TECHNICAL MEMORANDUM

24

Evaluation of the ECMWF operational model precipitation forecasts in October-November 1980

O. Åkesson

Research Department

February 1981

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

Evaluation of the ECMWF Operational Model precipitation forecasts in October - November 1980

by O. Åkesson

1. Introduction

Precipitation forecast by the ECMWF operational gridpoint model during the months of October and November 1980 has been verified over five selected areas in Europe, located in Finland, Denmark, France, Spain and Yugoslavia, and also at two stations, Copenhagen (06180) and Belgrade (13272). For area verification, the total forecast precipitation has been averaged over an array of 2 x 3 grid-points located in these areas with grid length 1.875^{o}, see figure 1.

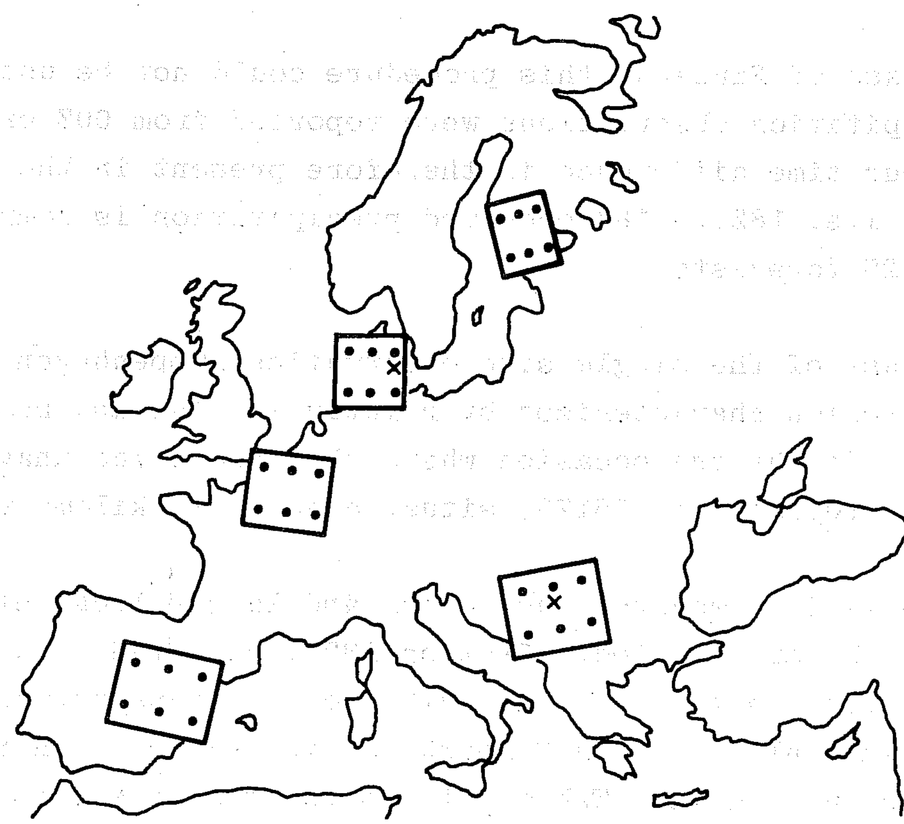


Fig. 1 Areas (rectangles) and grid points (dots) used for calculation of mean amounts of observed and forecast precipitation. The two crosses indicate the locations of the stations used: Copenhagen and Belgrade.

The observed precipitation has also been averaged, from all available observations within the corresponding areas covered by the grid-points plus half a grid distance added to each side. For station verification, the model precipitation forecast at the four nearest grid-points was linearly interpolated to the stations.

2. Observations

To obtain a 24 hour amount of precipitation accumulated between 12Z and 12Z is normally not a problem since most observation stations in Europe report at 6 hourly intervals, reporting 12 hour accumulated precipitation at 06Z and 18Z and 6 hour accumulated precipitation at 12Z and 00Z. The data were monitored manually as well as automatically in order to exclude a few incorrect reports. However, difficulties arose on occasion since some stations did not report regularly. This may have given rise to small errors normally not exceeding 1mm.

In the case of Finland, this procedure could not be used since no precipitation observations were reported from 00Z or 12Z. A six hour time difference is therefore present in the data for Finland, i.e. 18Z to 18Z observed precipitation is compared with 12Z to 12Z forecasts.

In the case of the single site observations, Copenhagen (06180) is in a region characterized by relatively flat and homogeneous terrain. On the one occasion where the report was unavailable, 06180 was replaced by 06179, situated about 75 kilometres distant.

Belgrade is in a mountainous region and larger local effects can therefore be anticipated. Station 13272 reported uninterruptedly during these two months. The number of stations normally reporting were 13 in Finland, 18 in Denmark, 24 in France, 7 in Spain and 30 in Yugoslavia. Table 1 lists the stations from which data were used to obtain area averaged precipitation. Studies of individual precipitation forecasts over Europe have been reported on previously in a Technical Note, "Case Studies of Precipitation Forecasts during the period February - July 1980 by the ECMWF Operational Model", by O. Åkesson, Operations Department File 30.2, 25 September 1980.

WMO Block	02	06	07	08	13
	986	089	140	348	493
	974	059	168	221	477
	976	052	143	280	462
	973	147	181	284	455
	972	119	169	238	452
	963	179	180	233	481
	966	151	149	160	473
	965	110	172		465
	944	152	147		389
	952	120	179		397
	942	180	150		382
	945	104	038		356
	935	060	157		388
		070	090		363
		071	057		352
		087	070		367
		030	037		376
		021	055		353
			061		384
			005		289
			017		295
			015		269
			002		279
			010		257
					262
					272
					183
					150
					168
					174

Table 1. List of SYNOP stations normally received from WMO BLOCKS 02 (Finland), 06 (Denmark), 07 (France), 08 (Spain), 13 (Yugoslavia) which were used to obtain area averaged precipitation in this study.

