StartClim.14

Hochwasser 2002
Prognosegüte meteorologischer Vorhersagemodelle

Zentralanstalt für Meteorologie
und Geodynamik

Projektleitung: Dr. Thomas Haiden
Zentralanstalt für Meteorologie und Geodynamik
Hohe Warte 38
1190 Wien
E-mail: thomas.haiden@zamg.ac.at
Tel.: 01-36026-2322
Fax: 01-36026-73
Web: www.zamg.ac.at

Projektmitarbeiter: Mag. Alexander Kann
Zentralanstalt für Meteorologie und Geodynamik
Hohe Warte 38
1190 Wien
E-mail: alexander.kann@zamg.ac.at
Tel.: 01-36026-2323
Fax: 01-36026-73
Web: www.zamg.ac.at

Wien, November 2003

Teilprojekt von StartClim
“Startprojekt Klimaschutz: Erste Analysen extremer Wetterereignisse
und ihrer Auswirkungen in Österreich“

Projektleitung: Institut für Meteorologie und Physik
der Universität für Bodenkultur Wien
Türkenschanzstr. 18, 1180 Wien
URL: http://www.austroclim.at/startclim/

Table of contents

Kurzfassung / Abstract

Die detaillierte Verifikation meteorologischer Vorhersagemodelle in Bezug auf das Hochwasserereignis 2002 ist nicht zuletzt für die Entwicklung eines effektiven Warnsystems unerlässlich. Im vorliegenden Bericht wird quantitativ gezeigt, dass die Prognosegüte eine starke Abhängigkeit sowohl von der räumlichen und zeitlichen Skala, als auch von den verwendeten Beobachtungsdaten und vom betrachteten Gebiet aufweist. Generell sind Vorhersagen für Regionen, die sich in alpinen Staulagen befinden, verlässlicher als für Flachlandregionen, da hier numerisch schwierig zu erfassende konvektive Prozesse maßgeblichen Anteil an Starkniederschlagsereignissen haben. Es kann eine signifikante Reduktion des relativen Fehlers durch Erhöhung der Vorhersagedauerstufe erreicht werden, da sich Prognosefehler im zeitlichen Verlauf des Ereignisses zu einem gewissen Teil kompensieren. Eine Reduktion des relativen Fehlers durch Übergang auf größere Gebiete wird erst im Bundesländer-Maßstab deutlich. Maßgeblichen Einfluss auf die Prognosequalität hat nicht nur die räumliche Ausdehnung des Vorhersagegebietes, aber auch jene des synoptischen Systems selbst. Verglichen mit anderen Ereignissen der letzten 4 Jahre wurden die Niederschlagsmengen beim ersten Teil des Hochwassers im August 2002 eher schlecht, beim zweiten Teil jedoch relativ gut erfasst, wobei jedoch einzelne stündliche Maxima generell problematisch sind und häufig unterschätzt wurden. Wahrscheinlichkeitsaussagen mittels Ensemble-Prognosen können einen Beitrag in der Früh- oder Vorwarnung leisten, indem Aussagen über die Bandbreite möglicher Niederschlags-Szenarien gemacht werden können.

A detailed verification of meteorological forecast models with regard to the August 2002 flood event is a necessary requirement for the development of effective warn systems. This report shows quantitatively that the forecast skill strongly depends on the temporal and spatial scale, as well as on the observational data used, and the area under consideration. In general, forecasts for alpine areas affected by orographic upslope precipitation are more reliable than those for lowland reagions because in the latter convective processes make a largeer contribution to heavy precipitation events. A significant reduction of the relative forecast error can be achieved by increasing the duration for which a forecast is made. This is because forecast errors partially compensate within the duration of an event. A reduction of the relative forecast error through increasing area size can be achieved only when one approaches the typical scale of a province. It is not just the size of the catchment area but also the size of the synoptic disturbance itself that determines forecast skill. Compared to other events of the last 4 years, precipitation amounts during the first part of the August 2002 flood were forecasted poorly, whereas during the second part forecasts were rather better than average. Hourly maxima are still hardly predicted, and generally underestimated. Probability forecasts based on ensemble predictions can contribute to improved pre-warnings (or ‘watches’) in the sense that they give the possible spectrum of precipitation szenarios.

14-1 Introduction

The severe flooding of August 2002, which affected large parts of Austria, raised public awareness towards heavy precipitation events and caused concern about possible changes in heavy precipitation climate during the next decades (Haiden and Schultheis, 1995). The project StartClim aims at addressing these concerns, providing an assessment of the August 2002 event and its repercussions, as well as a summary of the scientific state-of-the-art on the issue.

The main objective of sub-project StartClim.14 was to analyse in detail the skill of meteorological forecast models during the August 2002 event (Haiden, 2003), and to compare the results with other heavy precipitation cases of the recent past (Haiden et al., 1997). A good summary of the general performance of meteorological models with regard to the August 2002 event has been presented as part of the comprehensive documentation by Habersack und Moser (2003). Here we provide, as a further step, a detailed analysis of the forecast errors of different models, their dependence on parameters like area size and duration, and a comparison with earlier events.

For hydrological purposes the prediction of *areal* precipitation is even more important than point precipitation values. Areal averages (both observed and predicted) are also more robust and more suitable for inter-model comparison. Since 1999, ZAMG is operationally producing areal precipitation forecasts and analyses for 26 areas in the eastern alpine region (Andrade-Leal et al., 2002; Haiden and Stadlbacher, 2002; see also Figure 14-1.1). The dataset is continuously being archived and forms the basis of this study.

In Section 2 the uncertainty of observed areal precipitation is analysed by comparing results obtained from different data sources and station networks (ZAMG, HZB), as well as different interpolation methods. Here we also present a direct comparison of our results with those from sub-projects StartClim.12 and StartClim.13. Section 3 gives detailed validation results of the August 2002 flood in the form of time series of observed and predicted rainfall amounts, and error distributions. Section 4 provides the same kind of information for other heavy precipitation events out of the period 1999-2002. Based on the analysis of model errors, the problem of flood warnings is discussed from a meteorological point of view in Section 5.

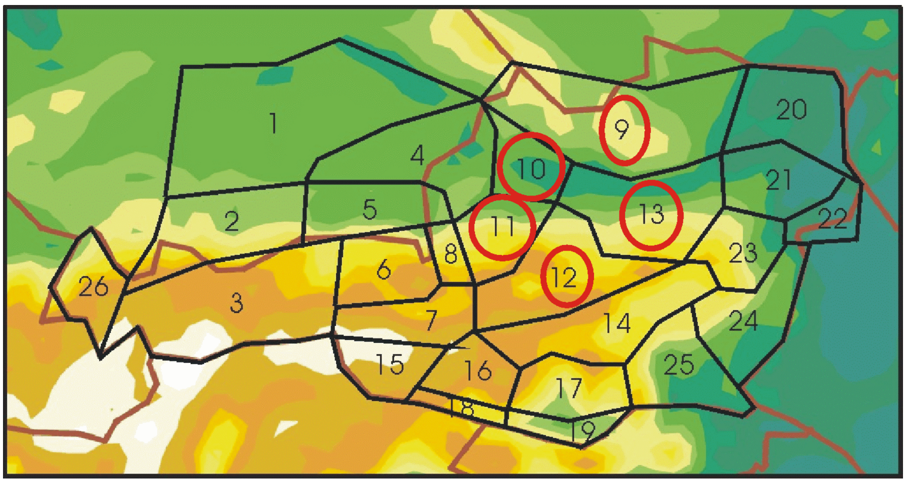


Fig. 14-1.1: Definition of catchment-type areas for operational precipitation analyses and forecasts. Areas marked in red were heavily affected by precipitation in August 2002 and chosen for more detailed studies.

14-2 Precipitation analysis uncertainty

The validation of areal precipitation forecasts contains not only model errors but also errors due to the uncertainty of observations and their spatial interpolation (Haiden, 1994).

In order to estimate the robustness of our forecast validation of the August 2002 flood event we compare different observation data sets and analysis methods. ZAMG operates about 140 TAWES stations in Austria which measure preciptation on a temporal resolution of ten minutes. On average, there is one station per 600 km² which gives only a crude estimate of areal precipitation amounts. This is especially true for mountaineous terrain, and generally during the summer season, when small-scale (10-20 km) convective cells dominate the heavy precipitation climate.

ZAMG's climate stations measure precipitation amount in three hours interval. The Hydrographische Zentralbüro (HZB) maintains roughly 1000 rain gauges in Austria, with a temporal resolution of 24 hours.

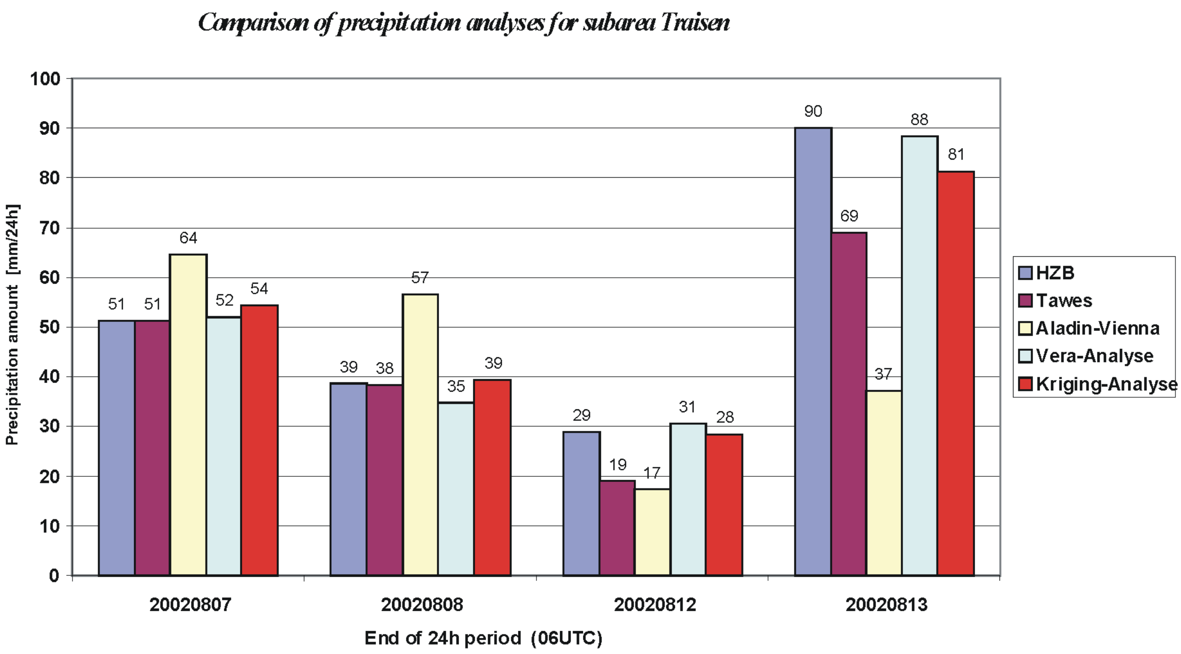


Fig. 14-2.1: Comparison of 24-hourly precipitation amounts obtained using different datasets and analysis methods for a sub-area of catchment area 13 (Traisen) for four different days of the August 2002 flood event. The columns denoted 'HZB', 'Vera-Analyse', and 'Kriging-Analyse' are based on HZB data. Precipitation amounts predicted by the ALADIN model are also shown.

Fig. 14-2.1 shows the effect of using different observational datasets and analysis methods for a spatial averaged precipitation sum. The chosen sub-area of the drainage area Traisen represent a domain of about 2000 km². The first part of the event (7-8 August 2002) was characterized by relatively homogeneous and widespread rainfall in this area. The precipitation values obtained from the HZB dataset and from the TAWES datasets agree to within 10% or better. The second part of the event (12-13 August 2002) was characterized by small scale convective cells which led to more significant differences between the spatially higher resolving HZB measurements and the TAWES data. For instance, the 24-hour precipitation total for 13 August 2002 increases by about 25% when HZB-measurements are used rather than TAWES. It is interesting to compare the spatially averaged precipitation of the VERA (5km resolution) and Kriging analyses (StartClim.13) to the sums gained by a simple distance-weighting algorithm of HZB-measurements. The resulting differences are small compared to the uncertainty between ZAMG and HZB datasets. For estimation of areal precipitation amounts the choice of the interpolation method seems to be definitely less critical than the choice of the input dataset. Figure 14-2.1 also shows that the forecast error is usually, but not always, larger than the differences between the observations.

The effect of using different datasets is of course strongly scale-dependent. Figure 14-2.2 shows the sensitivity of the forecast verification to the analysis method for areas of different size.

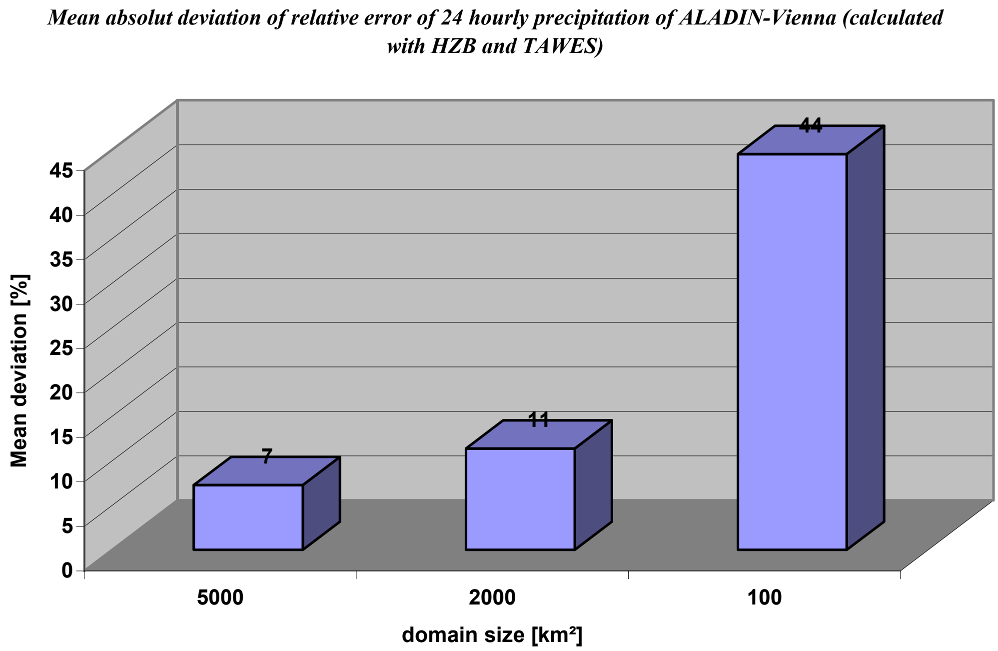


Fig. 14-2.2: Differences of the relative error of the ALADIN forecast verified with HZB vs. TAWES analyses. Average over 4 days from the August 2002 flood event, i.e. 7.,8.,12.,13.8.2002 (four 24h-totals) for the drainage area Traisen (5000 km²) and two sub-areas (2000 km² and 100 km²).

