MAP D-PHASE

Real-Time Demonstration of Weather Forecast Quality in the Alpine Region

BY MATHIAS W. ROTACH, PAOLO AMBROSETTI, FELIX AMENT, CHRISTOF APPENZELLER, MARCO ARPAGAUS, HANS-STEFAN BAUER, ANDREAS BEHRENDT, FRANÇOIS BOUTTIER, ANDREA BUZZI, MATTEO CORAZZA, SILVIO DAVOLIO, MICHAEL DENHARD, MANFRED DORNINGER, LIONEL FONTANNAZ, JACQUELINE FRICK, FELIX FUNDEL, URS GERMANN, THERESA GORGAS, CHRISTOPH HEGG, ALESSANDRO HERING, CHRISTIAN KEIL, MARK A. LINIGER, CHIARA MARSIGLI, RON MCTAGGART-COWAN, ANDREA MONTAINI, KEN MYLNE, ROBERTO RANZI, EVELYNE RICHARD, ANDREA ROSSA, DANIEL SANTOS-MUÑOZ, CHRISTOPH SCHÄR, YANN SEITY, MICHAEL STAUDINGER, MARCO STOLL, HANS VOLKERT, ANDRE WALSER, YONG WANG, JOHANNES WERHAHN, VOLKER WULFMEYER, AND MASSIMILIANO ZAPPA

A six-month project successfully tested real-time, end-to-end multimodel hydrometeorological forecasts for heavy precipitation and related flooding events in many different catchments in the Alps.

As the first research and development project (RDP) of the World Weather Research Programme (WWRP), the Mesoscale Alpine Programme (MAP) made important contributions to our knowledge on atmospheric processes determined by and influencing weather in mountainous terrain between 1994 and 2005 (Bougeault et al. 2001). A wealth of scientific results (Volkert and Gutermann 2007) was produced in research areas ranging from atmospheric dynamics to mountain hydrology. On the basis of these very positive results, the WWRP Scientific Steering Committee asked the MAP community to consider a forecast demonstration project (FDP). Generally, a FDP aims at demonstrating the advances a research and development activity (or any other trigger) has brought to operational atmospheric forecasting. Thus, a FDP ▶

Flooding in the small town of Laufen, Switzerland on 23 August 2007. Photograph courtesy of the State Police corps.

- deals with the forecast of weather with international relevance (high-impact weather);
- demonstrates a clear advance in forecasting capability;
- provides clear evaluation protocols; and
- is characterized by an expectation of success.

The first FDP was realized on the occasion of the Sydney Olympic Games in 2000 (Keenan et al. 2003). The MAP Steering Committee mandated a working group to investigate the prospects for a MAP FDP and supported the group's proposal to focus on the precipitation-related aspects (so-called wet MAP). The project acronym D-PHASE carries the double meaning of Demonstration of Probabilistic Hydrological and Atmospheric Simulation of Flood Events in the alpine region as well as the fourth phase of MAP. The Scientific Steering Committee of WWRP endorsed D-PHASE as an FDP in its meeting in October 2005.

The present paper gives an overview of the goals of D-PHASE and presents the project's central facility, the "visualization platform." Examples of applications are given as well as a summary of user feedback and an overall summary and outlook.

OVERVIEW ON MAP D-PHASE. Improvement of high-resolution numerical modeling was one of MAP's most successful achievements. For the first time in a project of comparable size, a high-resolution (3-km grid size) mesoscale model was used in the mission planning process during the intensive observation period (IOP; Benoit et al. 2002, 2003), and high-resolution numerical modeling was successfully used for different case studies in relation to orographic precipitation (Richard et al. 2007) or other processes (e.g., Weigel et al. 2006; Rotach and Zardi 2007). MAP's hydrological community pioneered the operational coupling of deterministic atmospheric and hydrological models (Ranzi et al. 2003) and investigated the performance of coupled systems (Bacchi and Ranzi 2003).

Radar observation of precipitation in complex terrain is extremely challenging (requiring correction of clutter and shading due to topography). MAP has triggered a substantial improvement in the performance of operational radar products (Germann et al. 2006). Finally, MAP has also substantially advanced our understanding of the processes related to orographic precipitation (Rotunno and Houze 2007).