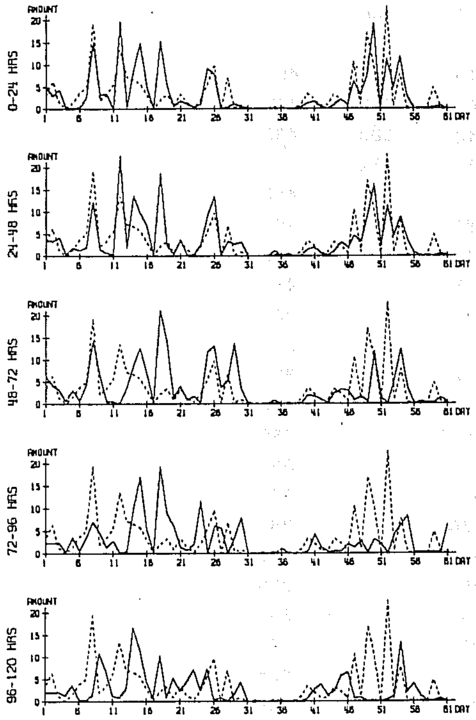
3. Results

3.a. Meteograms

Figures 2a to 2e show precipitation meteograms of observed and forecast 24-hour averaged accumulated precipitation for the different forecast days up to forecast day 5 for the five areas and Figures 3a and 3b show the same information for the two stations. It can be seen that distinguishing periods of precipitation from dry periods is very well handled by the model for the first two (or three) forecast days and even, with some success, to forecast day 5 (Compare Table 2). This was also observed by Bosart (1980) for the NMC LFM-2 model up to 36 hours.

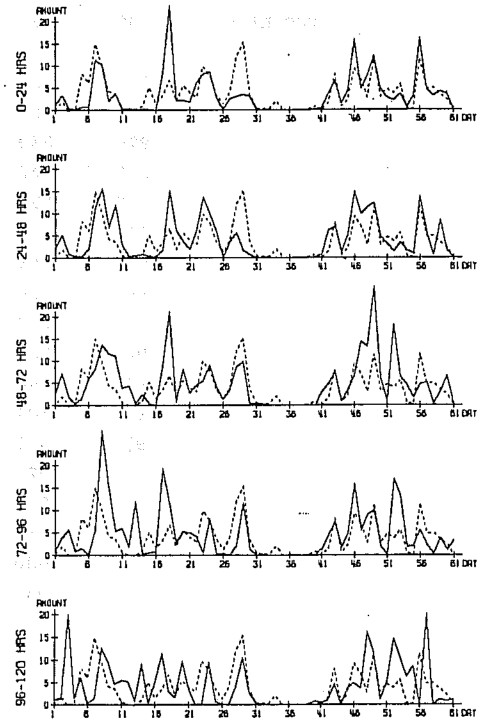
But given that precipitation is observed, how well does the model perform quantitatively?

a.



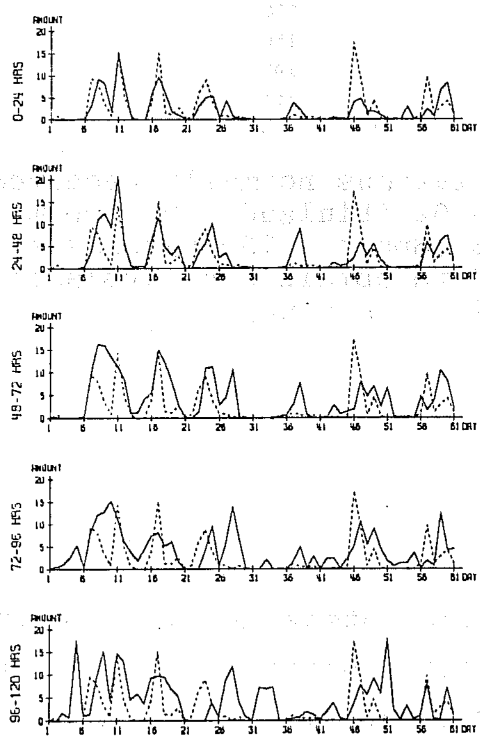
AREA AVERAGED PRECITATION FROM 1.10.80 TO 30.11.80 OVER FINLAND
SOLID LINE, FORECAST; DASHED LINE, OBSERVED

b.



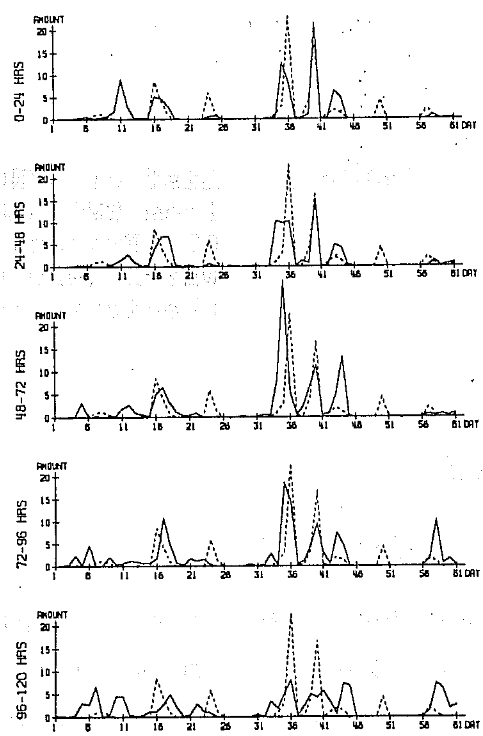
AREA AVERAGED PRECITATION FROM 1.10.80 TO 30.11.80 OVER DENMARK
SOLID LINE, FORECAST; ASHED LINE, OBSERVED

c.



