional coverage p determined from (22) and positive precipitation exponentially distributed, $f(z;\lambda)$, or the grid cell is fully covered with minimum intensity b, and exponentially distributed precipitation, $f(z;\lambda,b)$.
4) In case of intermittency, $p100$ values are simulated from $f(z;\lambda)$ and in the case of full coverage, hundred values are simulated from $f(z;\lambda,b)$.
5) The simulated values are then ordered, and the ranked pixels from 1) are assigned the simulated value of equal rank. In case of intermittency, the $(1-p)100$ lowest ranked pixels are assigned the value zero.

4. Results and discussion

The simulation procedure has been carried out for three HIRLAM events for the Gaula catchment (3248 km^2) in central Norway, somewhat south of Trondhjem. The events are picked because they represent a variety of situations that have to be adequately represented by the disaggregation procedure. The 24 hour accumulated event of 26th of September was intense and the mean precipitation of one of the grid cells (10×10 km) of the

HIRLAM field was estimated to be 67 mm. The 12 hour accumulated event of the 26th of September was characterised by intermittency and cell structure, while the second 12 hour accumulated event of the 26th of September obviously describes a passing (or an approaching) front. The figures 1 (a, b), 2(a, b) and 3 (a, b) show the HIRLAM– and the disaggregated fields of the three events. We want the disaggregation scheme to respect the following statistical properties of the precipitation fields, spatial mean, spatial variance, intermittency and spatial correlation structure. Table 1 provides the statistics of the HIRLAM- disaggregated fields. The statistical parameters are close for those generated by HIRLAM and the disaggregation scheme, although not perfect. We can note that the theoretical (by (22)) values of intermittency differ from the 'observed' for the event of 99092612 by approximately 50 %. This is interesting in terms of whether one should disaggregate the whole field using global estimates of the spatial mean and variance, or disaggregate grid cell by grid cell. In the first case, one would certainly obtain a smoother field without the discontinuities on the grid cell borders, but one would also, apparently, have errors in the estimated intermittency. We can further observe that the maximum values of the disaggregated fields are of hydrological importance especially for flood forecasting purposes. With a distributed hydrological model the timing and the magnitude of flash floods can, with the disaggregated precipitation field, be better forecasted.

Tab.1 Statistical parameters of precipitation from HIRLAM and disaggregation scheme

Date	990929		99092600		99092612	
Datasource	HIRLAM	Disagg.	HIRLAM	Disagg.	HIRLAM	Disagg.
mean	22.35	24.95	1.39	1.70	1.62	1.51
variance	169.25	209.95	8.49	12.37	3.87	3.29
observed fraction of zeros	0	0	0.68	0.56	0.44	0.34
estimated fraction of zeros by (22)	0	0	0.63	0.62	0.19	0.18
max	67	179.87	13	36.69	6	15.11
min	3	3.18	0	0	0	0

We see from figures 4 (a, b), 5 (a,b) and 6 (a, b), which show the spatial correlation structure of the HIRLAM- and the disaggregated fields, that the general correlation structure from the HIRLAM model is well reproduced by that of the disaggregated field. It is interesting to note that the correlation structure is quite different from event to event, which should be of consequence to the assessment of interpolation methods based on the spatial correlation structure (like various forms of Kriging). This is consistent with the findings of Skaugen (1997b), which used different correlation structures for kriging interpolation based upon a classification of the precipitation process into small scale (showers) and large scale (frontal precipitation) events.