For a domain size of about 5000 km², the impact of a highly resolved observational database of HZB is found to be small (concerning spatial averages only!). The difference of the relative model error does not exceed 7% (left bar). A catchment size of 2000 km² shows only a moderate increase in uncertainty (11%, center bar). However, the relative difference between errors of a precipitation forecast for a domain size of about 100 km² exceeds 40%. As a rule of thumb we conclude that verification with TAWES analyses becomes doubtful for catchments smaller than 1000 km². The limiting catchment size for verification with HZB measurements is about one order of magnitude smaller, i.e. 100 km². This ratio roughly agrees with the ratio between the number of stations of both networks.

Radar observations of precipitation can give a much more detailed picture of the spatial characteristics of an event (especially under convective conditions) but they are affected by uncertainties due to orographic effects, precipitation shadowing, varying Z_R relationships, among other things. At ZAMG, a system that combines surface measurements and radar information (Borga et al., 2000) to produce high-resolution precipitation analyses (Buzzi et al., 2003) is under development. For the purpose of this study, however, only surface station data was used.

14-3 The August 2002 flood event

14-3.1 Verification of precipitation time series

In order to quantify the error of areal precipitation forecasts, it is necessary to define specific areas. Since 1999, ZAMG operationally provides 1-hourly precipitation analyses and forecasts for the areas shown in Fig. 14-1.1. With regard to the flood events 2002, areas 9 to 13 in the provinces of Upper and Lower Austria were particularly affected by heavy precipitation. Figure 14-3.1 gives an overview of observed precipitation amounts during the August 2002 flood event, based on TAWES observations.

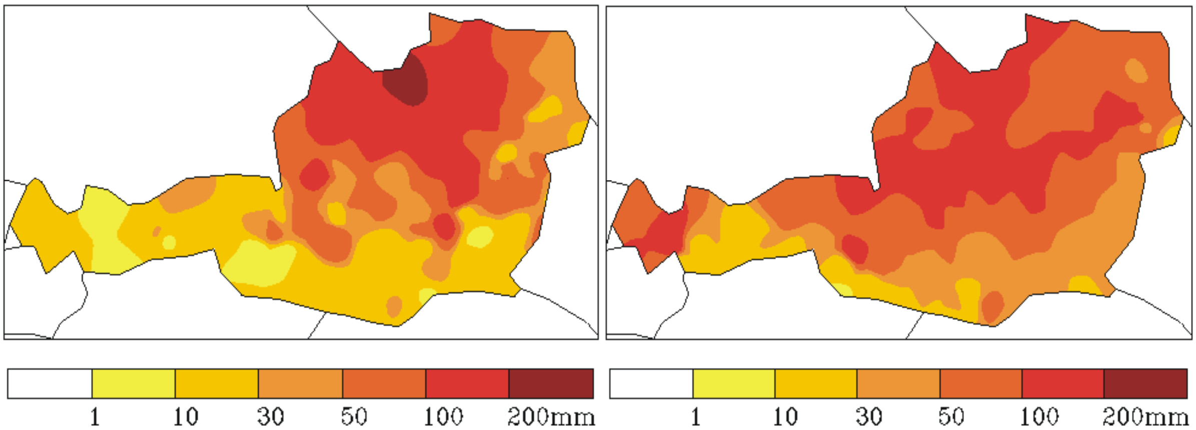


Fig. 14-3.1 TAWES precipitation sum interpolated on a regular 10x10km grid. Left: first part of the event (6.8.2002 12UTC – 8.8.2002 12UTC). Right: second part of the event (11.8.2002 12UTC – 13.8.2002 12UTC).

During the first part of the event, the center of precipitation was located in the area of Mühl- and Waldviertel. During the second part, these areas were hit again, but this time the more classical heavy precipitation regions along the northern alpine rim were affected as well.

In this study we we are able to make, for the first time, a direct comparison of heavy precipitation forecasts of two limited area models. The ALADIN model, which is run at ZAMG, and the Lokalmodell (LM) of the Deutsche Wetterdienst (DWD).

Figure 14-3.2 shows time series of forecasted and observed precipitation rate for the August 2002 event. During the first part of the flood event (7.-8.8.2002, top row) maximum intensities and total rainfall amounts were underestimated by both models. The rapid increase of rainfall intensity shortly after onset, and the bimodal temporal structure of the whole first part of the event were not captured. Note that the second peak within the first part of the event was not forecasted at all. It was this peak, however, which aggravated the already severe flooding situation in the catchment of the Kamp river.

Almost as important as the prediction of the onset of heavy rainfall is the prediction of its end. The upper right panel in Figure 14-3.2 shows that the end of the rainfall episode was more or less satisfactorily forecasted, with the ALADIN model giving a somewhat better indication of the actual ending than the LM.

The second event (bottom row of Figure 14-3.2), which was associated with a much larger low pressure system than the first one, shows generally better model results. The ALADIN model was able to forecast the hourly maximum at 11.8.2002 18UTC almost at the correct time, and with roughly the correct intensity. Other predicted intensity maxima, however, show little correspondence with observed intensity peaks. Again, the end of the epiosde is predicted more realistically by the ALADIN model.

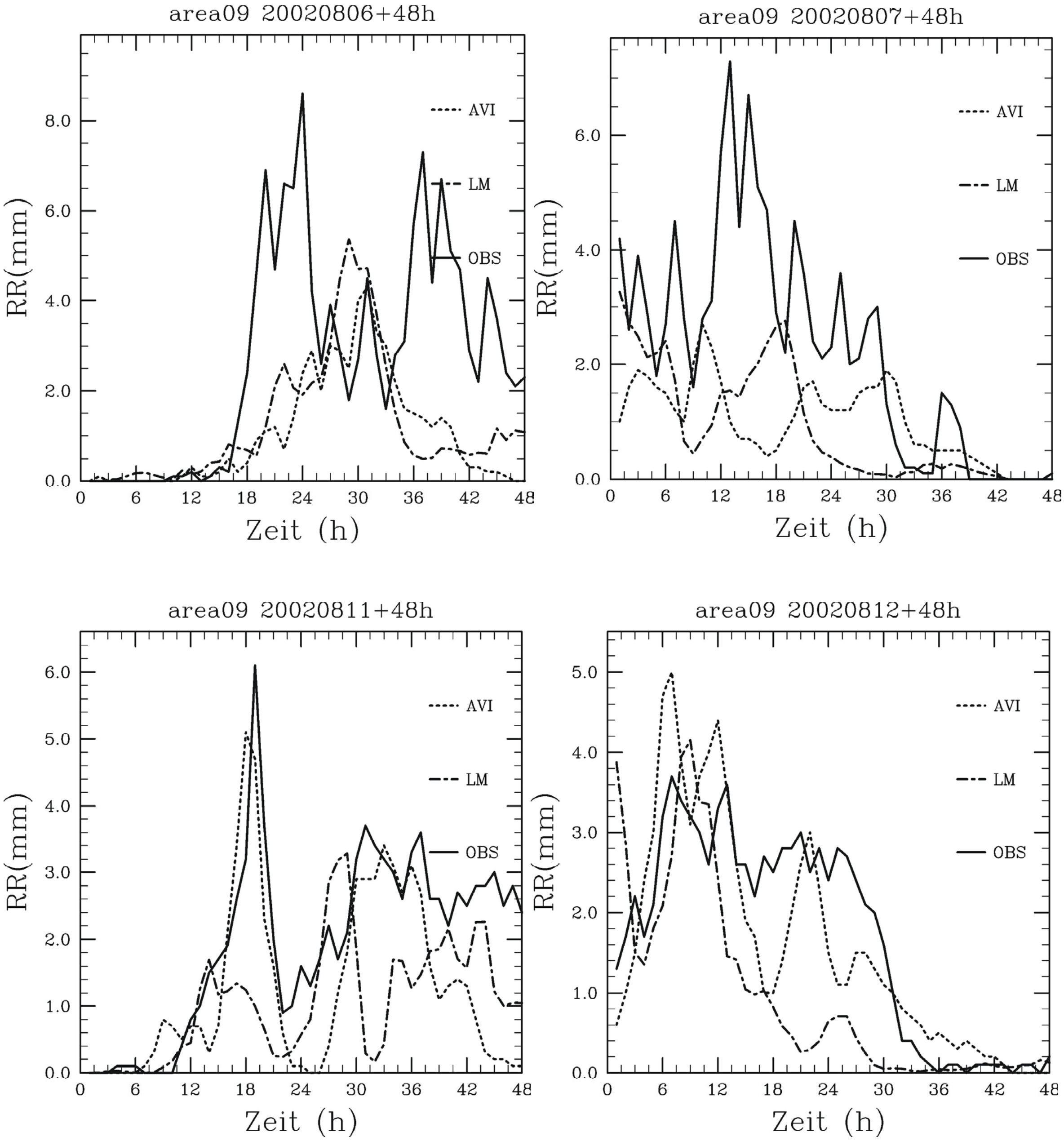


Fig. 14-3.2: Hourly intensity of spatially averaged precipitation for area 9 (Mühl-, Waldviertel). Solid line gives observations (OBS), dotted line the Aladin-Vienna forecast (AVI), and dashed-dotted line the forecast of the Lokalmodell (LM) of the DWD. Upper row shows results for the first flood event, for analysis times 6 and 7 August 2002, 00 UTC. Bottom row shows results for the second part of the event, for analysis times 11 and 12 August 2002, 00 UTC.

For hydrological purposes it is important to validate not just intensities but also cumulative rainfall amounts. Even if transient maxima and minima are not captured a model prediction can be useful as long as cumulative amounts are approximately correct. Figure 14-3.3 shows predicted and observed *cumulative* rainfall amounts, otherwise analogous to Figure 14-3.2.

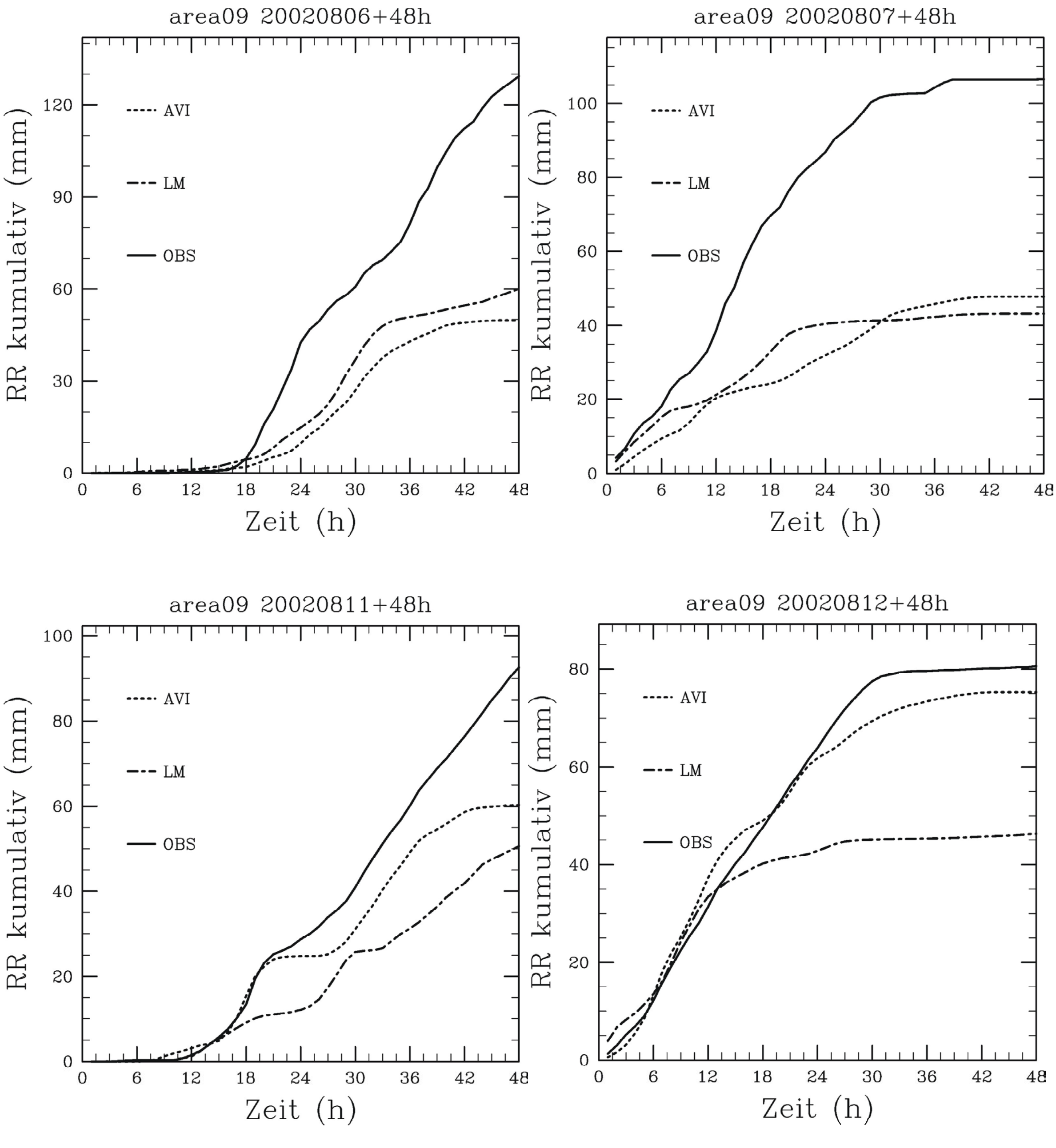


Fig. 14-3.3: As in Figure 14-3.2, but for cumulative rainfall amounts.

As was already indicated in the intensity diagrams, there is a general underestimation of the 48 hour precipitation total, especially on the 7th and 8th of August where both models end up with roughly 50% of the observed precipitation. The second part of the event (bottom row) was predicted more accurately, and errors after 48 h amount to about 10-30% for ALADIN and 30-50% for LM.

Previous verification studies at ZAMG have shown that in Austria’s lowland or hilly regions (including Mühl- and Waldviertel) heavy precipitation is generally more difficult to predict than in mountainous areas, especially along the alpine rim. This is because the blocking effect of the topography on the airflow introduces a deterministic element into the precipitation formation process. This tendency for reduced rainfall forecast errors in the ‘classical’ alpine upslope areas is illustrated in Figure 14-3.4 which gives results for area 13 (Traisen).