AREA AVERAGED PRECITATION FROM 1.10.80 TO 30.11.80 OVER FRANCE
SOLID LINE, FOHECAST; DASHED LINE, OBSERVED

d.

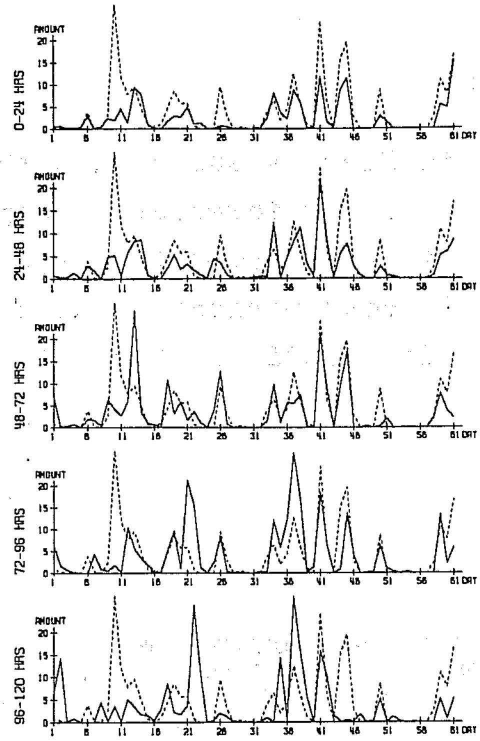


AREA AVERAGED PRECITATION FROM 1.10.80 TO 30.11.80 OVER SPAIN
SOLID LINE, FORECAST; DASHED LINE, OBSERVED

Figure 2. Area averaged precipitation forecast and observed for individual days from 1 October to 30 November 1980,

a. Finland, b. Denmark, c. France, d. Spain.

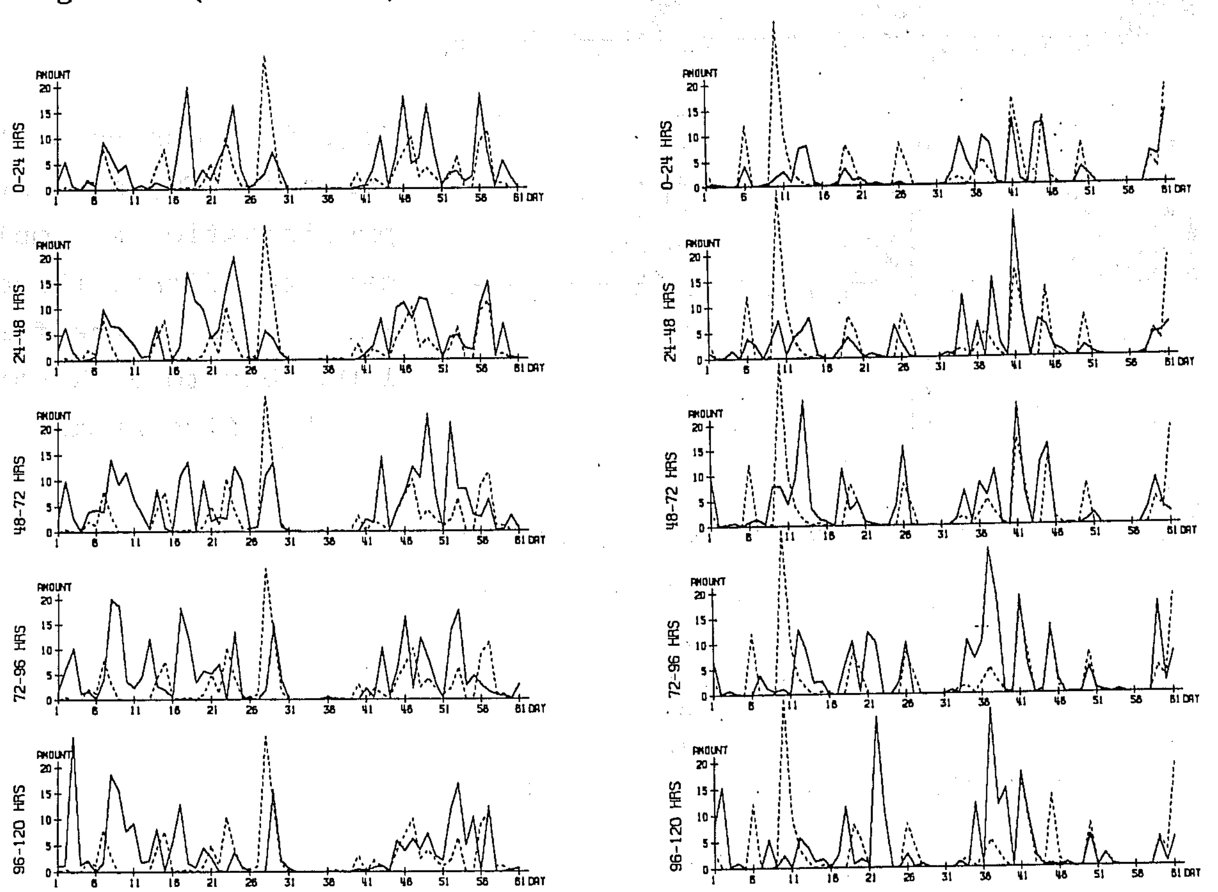
The forecasts are for precipitation accumulated during forecast day 1 (top) to day 5 (bottom) for each area.