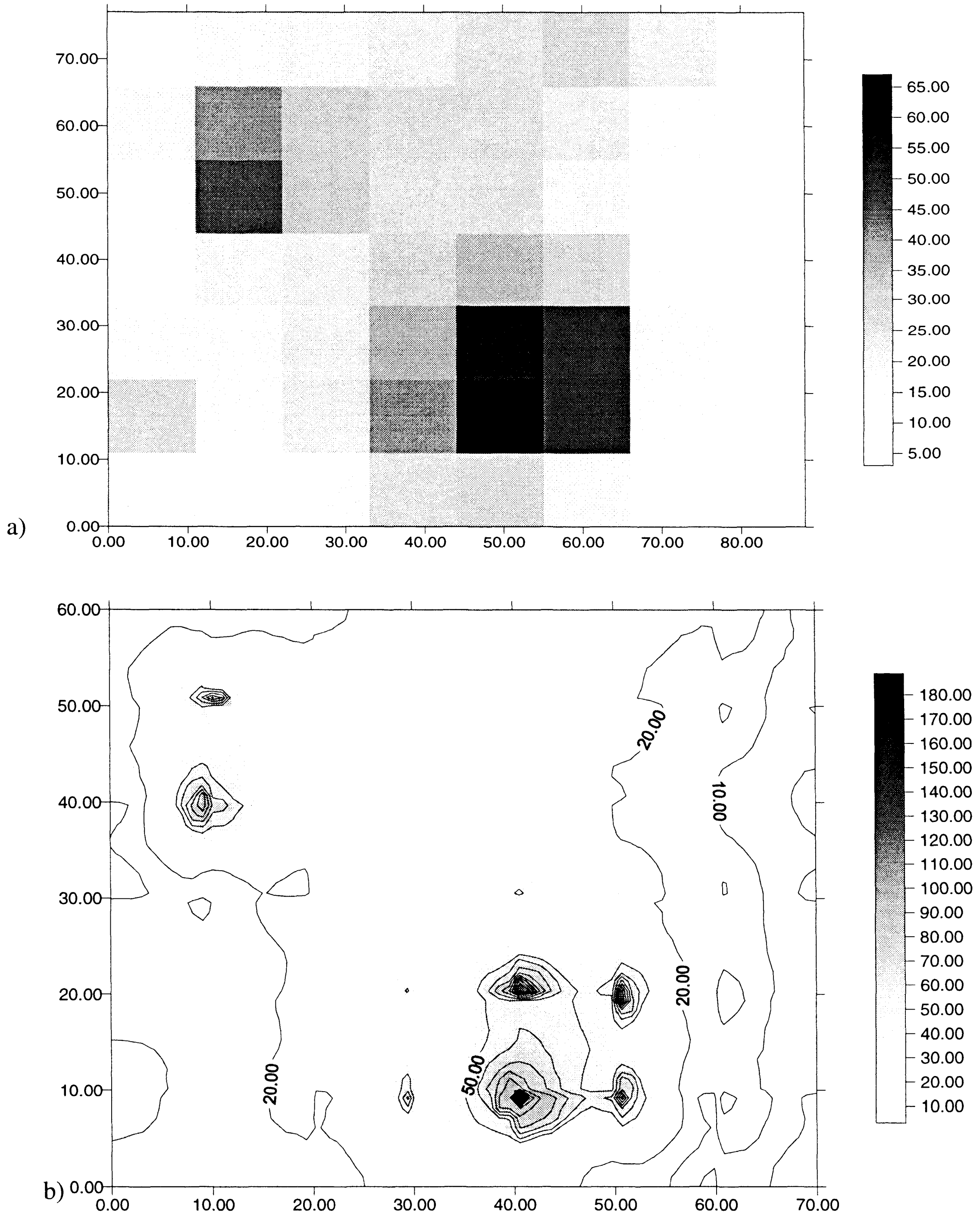
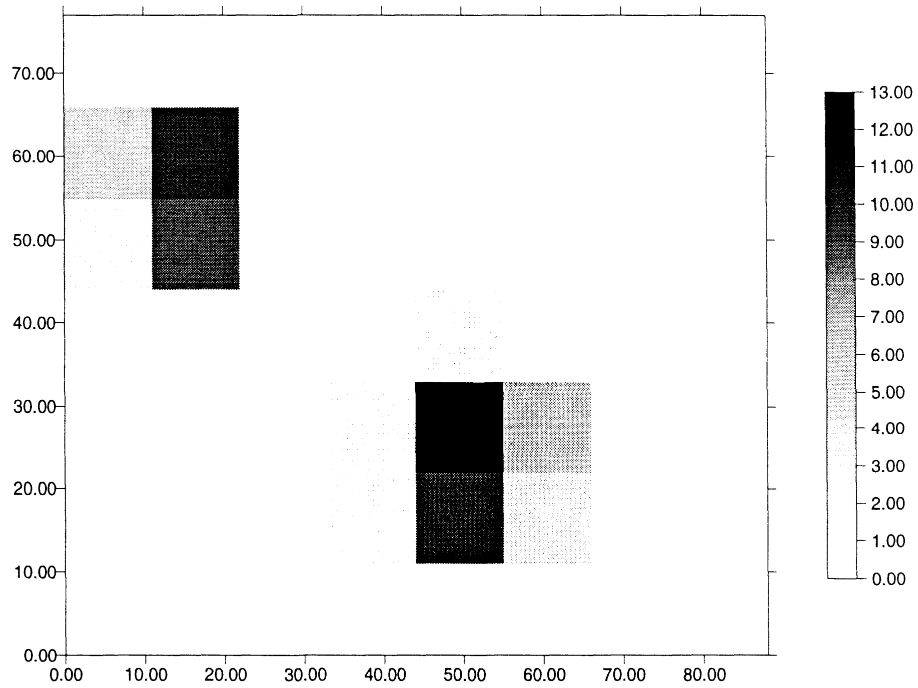