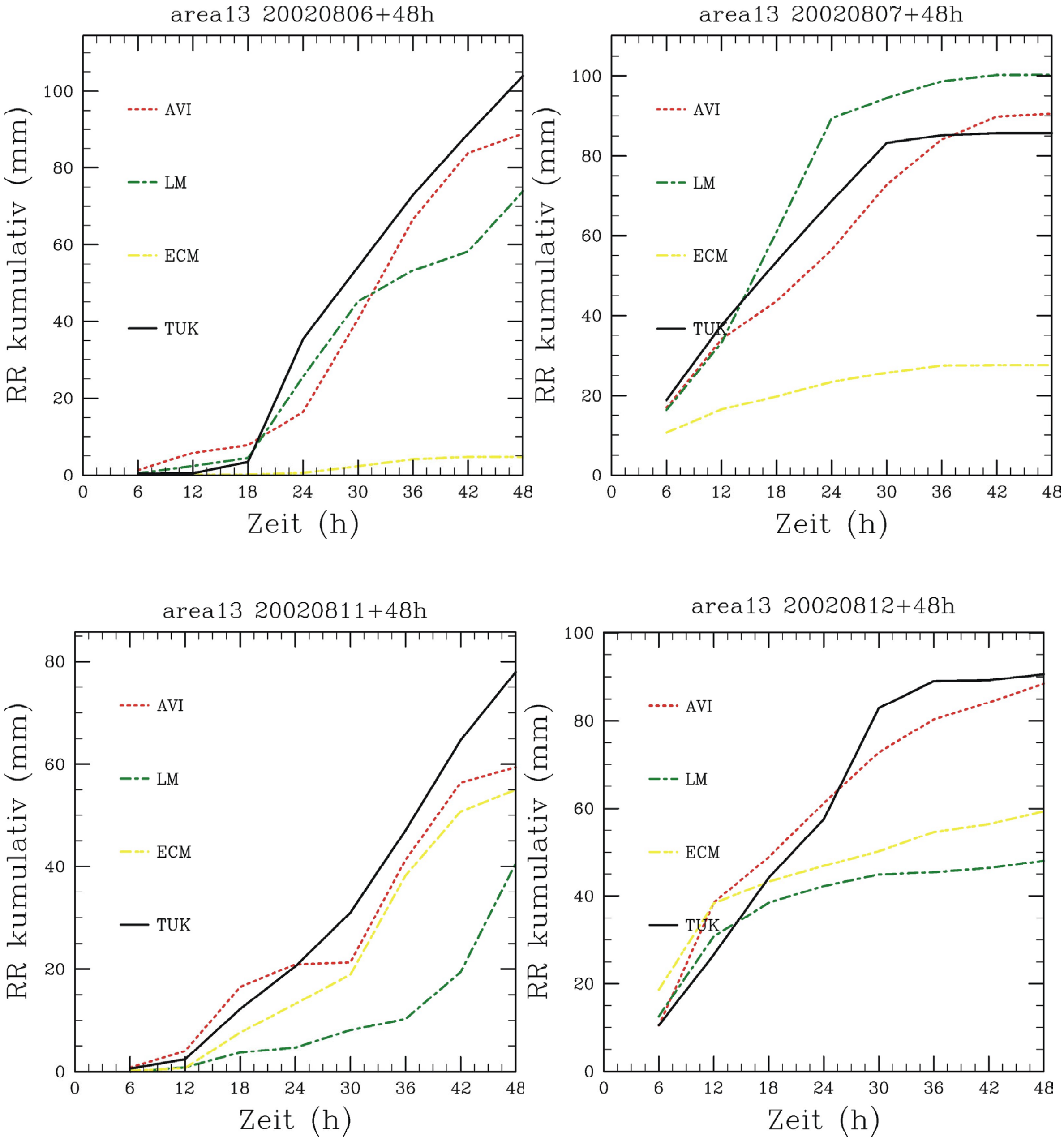


Fig. 14-3.4: Cumulative, spatially averaged precipitation for area 13 (Traisen). Black solid line gives observations derived from a combination of HZB and TAWES data (TUK), red dotted line the Aladin-Vienna forecast (AVI), green dashed-dotted line the forecast of the Lokalmodell (LM) of the DWD, and yellow dashed line the ECMWF forecast. Upper row shows results for the first flood event, for analysis times 6 and 7 August 2002, 00 UTC. Bottom row shows results for the second part of the event, for analysis times 11 and 12 August 2002, 00 UTC.

Here we show a comparison of three meteorological models with the 'best' observations (HZB and TAWES information combined). Unlike in the Kamp area, even the first part of the event (top row) was relatively well forecasted by the limited area models ALADIN and LM (relative errors in the range 10-30%) but not the global model ECMWF which missed out completely in this case. Overall, the ALADIN model gave the best results in this area.

14-3.2 A detailed study of the flood event 2002 for the drainage area Kamp

In general, heavy precipitation events do not often occur in the drainage area Kamp. Daily rainfall sums reach 100 mm only once per century, whereas in the alpine regions of Lower Austria 200 mm are observed within such a period (Nobilis et al., 1991). As large elevations are missing in this area, large precipitation amounts are mainly due to stationary frontal systems or upper level lows which are passing the Eastern Alpine region. In most of the cases deep convection, which is involved in or triggered by these systems, plays an important role. This is confirmed by the fact that more than 90% of annual maxima during 24 hours occur from April to September (Nobilis et al., 1991). (In the Alpine regions of Lower Austria the value is about 60-80%).

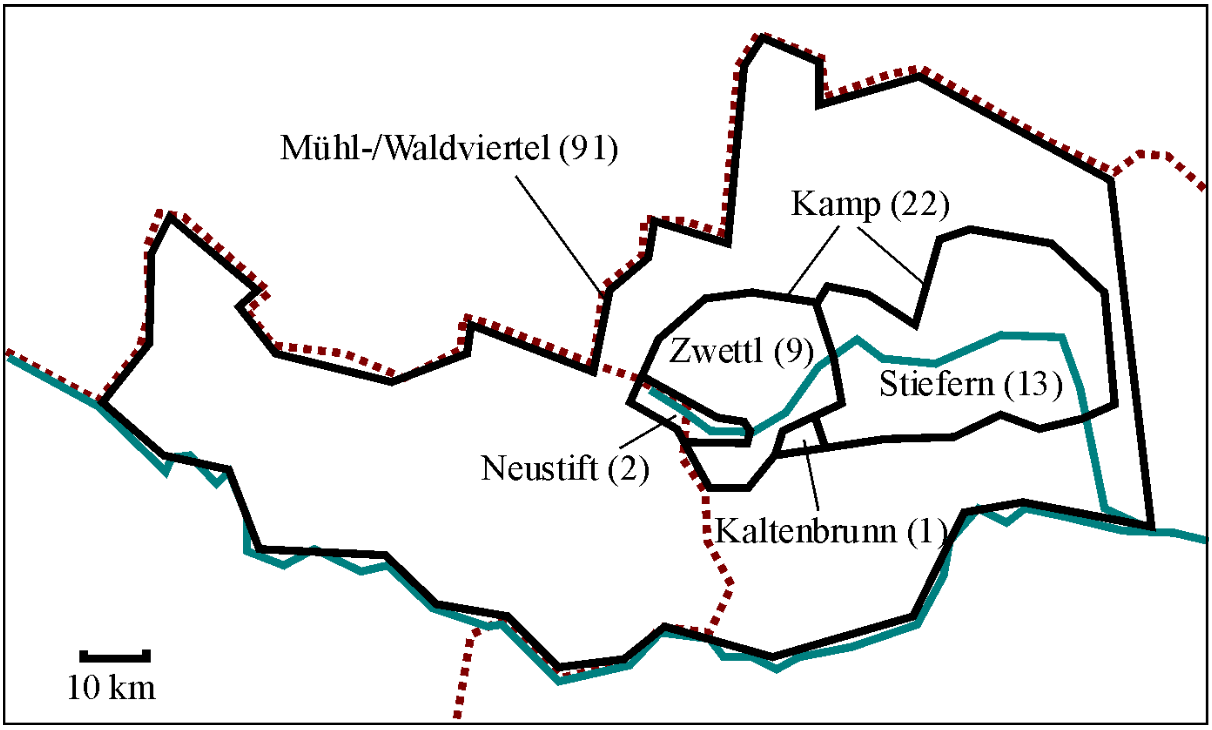


Fig. 14-3.5: Arrangement of area Mühl-/Waldviertel and of subareas within the drainage area *Kamp* that are used for the following verification. The values in brackets indicate the number of gridpoints of Aladin-Vienna forecast model (the mesh size is about 10km).

In order to study the precipitation forecast for different domain sizes, the area Mühl-/Waldviertel is divided into subareas as shown in Figure 14-3.5.

Tab. 14-3.1: Heavy precipitation events that are used in the verification for the area Kamp. Note that the precipitation amounts are 48 hour – sums, starting with 00UTC of the given day. In addition to observed heavy precipitation events those cases which were forecasted as intense events, were chosen too.

Datum		*ALADIN (mm)*	*TAWES (mm)*
1999	*8 Juli*	*43*	*55*
	9 Juli	*59*	*46*
	29 August	*29*	*3*
2000	*3 August*	*15*	*28*
	5 August	*50*	*35*
	6 August	*16*	*44*
	16 September	*10*	*27*
2001	*21 April*	*3*	*11*
	20 Juli	*25*	*50*
	31 August	*16*	*20*
	8 September	*12*	*12*
	29 Dezember	*2*	*8*
2002	*20 März*	*10*	*45*
	21 März	*14*	*35*
	6 Juni	*50*	*43*
	7 Juni	*21*	*35*
	6 August	*66*	*138*
	11 August	*53*	*70*
	12 Oktober	*19*	*40*

The hourly distribution of precipitation during the August 2002 flood events, forecasted by Aladin-Vienna, compared to observed precipitation amounts (both spatial averages) are shown in Figure 14-3.6.

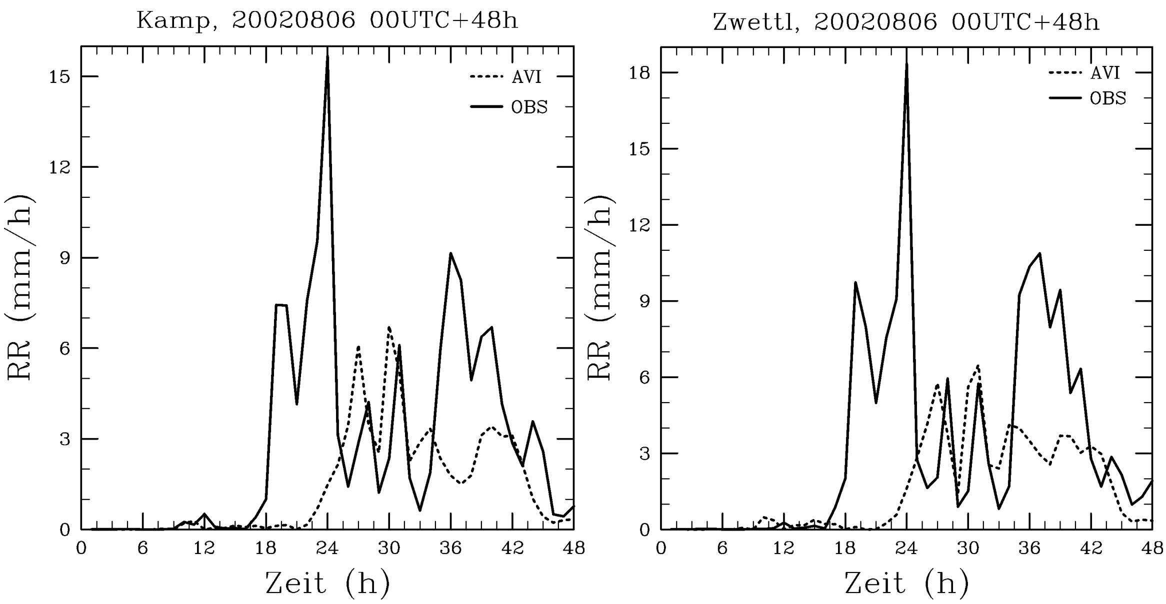


Fig. 14-3.6: Hourly intensities of precipitation from 6.8.2002 to 8.8.2002 for areas defined in Figure 14-3.5. Solid line indicates mean precipitation gained from TAWES measurements, dotted line represents ALADIN forecast (initial time: 6.8.2002 00UTC).

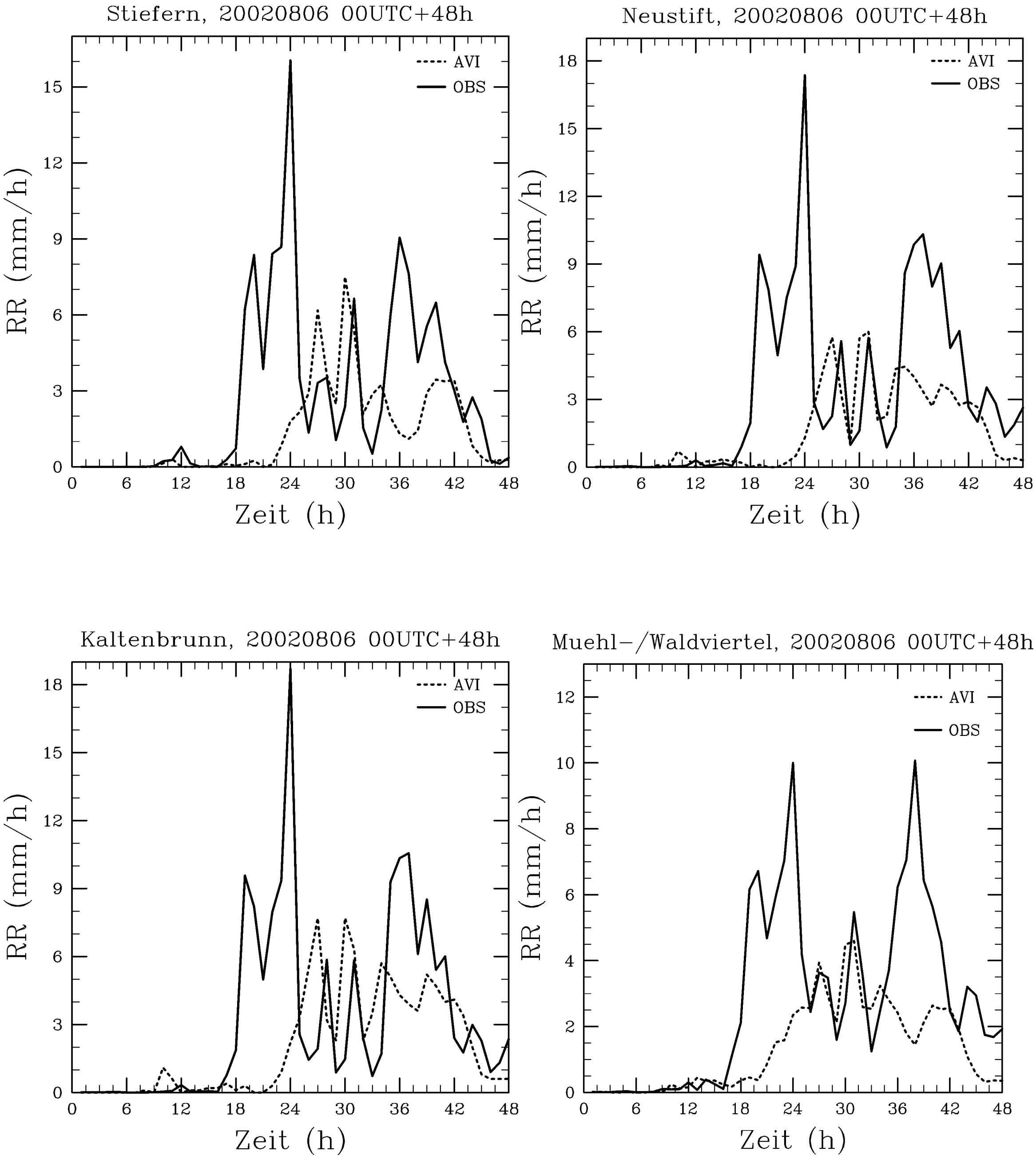


Fig. 14-3.6 (continued): Hourly intensities of precipitation from 6.8.2002 to 8.8.2002 for areas defined in Figure 14-3.11. Solid line indicates mean precipitation gained from TAWES measurements, dotted line represents ALADIN forecast (initial time: 6.8.2002 00UTC).