At the outset of MAP (Binder and Schär 1996), probabilistic modeling of atmospheric processes had neither explicitly been identified as a research topic, nor had follow-on hydrological ensemble modeling. Still, MAP has triggered a number of studies investigating the predictability of orographically influenced precipitation (Walser and Schär 2004; Walser et al. 2004a; Hohenegger et al. 2006; Hohenegger and Schär 2007). Also, in the aftermath of MAP, a high-resolution ensemble prediction system [the Limited Area Ensemble Prediction System, based on the Consortium for Small Scale Modeling (COSMO) model] has been developed (Molteni et al. 2001; Marsigli et al. 2001; Walser et al. 2004b) and used for first steps into hydrological ensemble prediction (Siccardi et al. 2005; Verbunt et al. 2007).

On the basis of these findings, and bearing in mind that orographic precipitation has often led to disastrous flooding events widespread over the Alps, it was decided to devote the MAP FDP to the demonstration of forecast capability with respect to heavy precipitation events in the Alps. The emphasis

AFFILIATIONS: Rotach, Ambrosetti, Appenzeller, Arpagaus, Fontannaz, Fundel, Germann, Hering, Liniger, Stoll, and Walser—Federal Office for Meteorology and Climatology (MeteoSwiss), Zurich, Switzerland; Ament—University of Hamburg, Hamburg, Germany; Bauer, Behrendt, and Wulfmeyer—University of Hohenheim, Stuttgart, Germany; Bouttier and Seity—Météo-France, Toulouse, France; Buzzi and Davolio—ISAC-CNR, Bologna, Italy; Corazza—ARPAL CRMI-PC, Genova, Italy; Denhard—German Weather Service (DWD), Offenbach, Germany; Dorninger and Gorgas—University of Vienna, Vienna, Austria; Frick, Hegg, and Zappa—Swiss Federal Institute WSL, Birmensdorf, Switzerland; Keil and Volkert—DLR, Oberpfaffenhofen, Germany; Marsigli and Montaini—ARPA-SIM, Bologna, Italy; McTaggart-Cowan—Canadian Meteorological Centre, Environment Canada, Dorval, Quebec, Canada; Mylne—Met Office, Exeter, United Kingdom; Ranzi—University of Brescia, Italy; Richard—University Paul Sabatier, Toulouse, France; Rossa—ARPAV, Centro Meteorologico Teolo, Italy; Santos-Muñoz—AEMet, Madrid, Spain; Schär—ETHZ, Zurich, Switzerland; Staudinger and Wang—Central Institute for Meteorology and Geodynamics, Vienna, Austria; Werhahn—Institute for Meteorology and Climate Research–Atmospheric Environmental Research (IMK-IFU), Garmisch-Partenkirchen, Germany
A supplement to this article is available online (10.1175/2009BAMS2776.2)
CORRESPONDING AUTHOR: Dr. Mathias W. Rotach, Federal Office of Meteorology and Climatology, Kraehbuehlstrasse 58, Zurich CH-8044, Switzerland.
E-mail: mathias.rotach@meteoswiss.ch

The abstract for this article can be found in this issue, following the table of contents.
DOI:10.1175/2009BAMS2776.1

In final form 5 March 2009

was put on high-resolution operational modeling, be it probabilistic or deterministic.

D-PHASE was set up as an end-to-end flood warning system as it is sketched in Fig. 1. "End to end" in this context means that the entire chain from the atmospheric forecast models to the decision maker as the end user is part of the system. Some 5 days before a possible event, atmospheric ensemble prediction systems may issue a "prealert"—that is, indicating that a threshold might be exceeded in a certain region in the Alps. At this stage, thresholds are primarily applied to precipitation, although the first hydrological models start to determine forecasts for runoff at various stations. As time comes closer to the possible event, high-resolution deterministic atmospheric models with a lead time of 18–36 h start, and so do the corresponding hydrological models. At all levels of the *visualization platform* (VP; see the next section) colored warnings are displayed, from which users may see immediately whether their region is in danger. At the time of the forecasted event, users also have access to nowcasting facilities to judge the "present" situation and to come to the most beneficial decision.