AREA AVERAGED PRECITATION FROM 1.10.80 TO 30.11.80 OVER JUGOSLAVIA
SOLID LINE, FORECAST; DASHED LINE, OBSERVED

Figure 2 (continued) e. Yugoslavia

a. b.



POINT OBS.AND FC.PRECIPITATION 1.10.80 TO 30.11.80 FOR COPENHAGEN
SOLID LINE, FORECAST; DASHED LINE, OBSERVED

POINT OBS.AND FC.PRECIPITATION 1.10.80 TO 30.11.80 FOR BELGRADE
SOLID LINE, FORECAST; DASHED LINE, OBSERVED

Figure 3. As Fig. 2, but single station instead of area-averaged precipitation for a. Copenhagen and b. Belgrade.

3.b. Scatter Diagrams

The scatter diagrams in Figs. 4 and 5 show the area-averaged observed versus area-averaged forecast precipitation, and point observed versus point forecast amounts respectively for the 0-24, 48-72, and 96-120 hour periods. The number on each diagram is the linear correlation coefficient of all events.

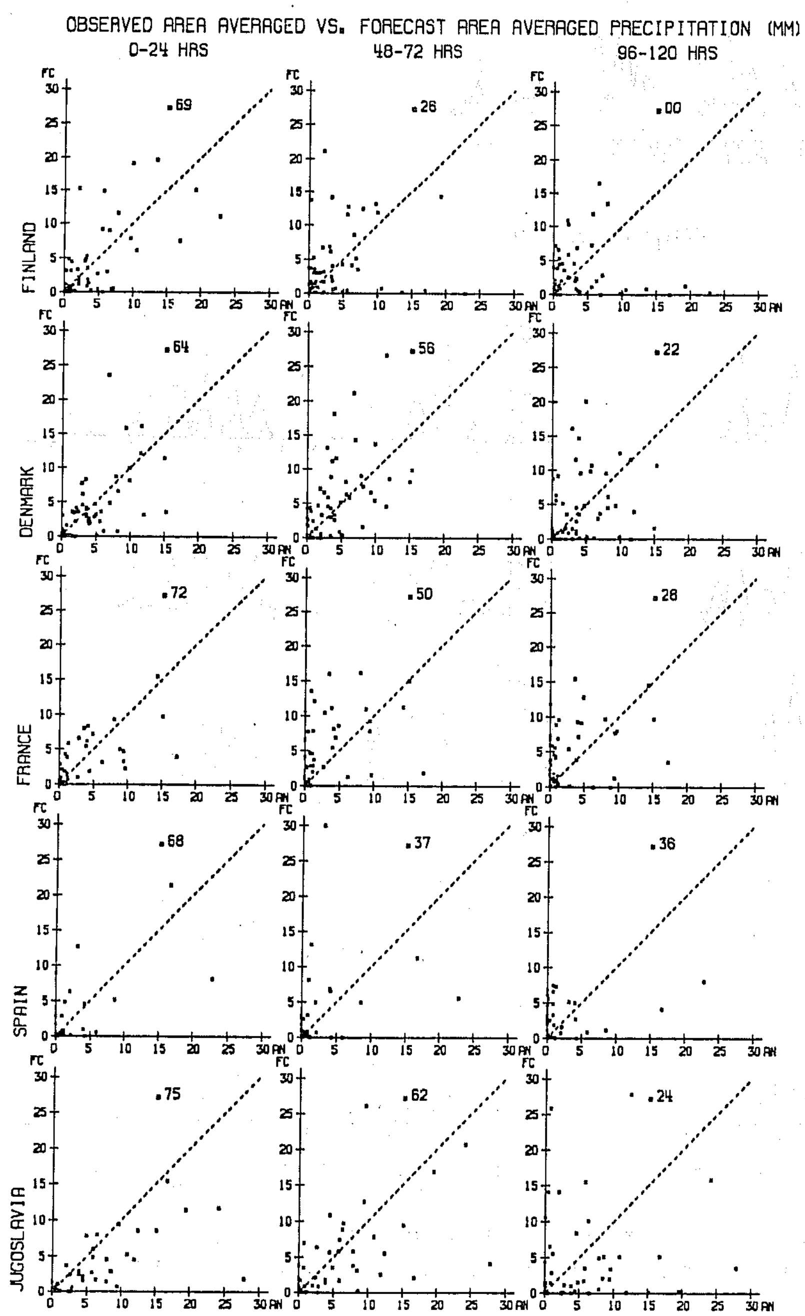


Fig. 4. Observed (AN) vs. forecast (FC) area averaged precipitation accumulated over the first, third and fifth days of the forecasts; 1 October to 30 November 1980 for the five areas.

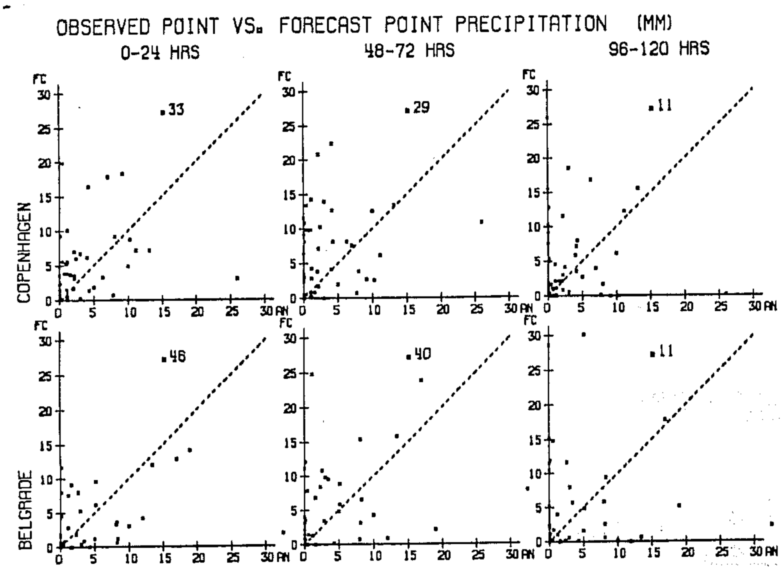


Fig. 5. As Fig. 4 but for the two stations.