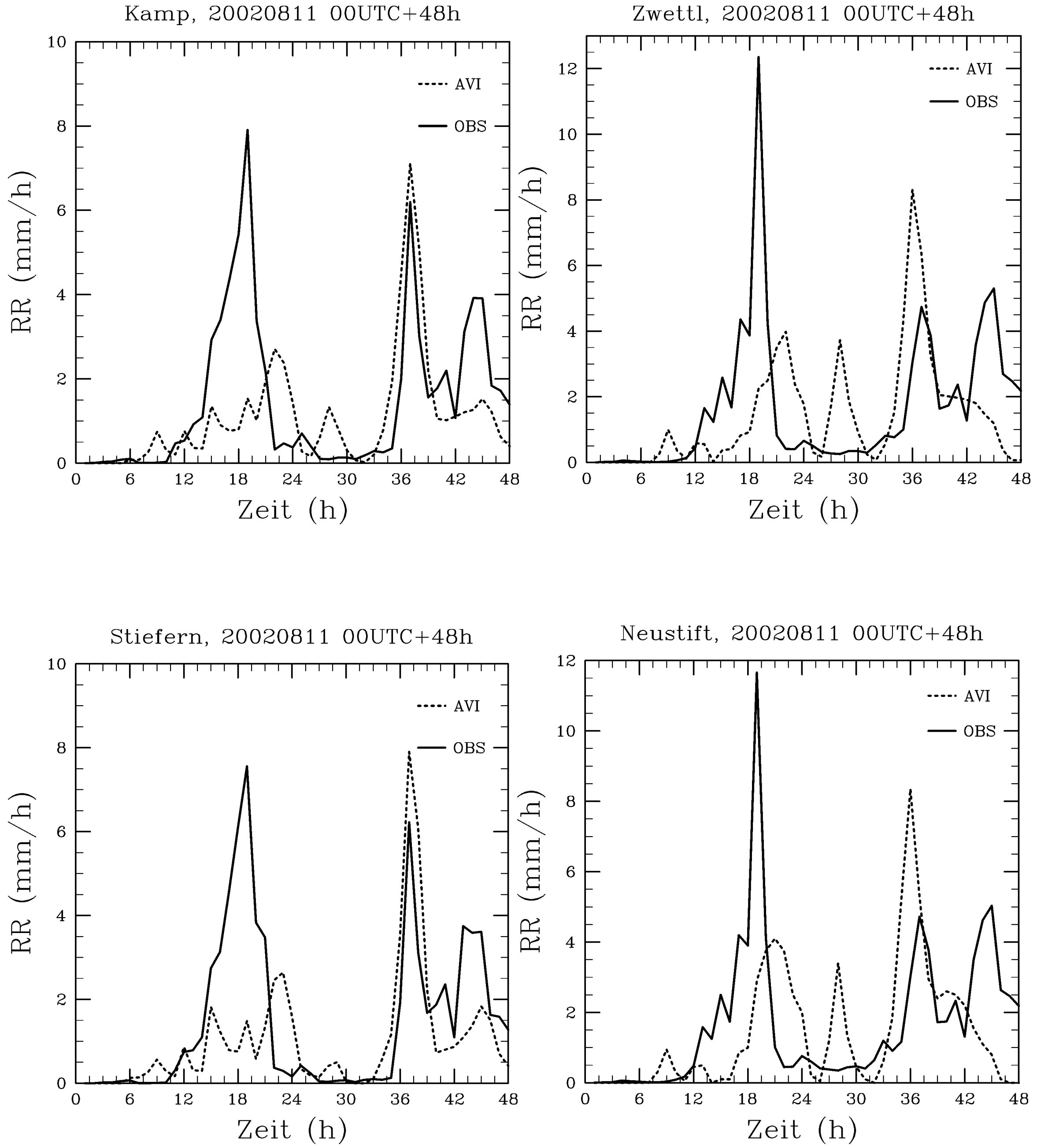


Fig. 14-3.7: Hourly intensities of precipitation from 11.8.2002 to 13.8.2002 for areas defined in Figure 14-3.5. Solid line indicates mean precipitation gained from TAWES measurements, dotted line represents ALADIN forecast (initial time: 11.8.2002 00UTC).

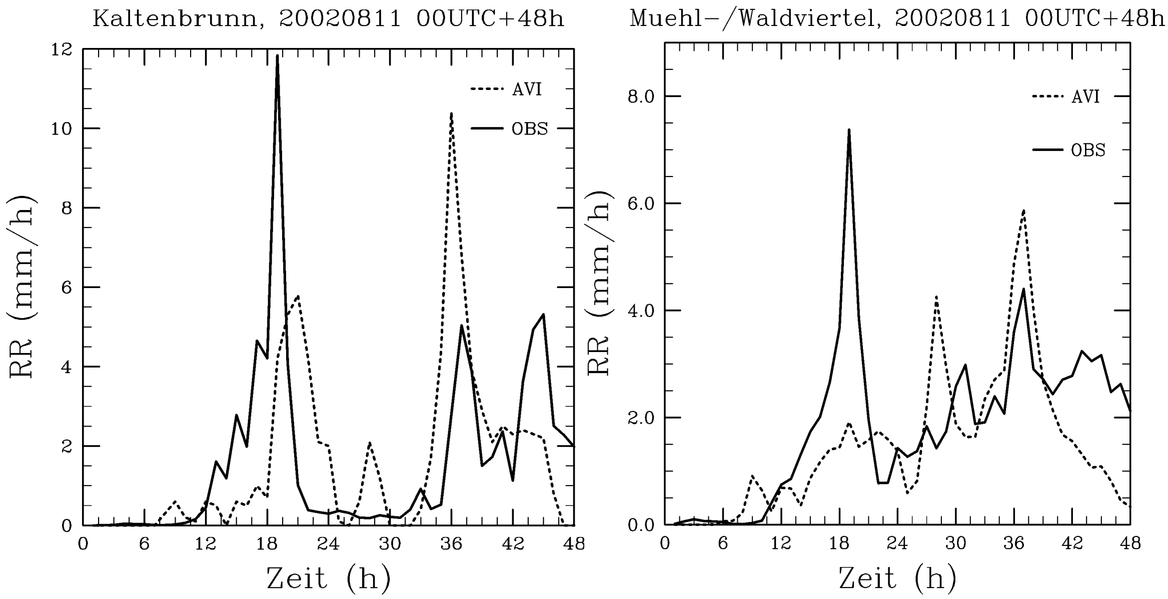


Fig. 14-3.7 (continued): Hourly intensities of precipitation from 11.8.2002 to 13.8.2002 for areas defined in Figure 14-3.11. Solid line indicates mean precipitation gained from TAWES measurements, dotted line represents ALADIN forecast (initial time: 11.8.2002 00UTC).

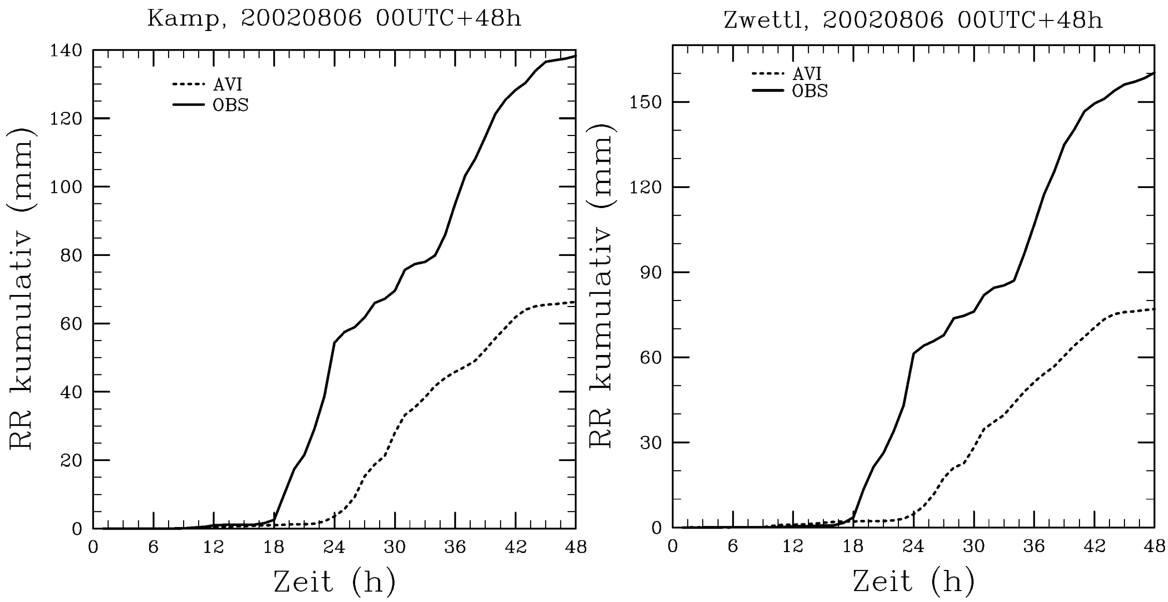


Fig. 14-3.8: As in Figure 14-3.6, but for cumulative rainfall amounts

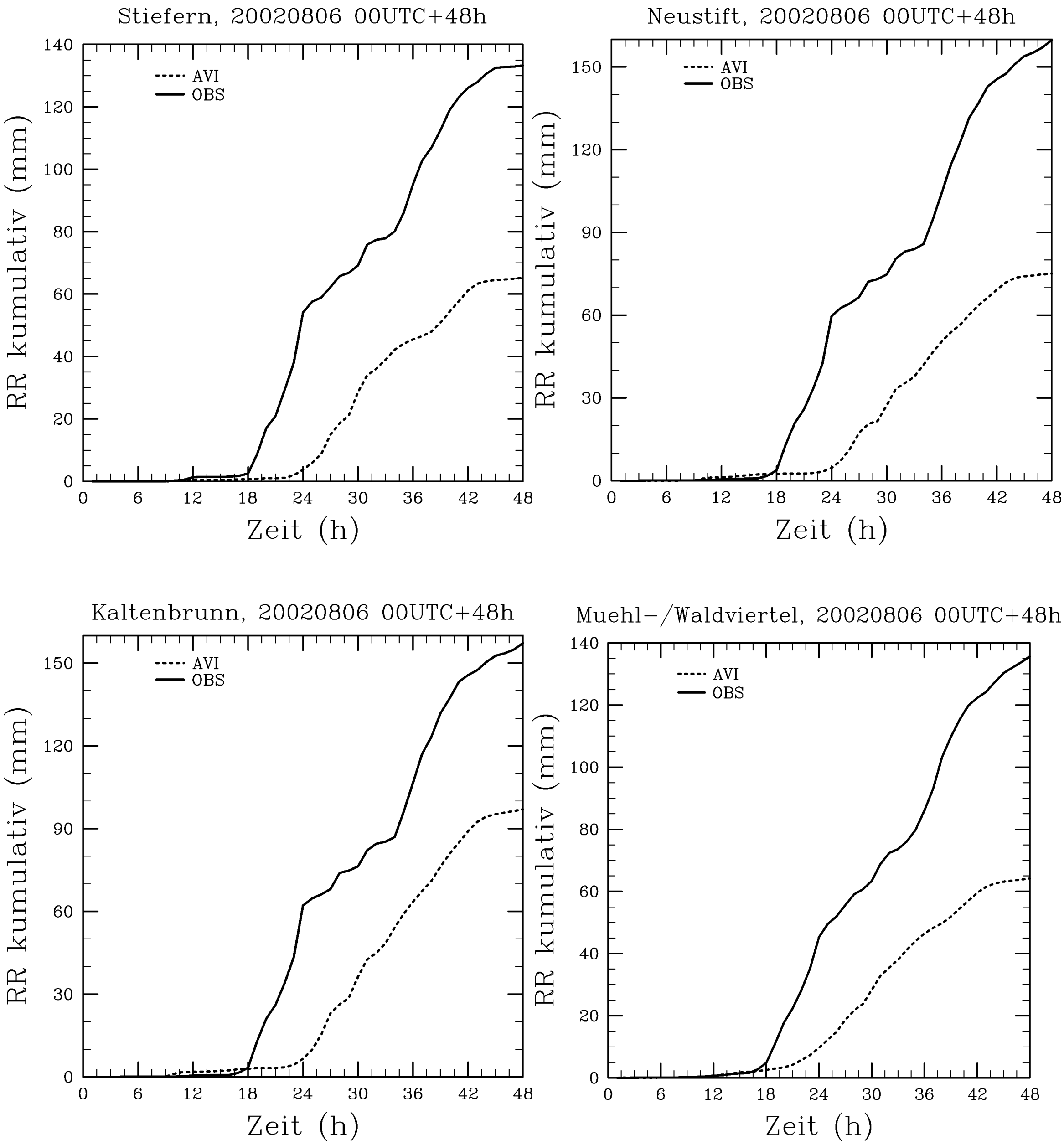


Fig. 14-3.8 (continued): As in Figure 14-3.6, but for cumulative rainfall amounts