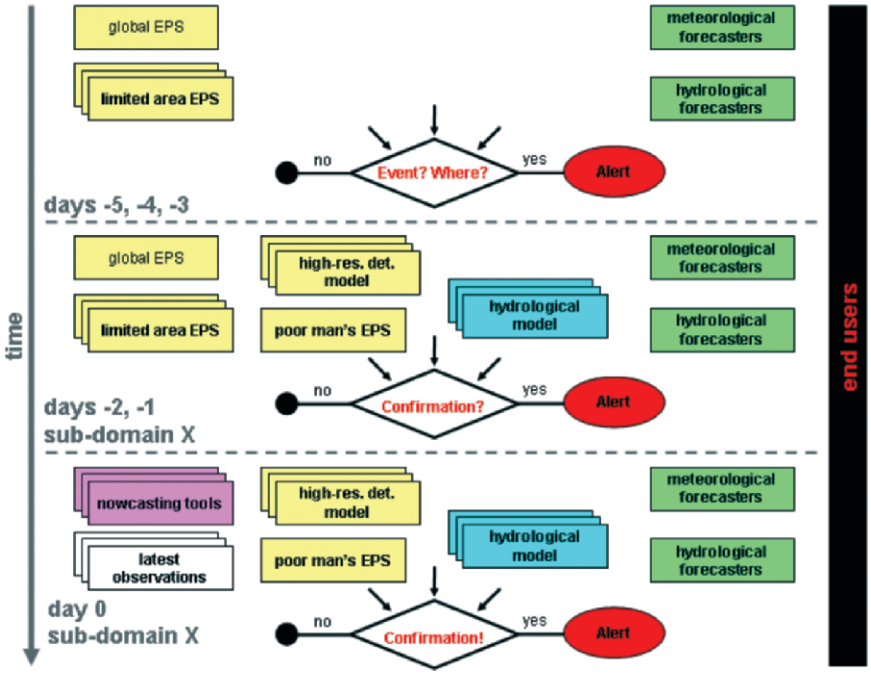

Fig. 1. Schematic representation of the various components of the D-PHASE end-to-end flood forecasting system. (top)–(bottom) Note the time running. The green, blue, purple, and white boxes denote tools that are employed within the indicated periods of time. The open diamond at the end of each period symbolizes the question that can be answered (such as "is there a possible event ahead?"). If no, then the alert cycle comes to an end; if yes, then the information is passed to the forecasters (green boxes) and/or the end users (black rectangle).

Participating atmospheric models (Table 1) include many of the high-resolution (i.e., a few kilometers grid size) deterministic operational models that are presently being developed in Europe as well as their lower-resolution driving models. In addition, a collection of ensemble prediction systems at intermediate resolution is on the list; for example, a poor man's ensemble prediction system (EPS; Micro-PEPS), like that of the short-range numerical weather prediction project (SRNWP; Denhard and Trepte 2006) that has been constructed from the participating high-resolution models especially for D-PHASE (refer to http://dx.doi.org/10.1175/2009BAMS2776.2).

The collection of hydrological models (Table 2) includes both deterministic and ensemble prediction systems (Zappa et al. 2008). The latter systems constitute advancement in the *conceptual treatment* of hydrological forecasts, and the D-PHASE operations period was an excellent opportunity to demonstrate whether they also constitute advancement in the *quality* of hydrological forecasts.

All nowcasting products on the VP are summarized in Table 3. Some of the products are offered on the entire "D-PHASE domain" (Fig. 2). Additionally, some operational institutions in the Alpine region (MeteoSwiss, Météo-France, and the regional meteorological service of Emilia-Romagna, Italy) offered their radar and nowcasting tools specifically designed for certain regions and applications.

An important group of project participants are the end users; that is, those people who use information on the VP for their decisions or for further elaboration of data. Different from MAP when "target areas" had been specified beforehand according to scientific criteria (Bougeault et al. 2001), the presence of an interested end user and his/her liaison with a hydrological modeler defined a "participating catchment" for D-PHASE. In this spirit hydrological forecasts were produced for a total of 43 catchments. End users

Table 1. Atmospheric models participating in D-PHASE (real time only; limited-area ensemble prediction systems, and high-resolution deterministic models) with model specifications (computational domain).