Note the shift to higher forecast amounts in all diagrams of the cluster of points from day 1 to day 3, easily seen for France and Yugoslavia. Note also the relatively more scattered picture in the 0-24 hour diagrams for the two stations, Fig. 5, with significantly lower correlation than for the areas, Fig. 4.

We consider correlation coefficients further in Section 3d. Note that for observed large precipitation amounts, there is an under-prediction in the forecast.

The cluster of points near the origin in the diagrams of Figs. 4 and 5 reflects the forecast's ability to distinguish dry from wet days. This is summarized in Table 2 in contingency tables where only two categories are considered: $\leq$ 0.2mm (dry) and > 0.2mm (wet). Note the significant reduction of dry events for both forecast and observation from day 1 to day 3 in the forecast and the small reduction from day 3 to day 5. Note also that, on day 3, the frequency of events of wet forecast and dry observation is higher than the frequency of events of dry forecast and wet observation throughout. This is related to the spin-up of forecast precipitation intensity.

Skill scores from contingency tables with four classes are considered in section 3.e.

a. OBS (horizontal), FC (vertical)

	0-24		48-72		96-120 HRS	
FINLAND	15	4	8	6	10	8
	0	42	7	40	5	38
DENMARK	8	10	6	5	5	10
	5	38	7	43	8	38
FRANCE	21	6	16	4	8	5
	5	29	10	31	18	30
SPAIN	34	6	25	6	20	4
	5	16	14	16	19	18
YUGOSLAVIA	20	6	16	5	16	9
	4	31	8	32	8	28

b. AN (horizontal), FC (vertical)

	0-24		48-72		96-120 HRS	
COPENHAGEN	14	1	12	2	10	10
	12	34	14	33	16	25
BELGRADE	27	3	21	3	19	6
	8	23	14	23	16	20

Table 2.

Contingency table with two categories: 0.0 - 0.2mm and more than 0.2mm. Events of observed values are found horizontally and forecast events vertically.

a: Area averaged observed and forecast precipitation for day 1, 3 and 5 respectively.

b: Point observed and forecast precipitation for day 1, 3 and 5 respectively.

3.c. Forecast errors

In order to obtain a measure of forecast accuracy as a function of observed precipitation amount, the observed data have been divided into four classes: 0-0.9 mm, 1.0 - 3.9 mm, 4.0 - 9.9 mm and more than 9.9 mm per 24 hours and the standard error of the forecast computed for each class. This is depicted in Fig. 6 and 7 for the 0-24 hour, 48-72 hour and 96-120 hour time intervals respectively. The growth of error (standard error of estimate) with time appears to occur mainly in the lowest classes and remains more or less constant in the higher classes but note the limited sample of high precipitation cases. In Spain, for example, only two events of more than 10 mm were recorded during the two month period.

In the largest category, Figures 6 and 7 show that cases of heavy observed precipitation are underestimated by the forecast. This can also be seen in Figs. 2a - 2e and Figs. 3a and 3b, as well as Figs. 4 and 5.

On the other hand, for small or no observed amounts, there is a marked overprediction on day 3 and 5 (See also Fig. 4) except for Yugoslavia. This is in agreement with what has earlier been found by Åkesson (1980) from case studies where a tendency to forecast widespread small amounts of precipitation developed with time. For winter cases, there was also a tendency to produce somewhat too small local maxima of precipitation over land. The contrary has on occasion been observed over the Atlantic, where apparently excessive amounts of precipitation have been locally forecast, although verifying data are not generally available here.

Figure 7 shows the forecast errors for single stations are larger, especially in the two lowest classes, than for area averaged precipitation amounts and, on day 5, the errors are almost the same throughout the classes. For the 0-24 hour diagrams in Fig. 7, the errors for the stations are of the same magnitude as the amount, while the errors are somewhat lower for the areas shown in Fig. 6.