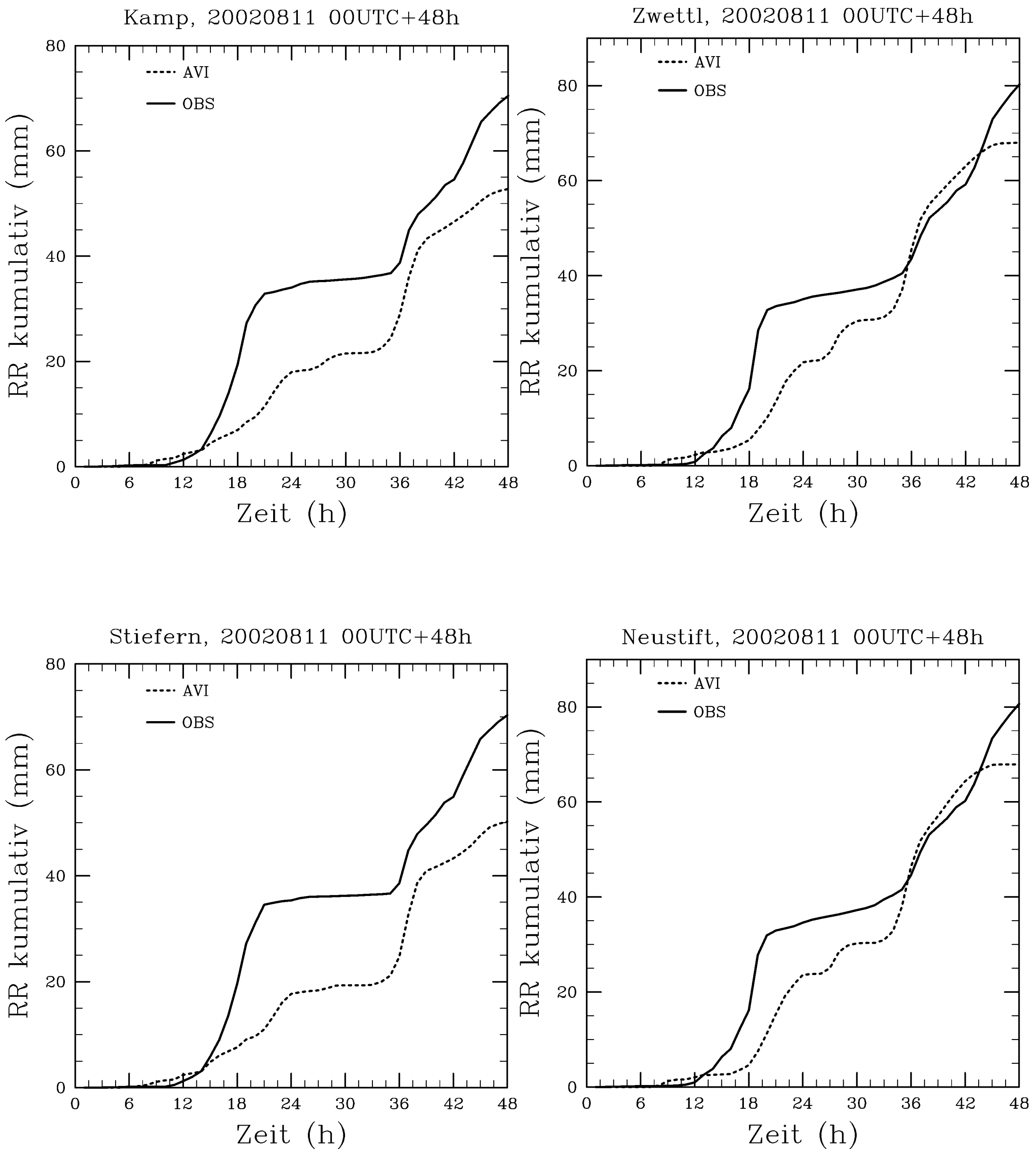
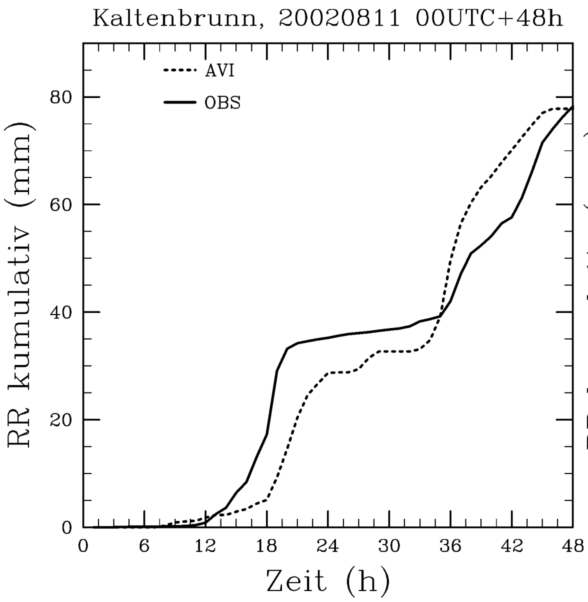


Fig. 14-3.9: As in Figure 14-3.7, but for cumulative rainfall amounts



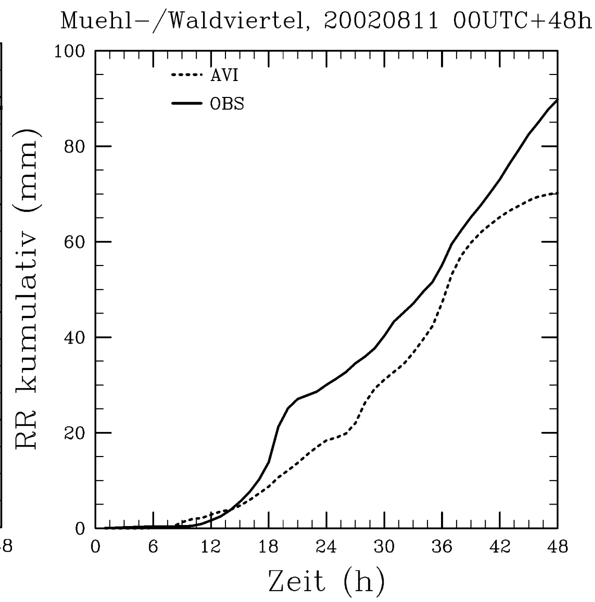


Fig. 14-3.9 (continued): As in Figure 14-3.7, but for cumulative rainfall amounts

The first maximum in both Figure 14-3.6 and Figure 14-3.7 is forecasted hardly, or not at all, regardless of the domain. The onset of the precipitation event seems to be delayed in many cases. On the other hand, the second peak during the second part of the flood event 2003 shows good forecast quality in terms of both intensity and timing. A significant reduction of errors by increasing the domain size does not occur, only the transition to the large area of Mühl-/Waldviertel yields an improvement. Mean precipitation forecasts for small areas like Kaltenbrunn or Neustift do not necessarily show less quality, as this special event was mainly characterized by large scale precipitation. However, for areas of small size, the analysed precipitation becomes uncertain.

In order to obtain information in addition to intensities depending on the domain size and forecast duration, a cumulative examination is chosen in Figure 14-3.8 and Figure 14-3.9. Precipitation amounts for the first part of the flood event 2002 are understimated by the model. The absolute error of the 48-hour precipitation forecast for the area Kamp is about 50%. In the case of the 11.8.2002 (second part of the flood event) the absolute error for all areas is smaller (about 25%), and the curve of cumulative rainfall amounts runs parallel to the observed one to a large extent.

Pointing out the dependency of the forecast error on duration and domain size, the absolute errors are calculated for 6-, 12-, 24- and 48-hour forecasts, averaged over the areas which are defined above (Figure 14-3.10). For most of the areas the mean error of precipitation forecasts exceed 70% for a duration of 6 hours, and drops below 50% by increasing the time scale of the prediction. Domain size influences the error just weakly, only the area Mühl-/Waldviertel is outstanding with lower values of mean errors (especially regarding predictions for 6 and 12 hours). The greater the time scale of prediction the less influence exerts the domain size on the result.

Regarding the area "Kaltenbrunn" (dark violet curve in Figure 14-3.10), which covers a region of about 100 km², the low values of absolute error are surprising. This seems to be a coincidence, as this domain is represented only by one single grid point in the model. Expanding the error distribution by using data from 1999 to 2002 (compare Table 14-3.1), supports this hypothesis (Figure 14-3.11). The mean error decreases by increasing the domain size and the forecast duration. Regarding predictions for 6 hours, the error exeeds 100%, whereas the prediction for the largest region (Mühl-/Waldviertel) is afflicted with an error of approximately 80%. Increasing the period of the forecast to up to 48 hours, almost all areal mean precipitation errors drop below 50%.

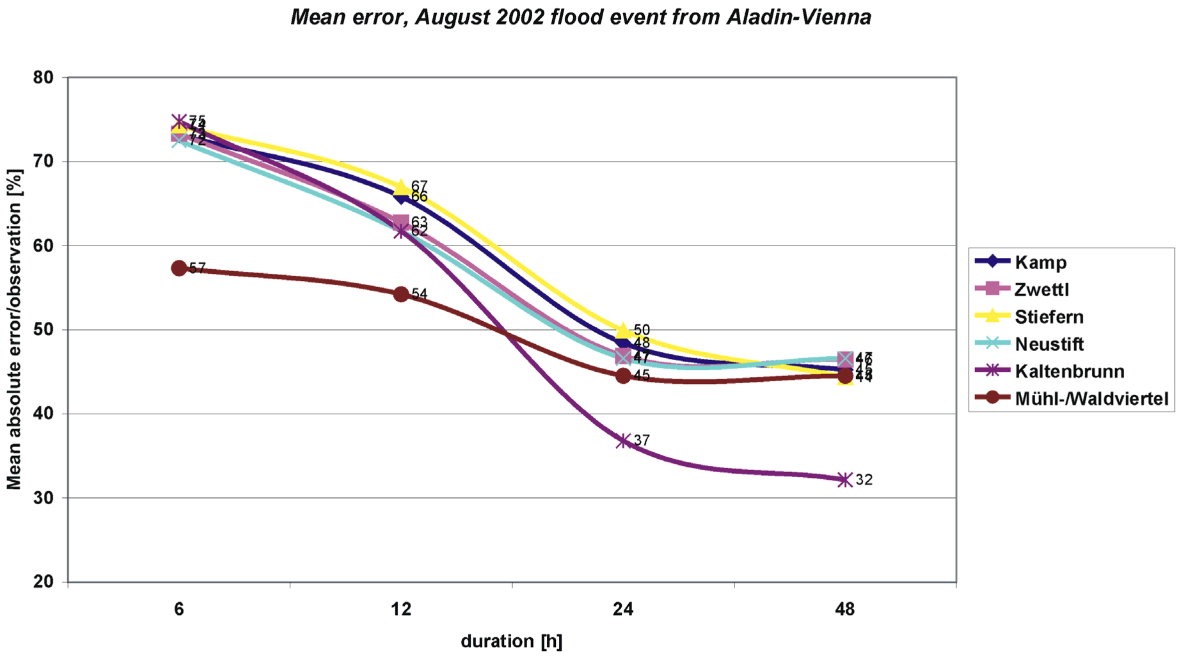


Fig. 14-3.10: Mean absolute forecast error in percent of precipitation forecast of ALADIN (August 2002 flooding) as a function of forecast duration for different subareas within area "Mühl-/Waldviertel".

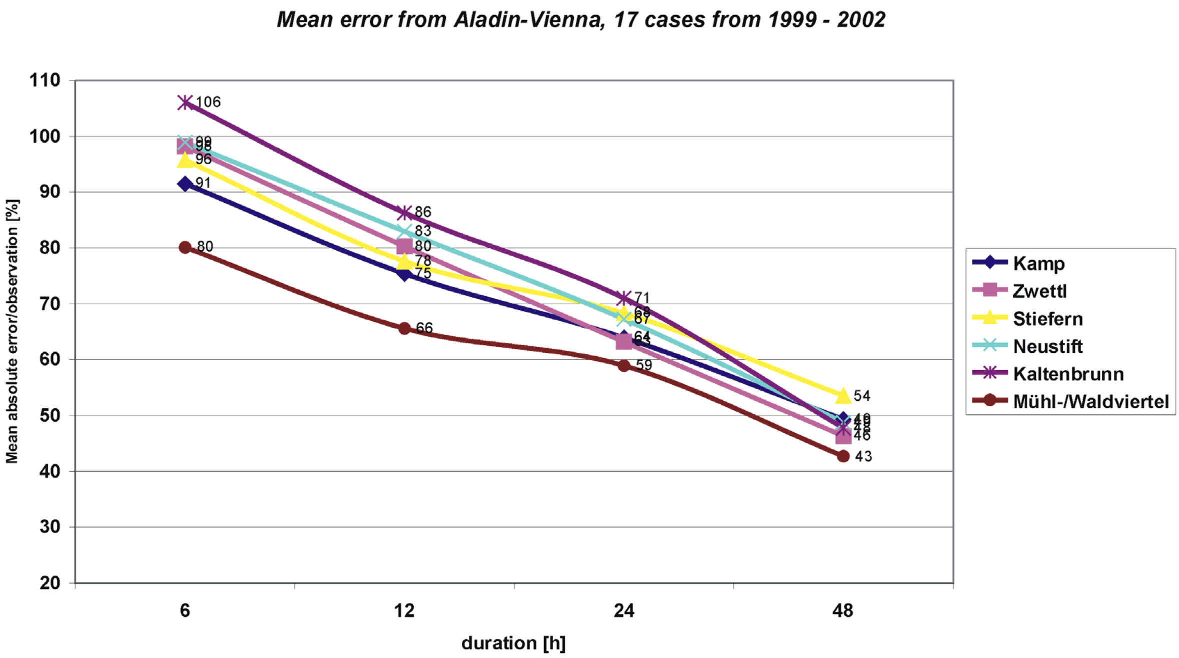


Fig. 14-3.11: Mean absolute forecast error in percent of precipitation forecast of ALADIN (17 cases from 1999 to 2002) as a function of forecast duration for different subareas within area "Mühl-/Waldviertel".

14-3.3 Error statistics

If errors in the forecast of precipitation intensity change sign within an event, it can be expected that there is a tendency for error compensation as we go from shorter to longer durations (Figure 14-3.5). It was investigated to what extent the prediction of 48-h totals has smaller relative errors that the prediction of 6-h totals (and totals for durations in between). This question is relevant for the design of flood warning systems.

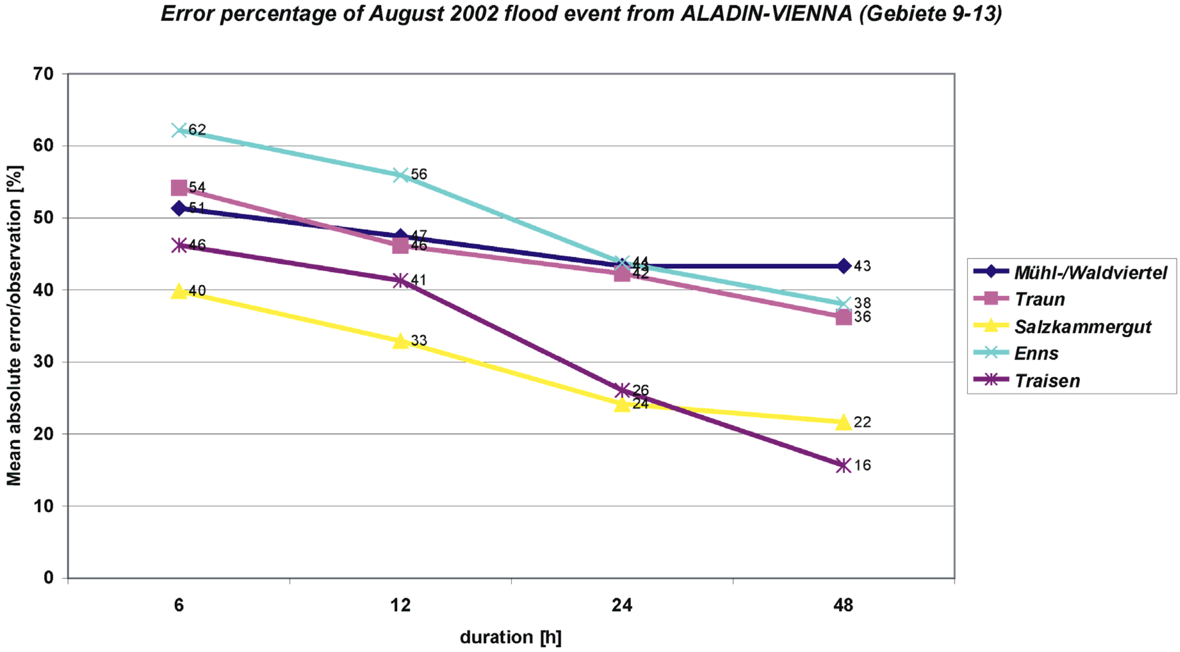


Fig. 14-3.5: Mean absolute forecast error in percent of precipitation forecast of ALADIN (August 2002 flooding) as a function of forecast duration for different areas.