Fig. 1 Hirlam field a) and disaggregated field b) of the 29th of september 1999.

a)



b)

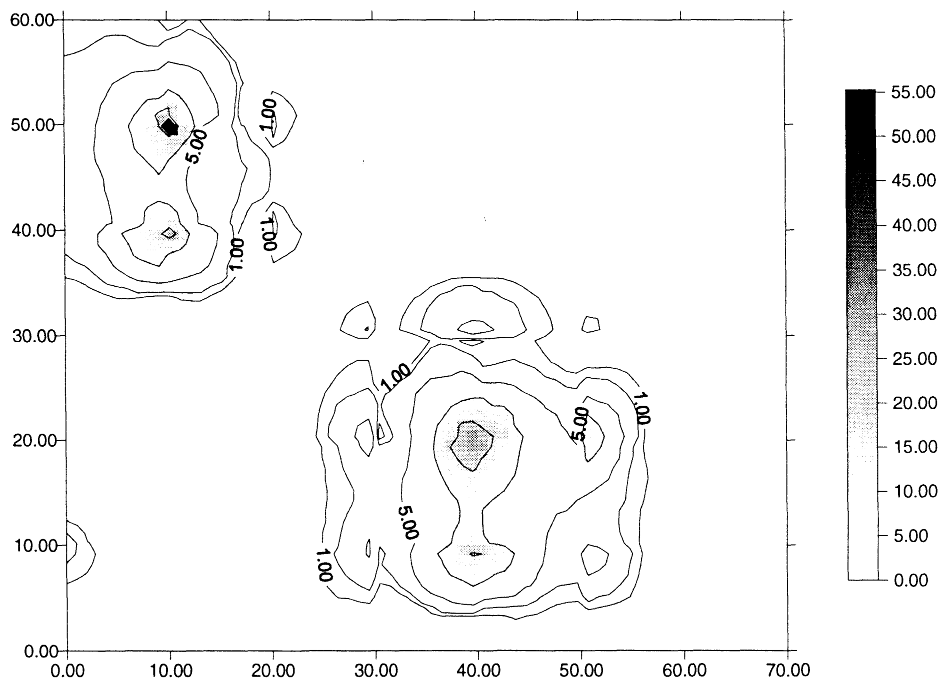
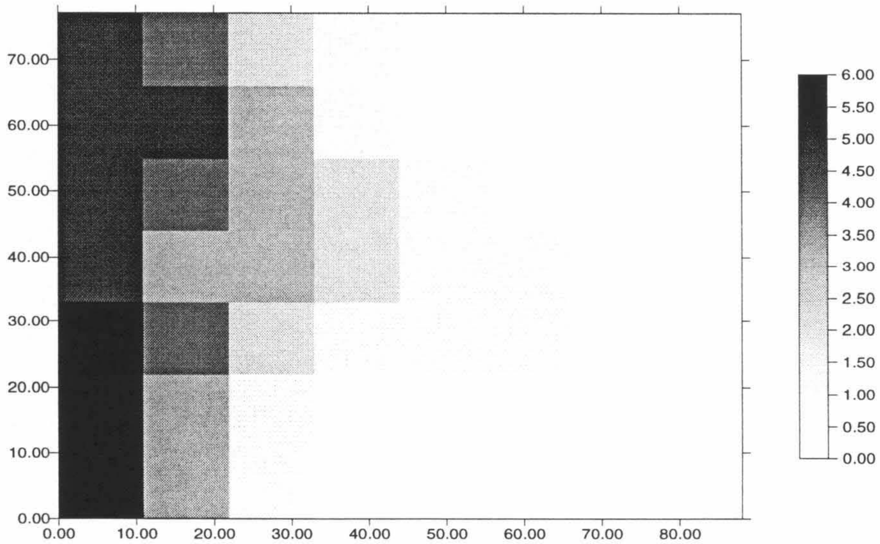