Model name on data archive	Model*	Ensemble size (1 = deterministic)	Mesh size [km (degrees)]	Forecast range (h)	Institution and contact person(s)
CLEPS	COSMO-LEPS	16	10 (0.09)	132	ARPA Emilia-Romagna, Andrea Montani
MOGREPS	MOGREPS	24	25 (0.22)	54	Met Office, Kelvyn Robertson
INMSREPS	SREPS	20	(0.25)	72	AEMET, Daniel Santos
CSREPS	COSMO-SREPS	16	10 (0.09)	72	ARPA Emilia-Romagna, Chiara Marsigli
LAMEPSAT	ALADIN-LAEF	17	(0.15)	60	ZAMG, Sabine Leroch
PEPS	PEPS	variable	(0.0625)	42	EUMETNET SRNWP, Michael Denhard
MPEPS	Micro-PEPS	variable	(0.02)	18	DWD, Michael Denhard
COSMOCH7	COSMO-7	1	7 (0.0625)	72	MeteoSwiss, Felix Ament
COSMOCH2	COSMO-2	1	2.2 (0.02)	24	
LMEURO	COSMO-ME	1	7 (0.0625)	72	CNMCA, Lucio Torrisi
LMITA	COSMO-IT	1	2.8 (0.025)	30	
LAMI7	COSMO-I7	1	7 (0.0625)	72	ARPA Emilia-Romagna, Davide Cesari
LAMI28	COSMO-I2	1	2.8 (0.025)	48	
LME	COSMO-EU	1	7 (0.0625)	78	DWD, Michael Denhard
LMK	COSMO-DE	1	2.8 (0.025)	21	
ISACMOL	MOLOCH	1 (BC from GFS)	2.2 (0.02)	39	ISAC-CNR, Silvio Davolio
ISACMOL2	MOLOCH	1 (BC from IFS)	2.2 (0.02)	48	ISAC-CNR, Silvio Davolio
ARPALMOL	MOLOCH	1	2.2 (0.02)	36	ARPA Liguria, Matteo Corazza
QBOLAM33	QBOLAM	1	33 (0.3)	60	APAT, Stefano Mariani
QBOLAM11	QBOLAM	1	11 (0.1)	48	
ALADFR	ALADIN	1	(0.1)	30	Météo-France, Eric Bazile (ALADFR), Yann Seity (AROME)
AROME	AROME	1	(0.04)	30	
MM5_2_CT	MM5	1	(0.02)	24	University of Hohenheim, Hans-Stefan Bauer, Matthias Grzeschik
MM5_2_4D	MM5	1	(0.02)	24	
MM5_60	MM5	1	(0.56)	72	FZK IMK-IFU, Johannes Werhahn, Andreas Marx
MM5_15	MM5	1	(0.14)	72	
MM5_375	MM5	1	(0.04)	60	
ALADAT	ALADIN	1	9.6 (0.07 × 0.11)	48	ZAMG, Yong Wang
CMCGEML	GEM-LAM	1	15 (0.135, 0.188)	24	Environment Canada, Ron McTaggart-Cowan
CMCGEMH	GEM-LAM	1	2.5 (0.0225, 0.0327)	18	

*For the high-resolution deterministic models, the coarser-resolution driving model(s) are also listed (if their output is provided).

as the "customers" of the D-PHASE information were granted free access to all products on the VP for the "price of feedback." One of the goals of the project was to systematically evaluate the user feedback as a subjective measure of performance, contrast this to the objective measures (model skill scores, among others), and make the results available to the community. Preliminary results are presented later.

D-PHASE profited from a successful collaboration with the WWRP RDP Convective and Orographically-induced Precipitation Study (COPS; Wulfmeyer et al. 2008, unpublished manuscript). Having similar objectives, it was decided to pool resources and coordinate efforts for mutual benefit. Therefore, the D-PHASE operations period (DOP) was defined to last from June to November 2007, including the COPS field

Table 2. Hydrological models participating in D-PHASE (real-time only). A hydrological model running in ensemble mode is denoted as e-hm; a hydrological model running deterministically is denoted as d-hm.

Model name on data archive	Applications (No. of catchments)	Driving model	Institution and contact person(s)
PREVAH (e-hm, d-hm)	6 in Switzerland 2 in Italy	CLEPS, COSMOCH2, COSMOCH7, MM5_15	ETHZ/WSL, Simon Jaun, Massimiliano Zappa
HBV (e-hm, d-hm)	12 in Switzerland	CLEPS, PEPS, COSMOCH2, COSMOCH7, (IFS)	BAFU, Stephan Vogt
LAMBRO (e-hm, d-hm)	2 in Italy	CLEPS, LAMI28, ISACMOL	ARPA Lombardia, Matteo Cislaghi
DIMOSOP (e-hm, d-hm)	10 in Italy	CLEPS, ISACMOL, ISACMOL2	University of Brescia, Roberto Ranzi
FEST (e-hm)	2 in Switzerland 1 in Italy	CLEPS	POLIMI, Giovanni Ravazzani
LARSIMBY (d-hm)	10 in Germany	LME, GFS, GME	WWA Kempten, Uwe Ehret
LARSIMBW (d-hm)	2 in Germany	LMK, LME	LUBW, Werner Schulz