For all areas the mean absolute error decreases with increasing forecast duration. Typically, the relative absolute error drops from 40-60% at 6-hourly duration to 20-40% at 48-hourly duration. This is especially obvious for the areas Traisen and Enns, which contain large mountainous areas. Another area where orographic blocking effects play an essential role is the region "Salzkammergut", where the model shows the best results for short forecast periods. On the other hand the model output does not significantly improve with increasing duration in this area. The region Mühl-/Waldviertel shows the smallest temporal compensation effect.

Figure 14-3.6 indicates similar tendencies for the LM model, apart from the error in the region Salzkammergut which is much higher than in ALADIN. A closer look at the precipitation intensities on single forecast runs showed that LM extremely overestimated the 48-h precipitaion sum at the beginning of the flood event.

In contrast to limited area models like LM or ALADIN, the global model of the European Center is not able to simulate the actual precipitation amounts occurring during severe events. This is mainly due to the lower horizontal resolution which leads to smoother precipitation fields. For example, heavy precipitation due to blocking effects is smoothed out and therefore locally understimated. As a result, the relative mean absolute error does not vary much with duration and location (Figure 14-3.7). Even for a duration of 48 hours the model error does not drop below 40%.

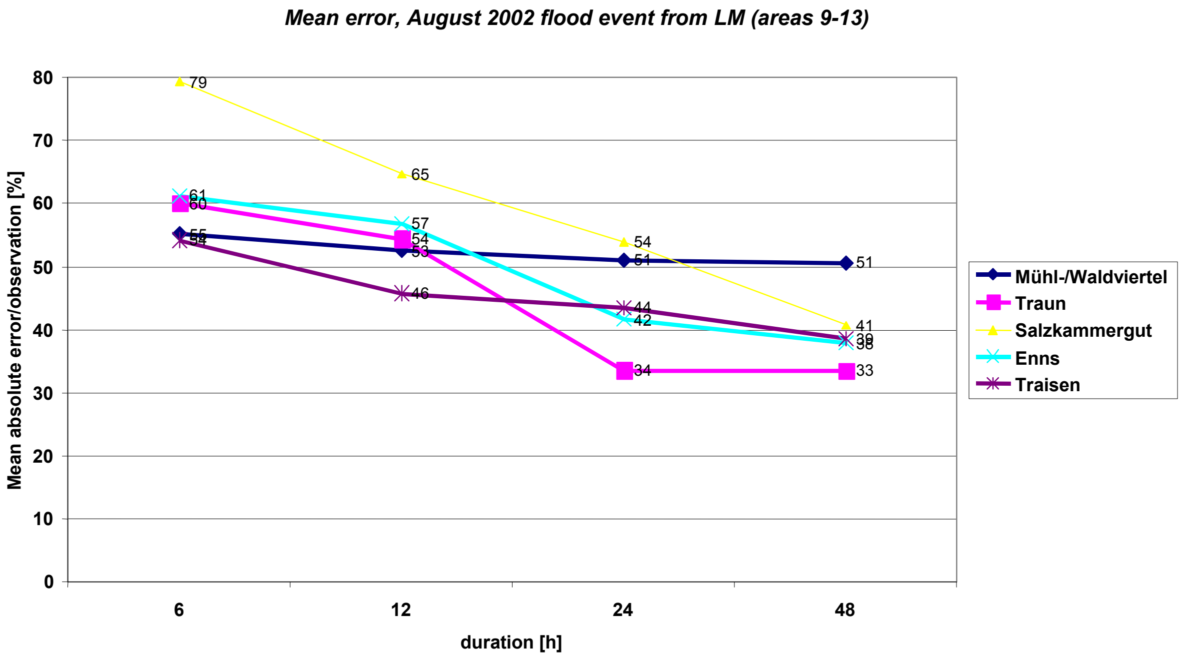


Fig. 14-3.6: Mean absolute forecast error in percent of precipitation forecast of LM (August 2002 flooding) as a function of forecast duration for different areas.

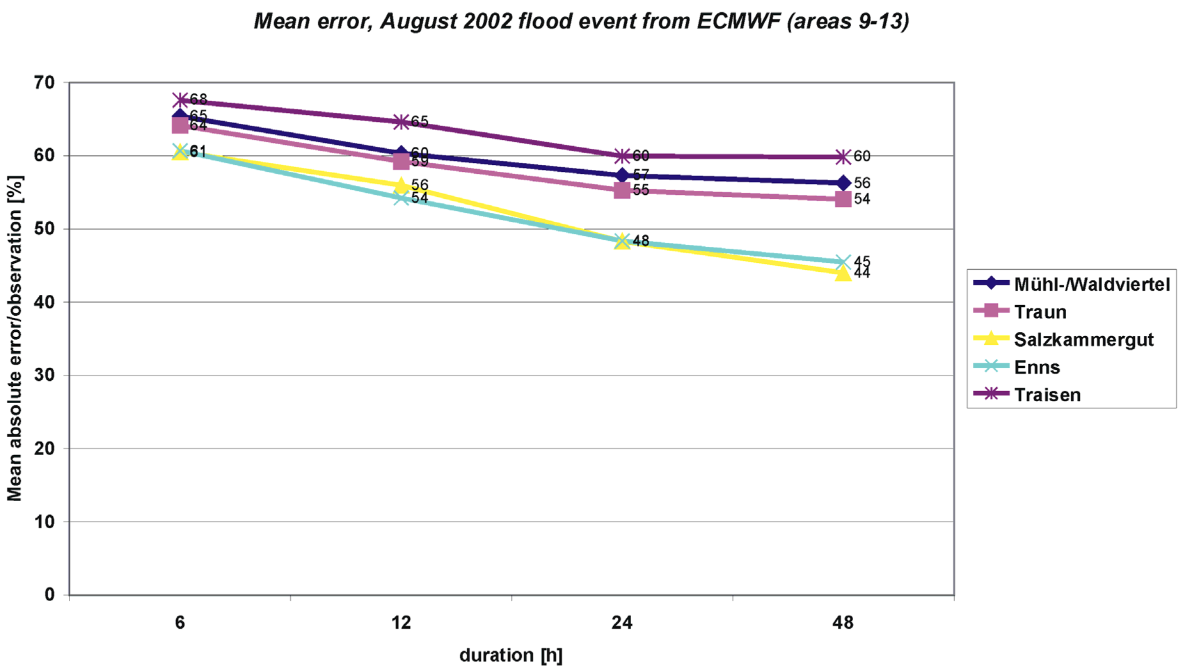


Fig. 14-3.7: Mean absolute forecast error in percent of precipitation forecast of ECMWF operational run (August 2002 flooding) as a function of forecast duration for different areas.

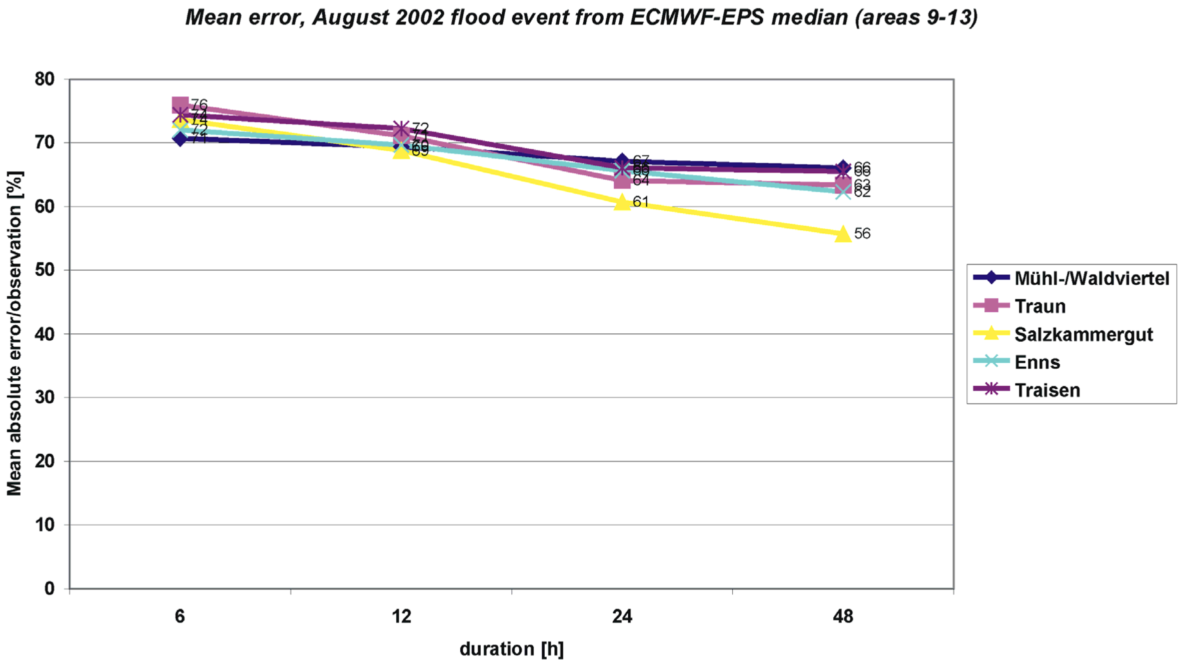


Fig. 14-3.8: Mean absolute forecast error in percent of precipitation forecast of ECMWF EPS-MEDIAN (August 2002 flooding) as a function of forecast duration for different areas.

In order to estimate the uncertainty of a forecast, the use of the Ensemble Prediction System (EPS) of the ECMWF has become indispensable in recent years. The system consists of 51 different forecasts (operational run included), with each one slightly disturbed in its initial state compared to the reference run. The resulting bundle of forecasts can be used to derive percentiles, or probabilities, of precipitation exceedance.

During the flood event 2002, the median, or 50% percentile, of ECMWF-EPS gives little indication for a heavy preciptation event and produces roughly the same error values as the reference run (cf. Figures 14-3.7 and 14-3.8). For the August 2002 event we must increase the percentile up to 90% in order to gain a signal for extreme precipitation and obtain reduced errors (Figure 14-3.9). Interestingly, this percentile also gives a much more pronounced reduction of error with increasing duration. The mean absolute errors of the 90% percentiles are about 55 – 75% for a 6 hour duration, but decrease up to 15 – 32 % for 48 hours. An exception is the region Enns, where the 90% percentile significantly overstimates the intensity of precipitation amount (50% error).

Concering an efficient warning system the prognostic value of using high percentiles is nevertheless doubtful, as it will lead to frequent false alarms if applied on a regular basis. However, it could be used to issue pre-warnings (or ‘watches’).

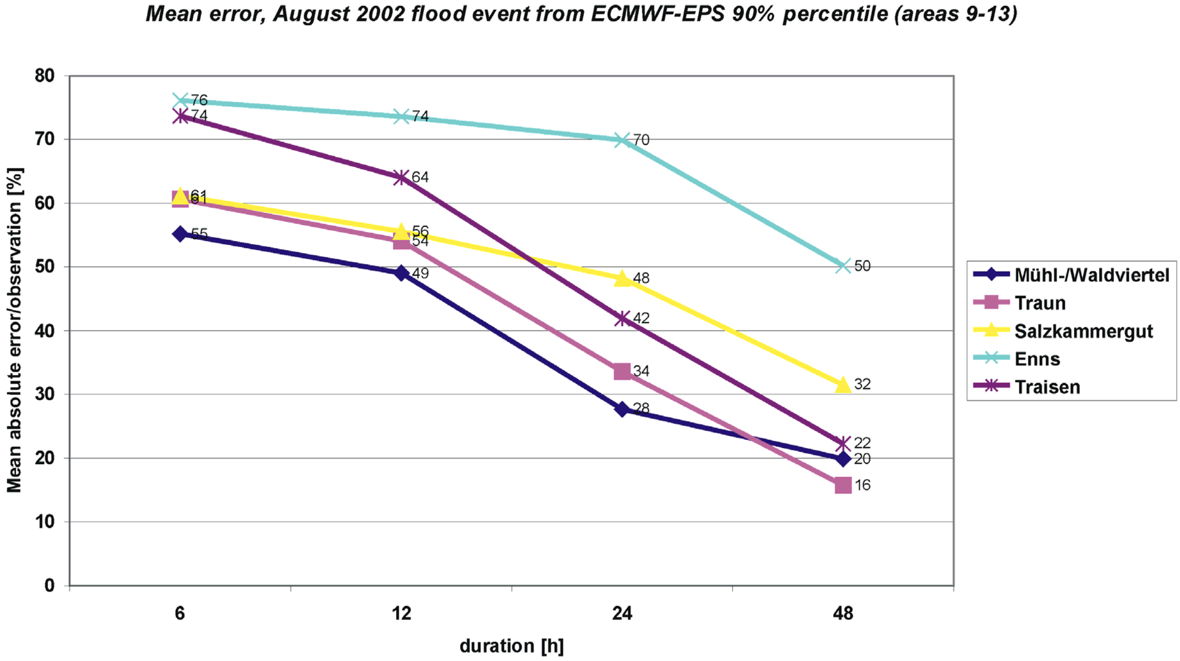


Fig. 14-3.8: Mean absolute forecast error in percent of precipitation forecast of ECMWF EPS-90% (August 2002 flooding) as a function of forecast duration for different areas.

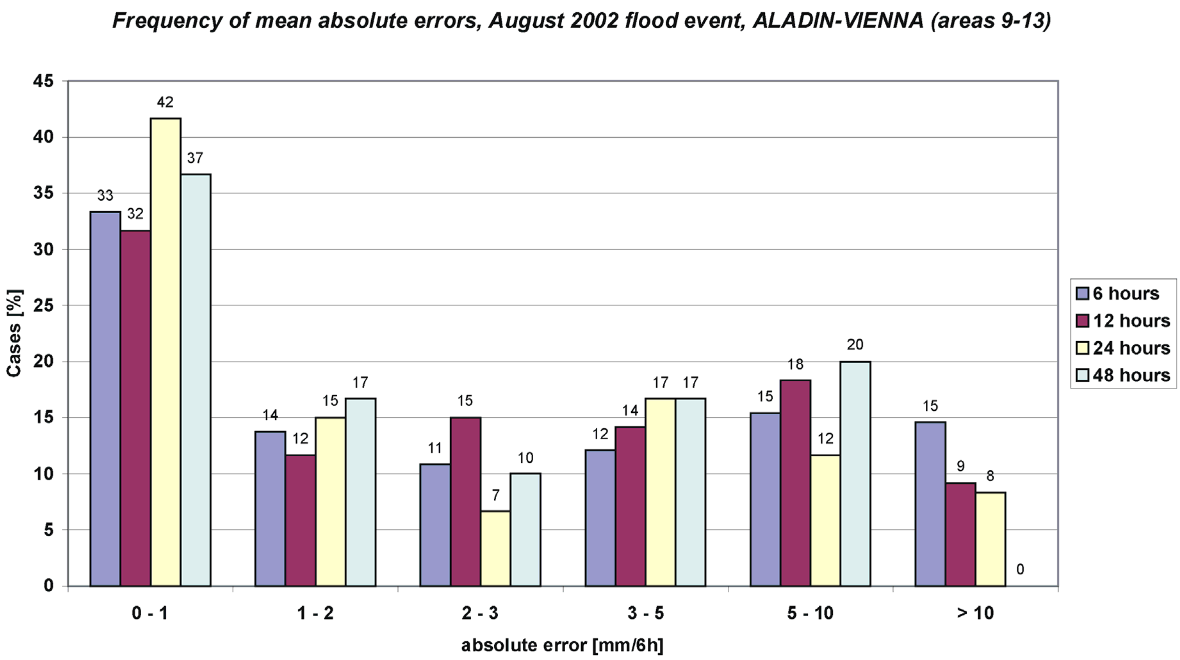


Fig. 14-3.9: Frequency distribution of mean absolute errors / 6 hours , divided in 6 intensity categories, for four forecast durations (6, 12, 24, 48 hours).