Fig. 2 Hirlam field a) and disaggregated field b) of the 26th of September at 00 hours, 1999.

a)



b)

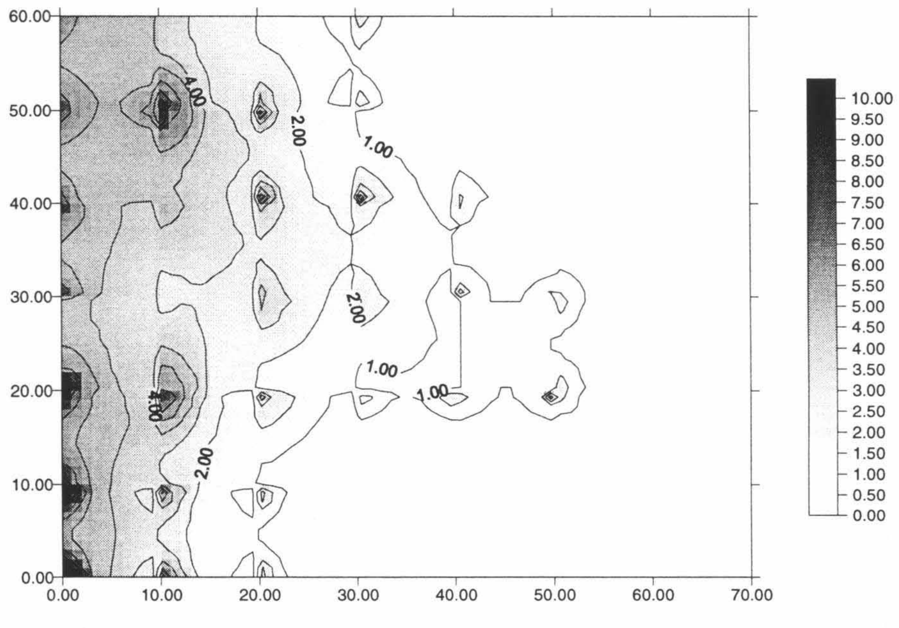
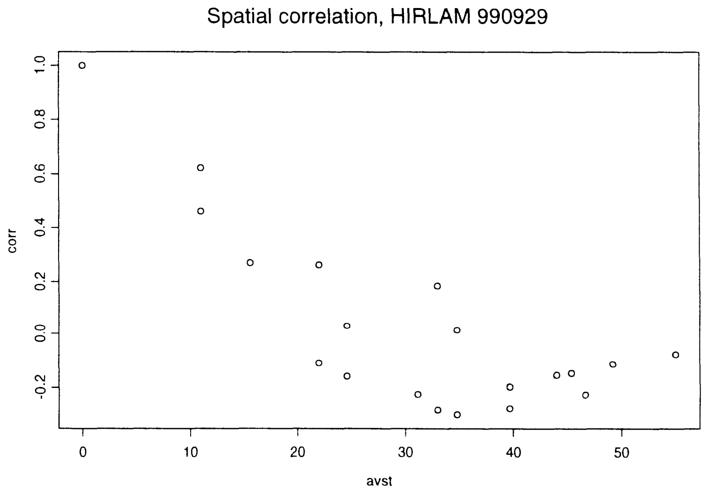
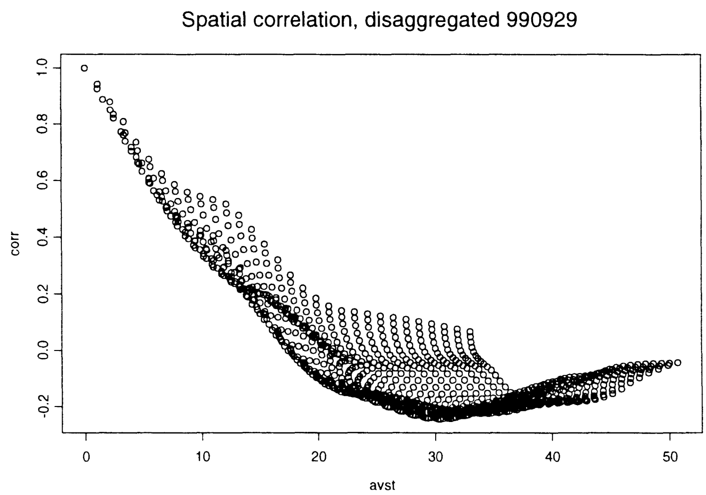


Fig. 3 Hirlam field a) and disaggregated field b) of the 26th of September at 1200 hours 1999.