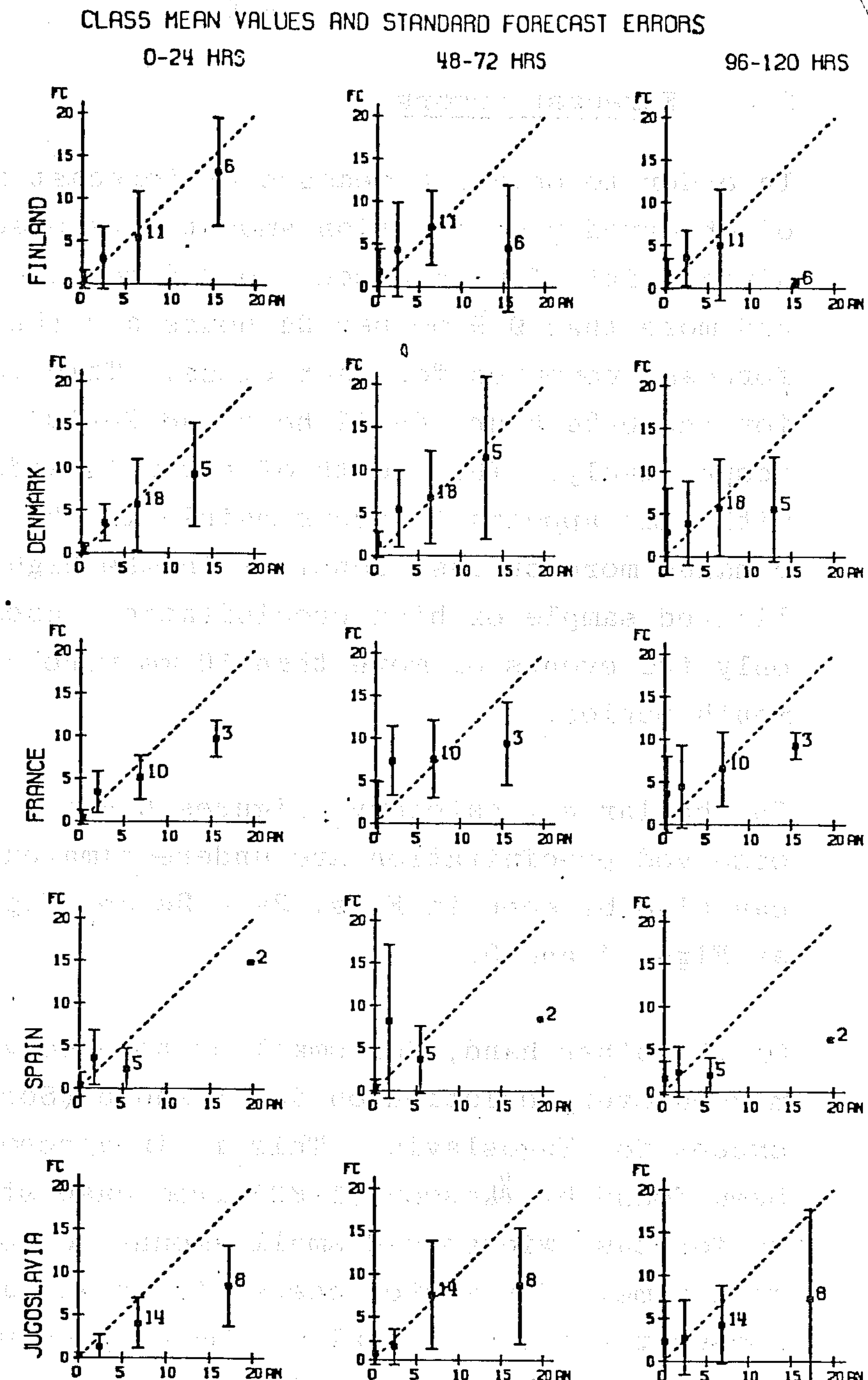


Figure 6.

Forecast precipitation as a function of observed for the 4 classes of observed precipitation, 0-0.9mm, 1.0-3.9mm, 4.0-9.9mm and more than 9.9mm per 24 hours for the five areas. The numbers on the figures show the sample sizes for the two higher classes. The standard error of the forecast is also shown (vertical bars).

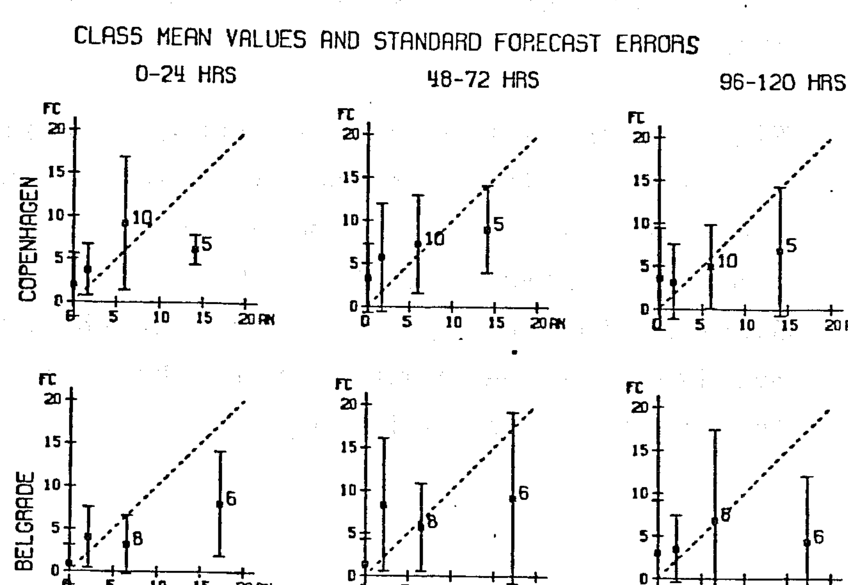


Figure 7.

As Fig. 6 but for the two stations

3.d. Correlation coefficients

The quality of the precipitation forecasts in terms of correlation coefficients (Table 3) shows a deterioration in time in general. The exceptional behaviour of Spain is associated with the small number of events of significant precipitation.

a.

	0-24	24-48	48-72	72-96	96-120
Finland	0.69	0.67	0.26	0.04	0.00
Denmark	0.64	0.64	0.56	0.42	0.22
France	0.72	0.65	0.50	0.43	0.28
Spain	0.68	0.71	0.37	0.59	0.36
Yugoslavia	0.75	0.71	0.62	0.49	0.24

b.

0-24	24-48	48-72	72-96	96-120
0.58	0.58	0.10	-0.13	-0.17
0.50	0.51	0.44	0.30	0.06
0.55	0.43	0.29	0.32	0.13
0.55	0.59	0.15	0.50	0.20
0.63	0.57	0.48	0.34	0.05

c.