For the study of the error distribution as a function of forecast duration, we defined 6 categories of 6-hourly rainfall intensity. The frequency of occurrence within these classes for different durations is shown in Figure 14-3.9 as percentage of number of cases.

Apart from the category with little precipitation (0-1 mm / 6 hours), which includes about 30 – 40% of all cases, the frequency is distributed rather homogenously. The effect of compensation of errors for increased forecast duration is most pronounced in the highest category (>10mm/6h). There, the relative occurrence decreases from 8-9 % to 0 %.

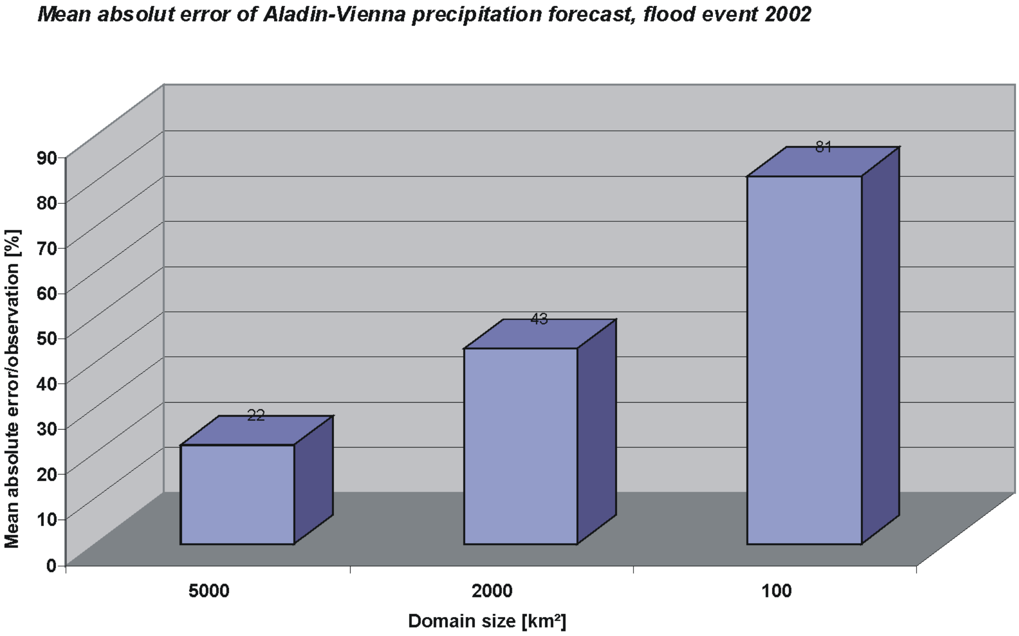


Fig. 14-3.10: Relative error of the precipitation forecast (ALADIN), averaged over the August 2002 flood event (7.,8.,12.,13.8.2002, 24-h totals) for drainage area Traisen (5000 km²) and two smaller sub-areas (2000 km² and 100 km²). Verified with HZB observations.

Similar to the error reduction associated with increasing time scales there is a spatial compensation effect that makes forecasts for larger catchments less difficult. Reducing the area size from 5000 km² to 2000 km² (within drainage area Traisen), the mean absolute error of the ALADIN precipitation forecast (using hydrological measurements for verification) almost doubles. Downsizing the domain further to 100 km² the error roughly doubles again (Figure 14-3.10).

14-3.4 Summary of forecast quality during the flood event

a) Dependency on forecast duration (ALADIN and LM forecast):

The mean absolute error of precipitation is reduced by 20 – 25 %, if duration is increased from 6 to 48 hours. Regarding spatial means the reduction of forecast errors is most distinct in areas where blocking effects play an essential role.

b) Dependency on domain size in a mountainous area (ALADIN forecast):

Decreasing the domain size from 5000 km² to 2000 km² approximately doubles the mean error. Decreasing the domain size from 2000 km² to 100 km² approximately doubles the mean error once again. Although this area dependency has been evaluated for a specific catchment only, we expect the order of magnitude to carry over to other areas

c) Use of ensemble prediction forecasts:

Mean, median, 50% and 75% percentiles are insufficient to reduce the precipitation error significantly, only 90% percentiles gave results that had some flood warning potential. Thus the problem of frequent false alarms must be addressed in the design of flood warning systems which are based on meteorological ensemble predictions.

14-4 Other events in 1999-2002

14-4.1 Verification of precipitation time series

The actual value of a model forecast for the August 2002 flood event is strongly connected to the model's ability to forecast extreme events within a longer time period. First, the results obtained from studying one episode might not be representative for the model's behaviour over a longer time period. Second, if two models have the same error for the August 2002 case but one of them has a significantly higher false alarm rate then the value of its forecasts is automatically reduced.

During the period from 1999 to 2002 the highest spatial averaged precipitation was observed from 21.05.1999 06UTC to 22.05.1999 06UTC in the area 26 (=Vorarlberg, compare allocation of areas in Figure 14-1.1) with 119 mm in 24 hours. For this interval the model output of the operational ALADIN model was 82 mm (absolute error 31%).

Figure 14-4.1 confirms the characteristic observed during the Agust 2002 event, namely that observed single-hourly peaks are rarely simulated by the model. However, the overall timing and temporal evolution of the event corresponds rather well to observations. This fact is confirmed by the value of absolute errors for different time periods (12, 24, 48 hours), which differ only between 23% and 26%.

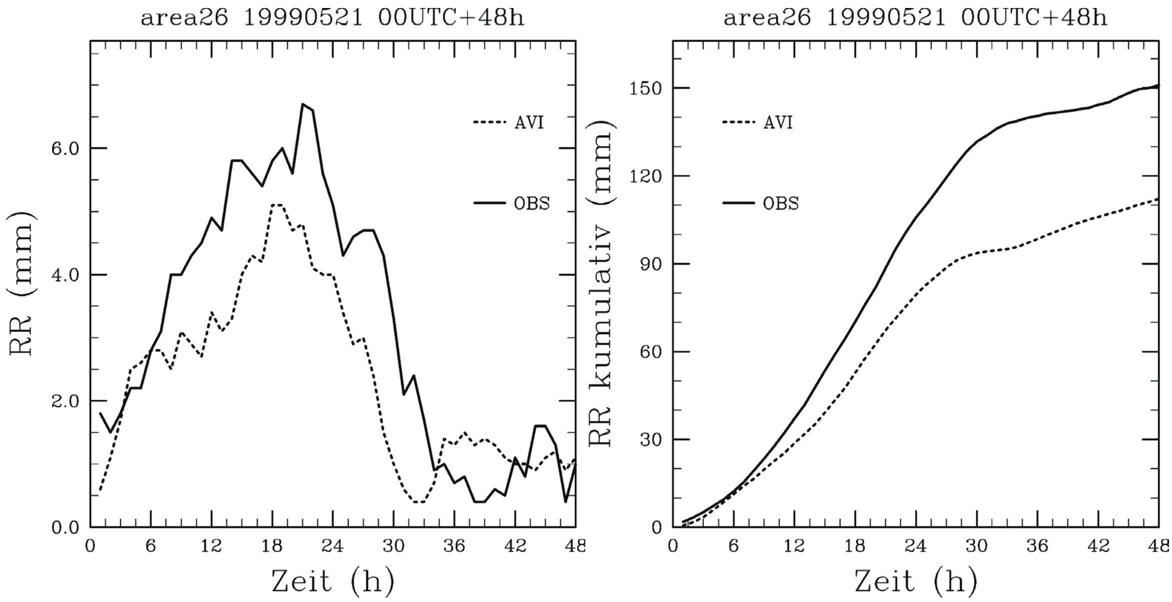


Fig. 14-4.1: Mean precipitation for area 26 (Vorarlberg) from 21.05.1999 00UTC to 23.5.1999 00UTC. Left panel shows intensities per hour, right panel the corresponding accumulated precipitation. Solid line indicates mean precipitation gained from TAWES measurements, dotted line represents ALADIN forecast (initialisation time: 21.05.1999 00UTC).

Tab. 14-4.1: Overview about the highest 24 hourly areal precipitation totals predicted by Aladin-Vienna 00UTC run compared to TAWES observation, on which the error statistics is based on. The sums are ranked in decreasing order, bold numbers indicate August 2002 cases. The verification period is January 1999 until August 2002.

Area	Strongest predicted cases			Strongest observed cases		
	Date	ALADIN	TAWES	Date	TAWES	ALADIN
01 Bayerisches Alpenvorland	1999-05-22	50	44	2002-03-20	46	30
	2002-03-21	38	22	1999-05-22	44	50
	2002-08-12	**36**	**38**	**2002-08-07**	**40**	**14**
	2000-03-18	34	16	**2002-08-12**	**38**	**36**
	2001-04-22	31	7	2000-09-22	38	16
02 Lech/Isar	**2002-08-12**	**89**	**61**	2002-03-20	73	69
	1999-05-21	79	24	**2002-08-12**	**61**	**89**
	2002-03-20	69	73	2000-08-07	48	46
	2000-03-18	59	47	2000-03-18	47	59
	2002-03-22	55	22	2001-06-11	39	29
03 Tiroler Inn	1999-05-21	67	19	1999-05-22	72	62
	1999-05-22	62	72	2002-03-20	56	40
	2002-08-12	**59**	**39**	2001-06-11	46	39
	2002-08-07	**55**	**24**	2000-03-18	44	37
	2002-06-28	44	15	2000-09-21	40	17
04 Deutscher Inn	**2002-08-12**	**69**	**50**	**2002-08-07**	**55**	**44**
	2002-03-21	53	40	**2002-08-12**	**50**	**69**
	1999-05-22	47	19	2002-03-21	40	53
	2000-03-18	45	21	**2002-08-13**	**37**	**29**
	2002-08-07	**44**	**55**	2002-03-20	36	41
05 Suedl. deutscher Inn	2000-03-18	76	53	**2002-08-13**	**33**	**29**
	2002-08-07	**74**	**69**	2002-03-20	60	69
	2002-03-20	69	60	2000-03-18	53	76
	2002-03-21	48	39	1999-05-21	43	48
	1999-05-21	48	43	2001-06-11	41	24
06 Saalach	**2002-08-12**	**83**	**73**	**2002-08-12**	**73**	**83**
	2002-03-20	70	55	2000-03-18	58	60
	2002-03-21	69	36	2002-03-20	55	70
	2000-01-31	67	26	2001-06-11	41	38
	2001-09-09	61	21	2001-12-07	40	42
07 Obere Salzach	1999-05-22	94	27	**2002-08-12**	**61**	**71**
	2002-08-07	**91**	**44**	2002-06-07	52	20
	2002-08-12	**71**	**61**	2002-03-20	50	22
	2000-03-18	68	40	2001-06-11	45	29
	2000-03-17	55	27	**2002-08-07**	**44**	**91**
08 Tennengau	**2002-08-12**	**97**	**77**	**2002-08-12**	**77**	**97**
	2002-08-07	**88**	**70**	**2002-08-07**	**70**	**88**
	1999-05-22	80	14	2002-03-20	48	45
	2000-03-18	63	42	2002-03-21	45	50
	2000-05-13	54	26	2001-12-07	44	44

Area	Strongest predicted cases			Strongest observed cases		
	Date	ALADIN	TAWES	Date	TAWES	ALADIN
09 Mühl-/Waldviertel	**2002-08-13**	**56**	**65**	**2002-08-08**	**84**	**32**
	2002-06-07	49	16	**2002-08-13**	**65**	**56**
	2002-08-06	39	12	**2002-08-07**	**61**	**27**
	2001-07-21	33	33	**2002-08-12**	**41**	**31**
	2002-08-08	**32**	**84**	2000-09-17	38	7
10 Traun	1999-07-11	60	25	**2002-08-08**	**79**	**22**
	2002-08-13	**58**	**63**	**2002-08-13**	**63**	**58**
	2000-08-06	56	3	**2002-08-12**	**54**	**47**
	2002-08-07	**49**	**49**	**2002-08-07**	**49**	**49**
	2002-03-21	49	42	2002-03-21	42	49
11 Salzkammer-gut	2002-03-21	89	68	**2002-08-12**	**71**	**84**
	2002-08-12	**84**	**71**	2002-03-21	68	89
	2000-03-18	80	43	2001-07-21	57	56
	2002-08-13	**66**	**57**	**2002-08-13**	**57**	**66**
	2002-03-20	63	41	2002-06-08	54	36
12 Enns	**2002-08-07**	**68**	**29**	**2002-08-13**	**76**	**42**
	2000-08-06	**64**	**14**	**2002-08-12**	**65**	**42**
	1999-07-11	44	16	2002-03-21	60	40
	2002-08-12	**42**	**65**	2001-07-21	47	38
	2002-08-13	**42**	**76**	1999-07-23	46	12
13 Traisen	**2002-08-13**	**62**	**84**	**2002-08-08**	**65**	**56**
	1999-07-11	58	24	2002-03-21	59	41
	2002-08-08	**56**	**65**	**2002-08-07**	**46**	**39**
	2002-06-07	44	27	2001-07-21	43	21
	1999-07-10	44	32	2002-06-08	39	33
14 Mur/Mürz	**2002-08-07**	**67**	**15**	**2002-08-12**	**57**	**45**
	2000-10-02	64	13	1999-07-23	51	22
	2000-08-06	62	28	1999-08-17	38	14
	2002-08-12	**45**	**57**	2001-09-15	31	29
	2002-06-07	39	31	2002-06-07	31	39
15 Osttirol	1999-09-21	71	37	2002-06-07	68	54
	2000-11-17	69	20	2001-01-08	42	24
	1999-04-17	57	26	2000-11-07	41	51
	2002-05-26	57	16	1999-07-23	40	25
	2002-06-07	54	68	2000-10-08	40	26
16 Mittelkärnten	**2002-08-12**	**89**	**61**	2000-11-07	56	55
	2000-08-06	66	32	2001-01-08	49	16
	2000-11-04	55	46	2000-11-01	49	39
	1999-09-21	55	48	1999-09-21	48	55
	2000-11-07	55	56	2000-10-02	46	23
17 Unterkärnten	2000-10-02	98	22	**2002-08-12**	**52**	**41**
	2002-08-07	**75**	**25**	1999-07-23	51	40
	2000-08-06	61	34	2001-09-15	47	32
	2002-07-16	50	27	1999-08-17	37	12
	2002-06-29	44	16	1999-07-11	36	19