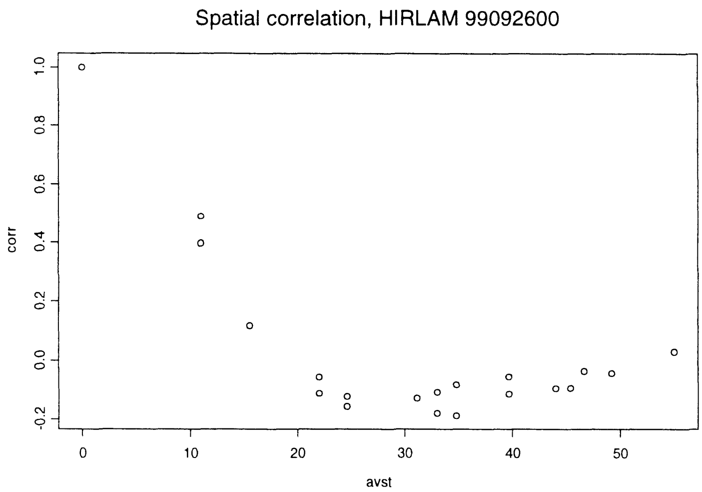


a)

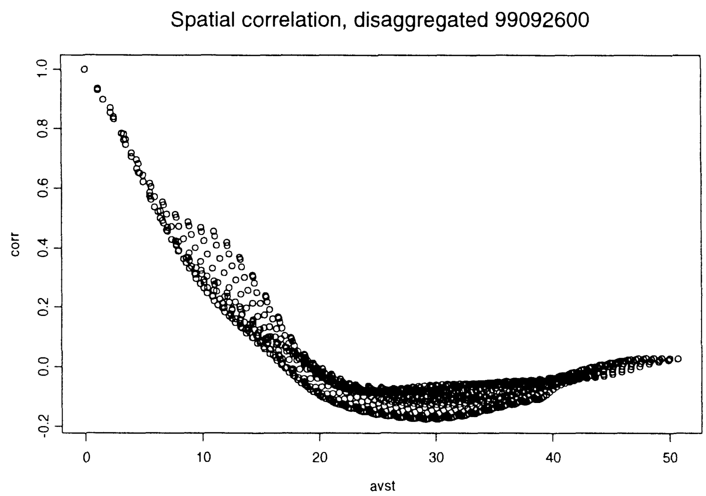


b)

Fig. 4 Spatial correlation structure of the Hirlam field a) and disaggregated field b) of the 29th of September 1999.

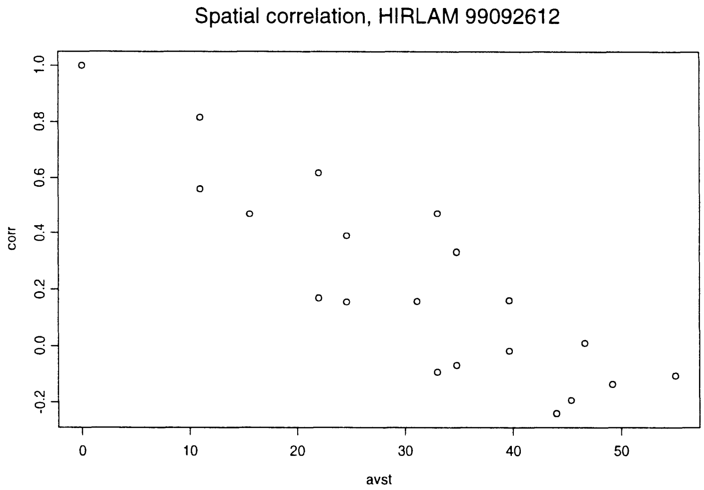


a)