	0-24	24-48	48-72	72-96	96-120
Denmark - Copenhagen	0.66	0.54	0.53	0.37	0.17
Yugoslavia-Belgrade	0.74	0.64	0.68	0.53	0.21

d.

0-24	24-48	48-72	72-96	96-120
0.28	0.34	0.26	0.12	0.19
0.48	0.49	0.34	0.26	0.14

e.

	0-24	24-48	48-72	72-96	96-120
Copenhagen	0.32	0.37	0.29	0.10	0.11
Belgrade	0.46	0.48	0.40	0.30	0.11

Table 3. Correlation coefficients calculated for:

a. Observed area averaged versus forecast area averaged precipitation.

b. Same as a. but for events where either of forecast or observed or both exceeds or equals 1mm.

c. Observed area averaged versus point forecast precipitation.

d. Point observed versus area averaged forecast precipitation.

e. Point observed versus point forecast precipitation.

Tables 3a to 3e show that the linear correlation coefficients between forecast and observed precipitation vary depending on whether point or area forecast and observations are used:

- area forecast vs. area observed. The linear correlation coefficients, about 0.7 on forecast days D+1 and D+2, show a marked fall to about 0.2 by D+5. If only events with significant precipitation (either forecast or observed or both) are considered, then the correlations are significantly lower, but they show the same trend.

- point forecast vs. area observed. Here, forecasts interpolated to Copenhagen and Belgrade are compared to area-averaged observations from Denmark and Jugoslavia. The correlations here are only slightly lower than area forecast vs. area observed.

- area forecast vs. point observed. The correlations here are notably lower and of the same low magnitude as for point forecast versus point observed.

- point forecast vs. point observed. The correlations are notably lower than for area forecast vs. area observed.

3.e. <u>Skill scores</u>

Table 4 shows the calculated Heidke-skill score based on the contingency tables using $S = \frac{R - E}{T - E}$, where R is the number of correct forecasts, T is the total number of forecasts and E is the number of forecasts expected to be correct based on chance. The highest skill possible gives a score of +1 and a score of zero indicates no skill. The numbers in Table 3 can be seen to be positive throughout. There is a low value for Spain on D+1 because of the few cases of large amounts. The contingency tables of forecast vs. observed precipitation, based on the class intervals used to compute the skill score are significant at better than the 99% level of significance by the chi-square test for area verification for the first three days, but fall to the 95% level or lower in some cases for forecast days 4 and 5.

a.

	0-24	24-28	48-72	72-96	96-120
Finland	0.47	0.45	0.24	0.09	0.06
Denmark	0.38	0.23	0.18	0.19	0.24
France	0.49	0.51	0.41	0.16	0.17
Spain	0.33	0.32	0.34	0.16	0.08
Yugoslavia	0.45	0.40	0.39	0.27	0.22

b.

	0-24	24-48	48-72	72-96	96-120
Copenhagen	0.21	0.19	0.19	0.10	0.16
Belgrade	0.39	0.29	0.35	0.29	0.15

Table 4. Skill scores computed from contingency tables with classes 0-0.9, 1.0-3.9, 4.0-9.9, and more than 9.9 mm.

a. Area averaged scores

b. Station scores

These skill scores can be compared to the scores reported by Andrews (1978) where subjectively modified numerical forecasts from a 6 Layer Primitive Equation Model (6L-PE) up to day 3 and a combined Barotropic-Reed Model for days 4 and 5 and precipitation probability areas based on MOS have been used. The skill scores for days 3 to 5 were reported to be 0.20, 0.16 and 0.12 respectively and the mean score for the day 1 to day 5 forecasts based on MOS was 0.23.

3.f. Bias in mean daily amounts of forecast and observed precipitation

The ratio of mean daily amounts of forecast over observed precipitation has been calculated and this is shown in Fig. 8.