Area	Strongest predicted cases			Strongest observed cases		
	Date	ALADIN	TAWES	Date	TAWES	ALADIN
18 Gail	**2002-08-07**	**76**	**30**	2000-11-07	77	41
	2000-06-25	74	34	1999-09-21	71	43
	1999-07-23	52	57	2001-01-08	67	18
	2000-11-17	49	34	2000-11-01	67	26
	2000-11-04	49	61	2001-09-15	63	37
19 Karawanken	1999-05-22	50	44	2001-09-15	64	46
	2002-03-21	38	22	**2002-08-12**	**54**	**32**
	2002-08-12	**36**	**38**	2001-09-26	43	40
	2000-03-18	34	16	2002-07-15	40	14
	2001-04-22	31	7	2000-11-07	37	11
20 Weinviertel	**2002-08-08**	**39**	**21**	**2002-08-13**	**36**	**26**
	1999-06-13	36	1	1999-05-21	34	5
	2002-08-13	**26**	**36**	**2002-08-14**	**31**	**22**
	1999-07-09	24	11	2000-08-06	31	7
	1999-07-10	23	17	2000-09-17	29	16
21 Wr. Becken Nord	**2002-08-08**	**55**	**12**	1999-05-21	50	4
	1999-07-10	52	26	**2002-08-13**	**42**	**20**
	1999-07-11	41	19	2000-08-06	35	14
	2001-09-18	39	27	2002-03-21	34	6
	2002-06-07	28	21	**2002-08-12**	**34**	**9**
22 Neusiedler See	**2002-08-08**	**54**	**7**	1999-07-10	36	41
	1999-07-10	41	36	1999-11-10	33	30
	1999-07-11	39	13	1999-06-22	30	17
	2001-09-18	32	22	1999-08-17	29	6
	1999-06-13	30	0	2000-09-17	29	14
23 Wr. Becken Süd	1999-07-11	83	32	1999-05-21	57	3
	2002-06-07	55	37	2002-03-21	56	6
	1999-07-10	47	28	2000-08-06	45	28
	2001-09-15	35	24	1999-07-23	39	21
	2002-08-07	**32**	**20**	**2002-08-12**	**38**	**22**
24 Raab	2002-06-07	58	22	**2002-08-12**	**47**	**23**
	2002-08-07	**51**	**13**	1999-05-21	46	7
	2000-10-02	45	17	2000-08-06	43	27
	1999-07-11	35	38	1999-07-11	38	35
	1999-05-05	34	15	1999-08-17	38	15
25 Südwest-Steiermark	2000-10-02	78	22	**2002-08-12**	**58**	**27**
	1999-07-10	63	33	1999-07-11	47	32
	2002-08-07	**56**	**11**	1999-05-21	43	18
	1999-05-05	46	14	2000-08-06	41	31
	2002-06-29	38	12	1999-07-23	41	36
26 Vorarlberg	1999-05-22	82	119	1999-05-22	119	82
	2001-12-30	82	31	**2002-08-12**	**75**	**71**
	2002-03-22	80	16	1999-05-12	56	77
	2000-01-31	80	21	2000-08-07	54	44
	2001-02-23	78	34	2000-09-21	52	38

14-4.2 Error statistics

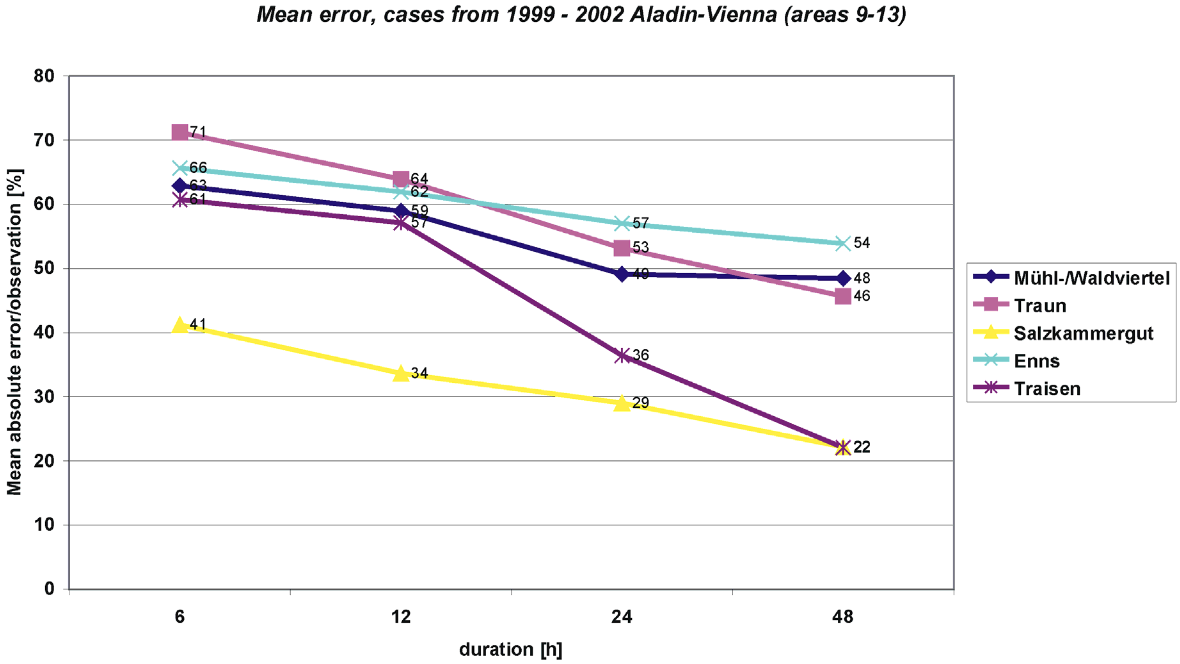


Fig. 14-4.2: Mean absolute error in percent of precipitation forecast of Aladin (cases from 1999 - 2002) depending on area and forecast duration.

Studying the performance of ALADIN from 1999 to 2002, five cases with highest mean precipitation amounts (for both observed and predicted) are chosen for each area in order to use a representative number of heavy rainfall events (compare the table of highest precipitation totals above). In Figure 14-4.2 areas 9 to 13 are picked again to point out similarities and/or differences of predicted precipitation amounts to the August 2002 flood event. For the region "Salzkammergut" the model's quality turns out to behave quite similar to the latter during August 2002 (compare with Figure 14-3.5). The mean absolute error is about 40% for a period of 6 hours and decreases almost linearly to 22% for durations of 48 hours. Forecasts for the regions Mühl-/Waldviertel, Traun, and Enns show an improvement of 10 – 20 % if the duration is extended from 6 hours to 48 hours. The behaviour of area Traisen seems to be an exception, as the reduction of mean absolute error is almost 40%. This corresponds to the model output in case of the flood event 2002, where a decrease of at least 30% was observed. Differences occur for the region "Enns", the mean error is only reduced by about 12%, whereas for August 2002 an improvement of almost 25% was reached. The Enns-valley is situated partly within the Alps and therefore not as "classical" a region affected by blocking effects as the rim area "Salzkammergut" (Haiden et al., 1992), where the forecast quality is higher. Predicted precipitation amounts for lowland regions (Mühl-/Waldviertel and Traun) show similar results, but both of them diverge from the August event (by 10 – 20%).

We also look into the error distribution for different forecast durations (4 categories) as defined in Section 14-3. Comparing Figure 14-4.3 to Figure 14-3.9, differences between August 2002 model results and historical ones become evident. Although a large number of cases show errors below 3 mm / 6 h, the frequency of errors between 3 and 10 mm is higher than for August 2002. Again, the effect of compensation due to longer periods is especially observed for errors larger than 10 mm.

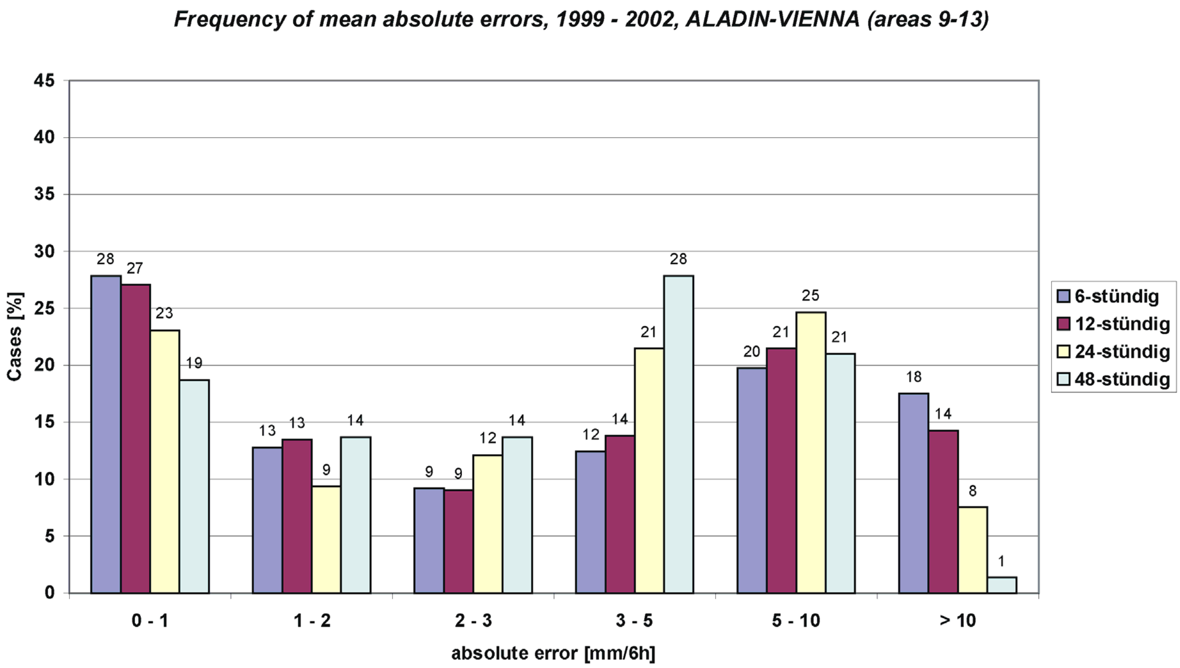


Fig. 14-4.3: Frequency distribution of mean absolute errors / 6 hours , divided in 6 intensity categories, for four forecast durations (6, 12, 24, 48 hours). Cases are chosen from 1999 – 2002.

14-4.3 Summary of ALADIN forecast quality during 1999-2002

The verification of precipitation amounts for cases from 1999 to 2002 confirms the fact that the model is only partly able to forecast extreme situations. The time of onset and of termination of precipitation episodes is generally simulated reasonably well. Hourly peaks are usually not reproduced. Intra-epsiode variations of intensity on the 3-6 hour timescale are sometimes, but not consistently, reproduced. The forecast is generally better in regions at the alpine rim, where orographic blocking effects dominate. In lowland regions convective systems are the cause of the the majority of heavy precipitation events. These systems are much more difficult to deal with in a numerical forecast. Consequently mean errors are higher and only marginally reduced by increasing duration. Furthermore, for convective events, which rarely last longer than a few hours, an extension of duration to 24 or 48 h makes little sense.

Comparing the verifications for both periods, it turns out that mean errors during the August 2002 event are somewhat smaller than the average of the other events out of the 3 years. An exception is the region Salzkammergut, where the error of the model remains rather constant. Summing up the most intense cases during the last 3 years, errors are typically in the range 50-70% for 6 hour duration, and 20–50% for 48 hour duration.

14-5 Concluding remarks

A detailed verification of meteorological forecast models with regard to the August 2002 flood event is a necessary requirement for the development of effective warn systems. This report shows quantitatively that the forecast skill strongly depends on the temporal and spatial scale, as well as on the observational data used, and the area under consideration. In general, forecasts for alpine areas affected by orographic upslope precipitation are more reliable than those for lowland reagions because in the latter convective processes make a largeer contribution to heavy precipitation events. A significant reduction of the relative forecast error can be achieved by increasing the duration for which a forecast is made. This is because forecast errors partially compensate within the duration of an event. A reduction of the relative forecast error through increasing area size can be achieved only when one approaches the typical scale of a province. It is not just the size of the catchment area but also the size of the synoptic disturbance itself that determines forecast skill. Compared to other events of the last 4 years, precipitation amounts during the first part of the August 2002 flood were forecasted poorly, whereas during the second part forecasts were rather better than average. Hourly maxima are still hardly predicted, and generally underestimated. Probability forecasts based on ensemble predictions can contribute to improved pre-warnings (or ‘watches’) in the sense that they give the possible spectrum of precipitation szenarios.

List of Figures and Tables

References

Andrade-Leal, R. N., M. Bachhiesl, U. Drabek, D. Gutknecht, T. Haiden, H. Holzmann, K. Hebenstreit, R. Kirnbauer, H. P. Nachtnebel and J. Precht, 2002: Hydrologische Vorhersagemodelle im operationellen Betrieb der Wasserkraftwirtschaft. *Österr. Wasser- u. Abfallwirtschaft,* **54,** 129-134.

Borga, M., E. N. Anagnostou and E. Frank, 2000: On the use of real-time radar rainfall estimates for flood prediction in mountainous basins. *J. Geophys. Res.,* **105,** 2269-2280.

Buzzi, A., M. D'Isidoro, S. Davolio and P. Malguzzi, 2003: Numerical assessment of MAP episodes of heavy precipitation using high resolution reanalyses and assimilation of surface data. *Preprints, ICAM and MAP Meeting,* Brig, Switzerland, 24-27.

Habersack, H., and A. Moser, 2003: Ereignisdokumentation Hochwasser August 2002. ZENAR / Plattform Hochwasser, 184p.

Haiden, T., 1994: Eine optimierte Starkniederschlagsauswertung IV: Niederschlagsinterpolation unter Berücksichtigung orographischer Effekte. *Mitt. Hydrogr. Dienst in Österr.,* **72,** 47-62.

Haiden, T., 2003: On the performance of ALADIN during the August 2002 floods. *ALADIN Newsletter, 23,* 191-193.

Haiden, T., M. Kerschbaum, P. Kahlig and F. Nobilis, 1992: A refined model of the influence of orography on the mesoscale distribution of extreme precipitation. *Hydrol. Sci. J.,* **37,** 417-427.

Haiden, T. and R. Schultheis, 1995: Verfahren zur Abschätzung der Auswirkungen von Klimaänderungen auf den Wasserhaushalt von Einzugsgebieten. *Mitt. Hydrogr. Dienst in Österr.,* **73,** 21-38.

Haiden, T., H. Seidl, G. Hermann and G. Skoda, 1997: Das Starkniederschlags-Ereignis 4.-8. Juli 1997 aus prognostischer Sicht. *ÖGM-bulletin* 97/2, 1-13.

Haiden, T. and K. Stadlbacher, 2002: Quantitative Prognose des Flächenniederschlags. *Österr. Wasser- u. Abfallwirtschaft,* **54,** 135-141.

Nobilis, F., T. Haiden and M. Kerschbaum, 1991: Statistical considerations concerning probable maximum precipitation (PMP) in the Alpine country of Austria. *Theor. Appl. Climatol.,* **44,** 89-94.