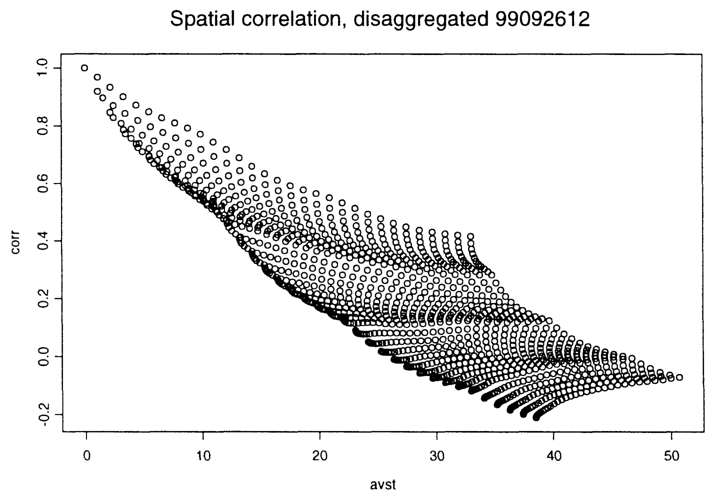


b)

Fig. 5 Spatial correlation structure of Hirlam field a) and disaggregated field b) of the 26th of September at 00 hours, 1999.



a)



b)

Fig. 6 Spatial correlation structure of Hirlam field a) and disaggregated field b) of the 26th of September at 1200 hours 1999.

5. Conclusions and future research

A disaggregation scheme is put forward which adopts the spatial dependencies from an interpolation method and gives exponential distributed precipitation values for each grid cell. The exponential distribution can be used both for events where intermittency in grid cells occurs, and for events where the minimum intensity in the grid cell is higher than zero

The disaggregated fields are promising with respect to the comparisons to the aggregated fields. The spatial mean, spatial variance and spatial correlation structure are comparable to those of the aggregated field, while the fraction of zero rainfall is underestimated.

A theoretical expression is derived for the fraction of dry area, given that the non-zero precipitation is exponentially distributed.

The future research of this study will have to deal with inaccuracies of the statistical parameters of the disaggregated field compared to that of the aggregated field, and the discontinuities on the border of the grid cells have to be investigated.

6. References

Beldring, S. L. T. Gottschalk, J. Seibert and L.M. Tallaksen, 1999 . 'Distribution of soil moisture and groundwater levels at patch and catchment scales'. *Agricultural and Forest Meteorology*, **98-99**, 305-324.

Dolman, A. and D. Gregory, 1992. ' the parameterization of rainfall interception in GCMs', *Q. J. R. Meteorol. Soc.*, **118**, 455-467.

Eagleson, P. S., 1978. ' Climate, soil and vegetation, " The distribution of annual precipitation derived from observed storm sequences'. *Water*

Resour. Res. **14**(5), 713-721.

Entekhabi, D. and P. Eagleson, 1989. 'Land surface hydrology arameterization for atmospheric general circulation models including subgrid scale spatial variability', *J. Clim.*, **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: Kalma, J.D. and M. Sivapalan (Eds.), *Scale issues in hydrological modelling*, Wiley, Chichester, 181-194.

Gao, X, and S. Sorooshian, 1994. ' A stochastic precipitation disaggregation scheme for GCM applications', *J. Clim.*, **7**(2), 238-247.

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

Lammering, B. and I. Dwyer, 2000. 'Improvement of water balance in land urface schemes by random cascade disaggregation of rainfall'. *Int. J. Climatol.* **20**: 681-695.

Onof, C., N.G. Mackay, L. Oh and H. S. Wheater. 1998. ' An improved rainfall disaggregation technique for GCMs', *Jour. of Geophys. Res.*, **103**, D16, 19 577-19 586.

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

Skaugen, T., 1997a. ' A spatial approach to the modelling and estimation of real precipitation' Phd. thesis, University of Oslo, Norway.

Skaugen, T., J-D. Creutin and L. T. Gottschalk, 1996. ' Reconstruction and frequency estimates of extreme daily precipitation'. *Jour. of Geophys. Res.*, **101**, D21, 26 287 – 26 295.

Skaugen, T., 1997b. 'Classification of rainfall into small- and large-scale events by statistical pattern recognition'. *J. Hydrol.* **200**, 40-57.

Denne serien utgis av Norges vassdrags- og energidirektorat (NVE)

Utgitt i Oppdragsrapportserien i 2001

Nr. 1 Liss M. Andreassen og Hans-Christian Udnæs: Snøfordeling gjennom en smeltesesong (20 s.)

Nr. 2 Thomas Skaugen: Disaggregation of precipitation (25 s.)

Norges
vassdrags- og
energidirektorat

N V E

Norges vassdrags- og energidirektorat
Middelthunsgate 29
Postboks 5091 Majorstua
0301 Oslo

Telefon: 22 95 95 95
Telefaks: 22 95 90 00
Internett: www.nve.no