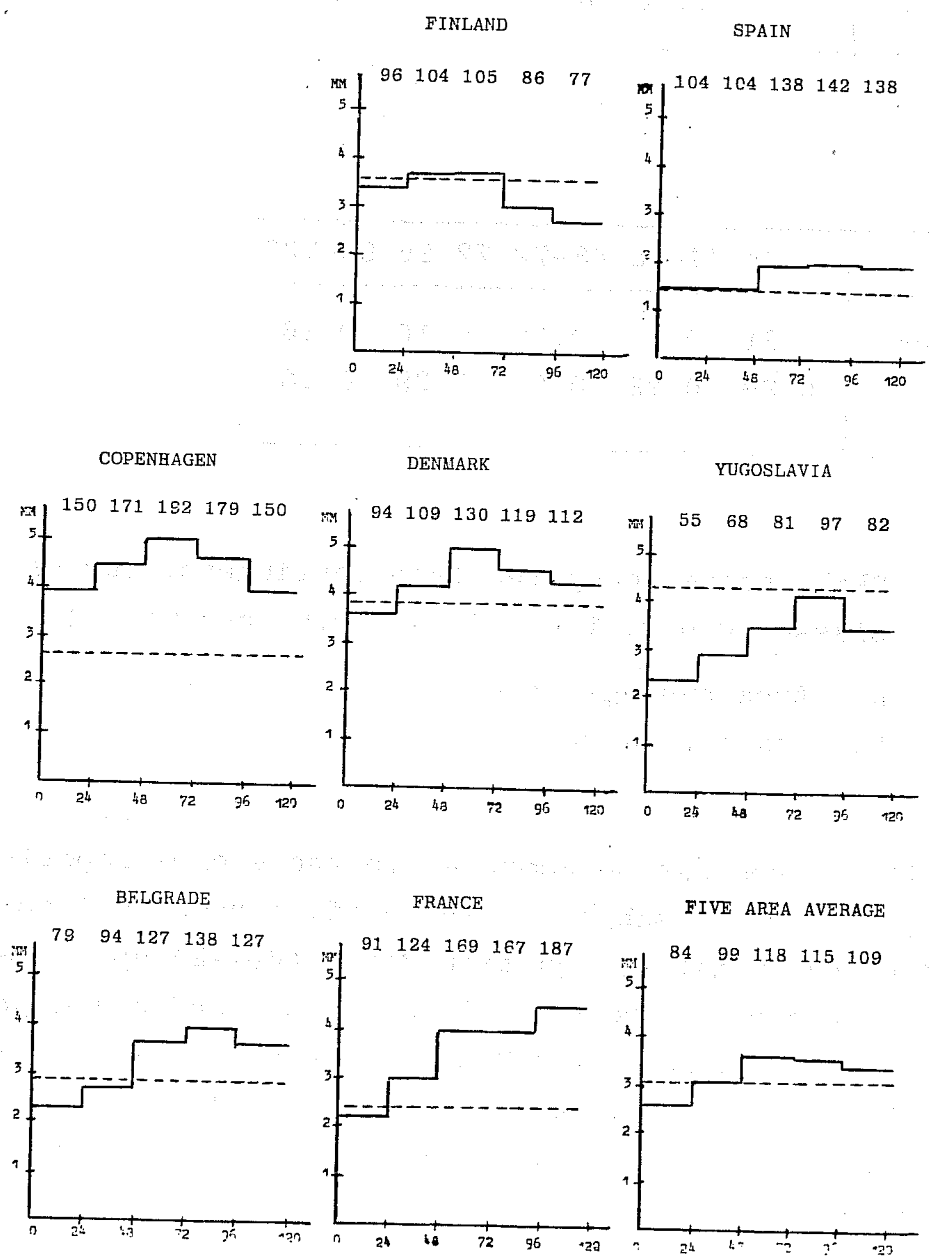


Fig. 8 Mean observed precipitation (dashed line) for Copenhagen and Belgrade (left column), Finland, Denmark and France (middle column), Spain, Yugoslavia and the five area average (right column) and mean daily forecast precipitation (solid line) for the first five days of the forecast. The numbers on the top of the diagrams show forecast precipitation expressed as percentage of the observed.

A certain increase of total forecast precipitation from day 1 to day 3 is obvious in all areas but the increase is weak for Finland and Spain. For Yugoslavia, the forecast peak is reached on forecast day 4 with 97% of forecast over observed precipitation, whereas for France, the area of maximum bias, the forecast maximum is reached on day 5 with 187% of forecast over observed amount in this two-month sample.

For all five areas together, there is an overall increase in forecast precipitation from 84% on day 1 to 118% on day 3 and, from day 3 to day 5, there is a slight decrease in the total amount.

4. Summary

Forecast precipitation for 24 hour periods has been quantitatively assessed for 5 areas (and 2 points) in Europe, during the two month period 1.10.80 - 30.11.80. Meteograms are shown in Figures 2a to 2g. The main features of the investigation can be summarized as follows:

1. Dry and wet periods are very well distinguished by the forecasts on D+1 and D+2, with a slow deterioration to D+5. Given that precipitation is observed, the forecast accuracy decreases with increasing observed amounts and there is a systematic underprediction of heavy precipitation which increases as the forecast proceeds.

2. The standard deviation of forecast error is of the same order of magnitude as the amount itself on day 1, whereas the forecast error grows for amounts lower than 4mm in longer range forecasts, For observed amounts larger than 10mm per 24 hours, the errors appear to be almost constant with time, but the sample sizes were small for heavy precipitation.

3. The correlation coefficients between forecast and observed precipitation are highest when area forecast precipitation is verified against area observed, and noticeably lower when single-station precipitation is verified.

4. Skill scores for area averages computed from contingency tables with four fixed classes 0.0-0.9, 1.0-3.9, 4.0-9.9 and more than 10.0 mm show a steady drop from day 1 to day 5 (typically 0.4 to 0.2). For the first three forecast days, the χ^2 test indicates that the contingency tables are significant at the 99% level.

5. In general for this two-month sample, the forecast bias is small for the areas of Finland, Denmark and Spain. For France, a significant enhancement of the accumulated precipitation amount occurs from day 1, with a forecast to observed ratio of 91%, to day 5, when there is an overprediction of 187%. Over the mountainous area of Yugoslavia, however, predicted amounts are less than the observed for all five time periods. Averaged over the five areas, there is an 84% forecast over observed ratio for day 1, rising to 118% for day 3 and then slowly decreasing until day 5.

6. The bias on day 3 through day 5 is mainly due to small and medium amounts of precipitation being predicted when none or only small amounts are observed, i.e. the frequency of events of significant forecast precipitation is higher than observed. The large observed amounts, however, are in general not forecast to the full extent.

References

ECMWF Forecast Model Documentation Manual, Vol. 1.

Åkesson, O., 1980: Case studies of precipitation forecasts during the period February to July 1980 by the ECMWF Model. Technical Note file: 30.2, Date: 25.9.80.

Andrews, J.F., 1978: Medium range weather forecasting in the National Meteorological Center. Preprints Conf. Weather Modification and Analysis and Aviation Meteorology, Silver Spring, American Met. Soc., 13-19.

Bosart, L.F., 1980: Evaluation of LFM-2 Quantitative Precipitation Forecasts. Monthly Weather Review 108, 1087-1099.

Hughes, L.A., 1978: Precipitation Probability Forecasts - Problems seen via a comprehensive verification. Monthly Weather Review 107, 129-139.

Klein, W., 1978: Proceedings from ECMWF Seminar 1978, 222-226.