phase [June–August (JJA)] as well as the MAP special observing period [September–November (SON)]. An essential accomplishment was the provision of a large set of model variables in a harmonized format [gridded binary edition 1 (GRIB1)] using the so-called TIGGE + table (for more information on TIGGE and the WWRP THORPEX program that defined this table, see www.wmo.int/pages/prog/arep/wwrp/new/thorpex_new.html). This enabled the production of plots of a large set of forecasts with exactly the same domains and colors for the same atmospheric variables. The D-PHASE domain entirely includes the COPS domain (Fig. 2). The COPS mission planning team successfully used the D-PHASE information on the VP (mainly atmospheric forecasts) for their operations between June and August 2007. All of the data, the observations due to COPS, and numerical model results produced by D-PHASE are stored at a joint data archive at the World Data Center for Climate (WDCC), run by the Max Planck Institute for Meteorology Hamburg (Wunram 2006).

THE VISUALIZATION PLATFORM. For the duration of the DOP, all graphical D-PHASE information (forecasts, warnings, nowcasting products) was displayed on the Web-based VP. On three levels [Alpine wide, region (country), catchment] users could choose between the following different options:

- *eye catching warning maps* in "traffic-light colors": as an example a screenshot of level 1 is given in Fig. 3 for the extraordinary event on 8/9 August 2007;
- *more detailed information* on duration, peak amounts, and more, featuring a comparison of all models;
- *harmonized model products* such as plots, cross sections or meteograms based on the TIGGE+ dataset;

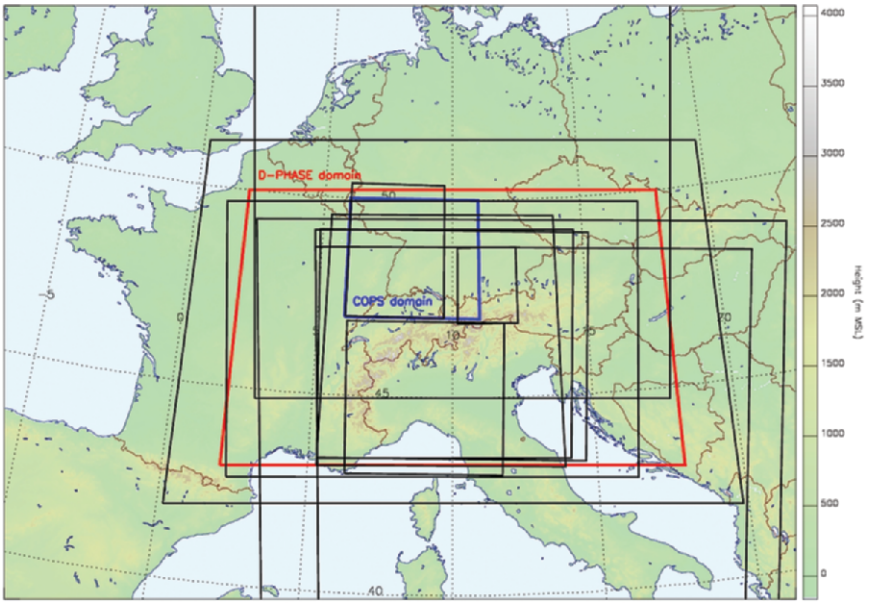

Fig. 2. Map of the Alps (color shading) with the outlines of the model domains of some of the high-resolution atmospheric D-PHASE models of Table 1. The red rhombus depicts the D-PHASE domain, while the blue square shows the COPS domain.

- *nowcasting products* such as radar loops, extrapolated trajectories, and composites; and
- *validation products* [Vienna Enhanced Resolution Analysis (VERA) analyses in particular, discussed later] and difference plots between a specific model and these analyses.

The probably most attractive single aspect of the VP was that all of the warnings were based on the same thresholds and procedures. All of the models agreed on a joint output format and used the same program to determine threshold exceedance (the routine was different for each category of models—deterministic or ensemble, atmospheric or hydrological forecasts). Thus, if a particular region or catchment was red (severest threshold) as forecasted by one particular model and only orange (second highest threshold) by another model, there was a difference in the model results and not in the analysis (how to define the area borders and the threshold, how to sum up, and so on). The philosophy of D-PHASE with respect to warnings was directed toward the highest probability of detection (POD): if only one model exceeded a threshold for a particular catchment, this catchment assumed the corresponding color on the summary plot.

Bearing in mind that D-PHASE was an experimental system (experimental to demonstrate operational fitness), the warning levels (thresholds; WLs) were set rather low:

- WL3 (red): return period of 10 yr,
- WL2 (orange): return period of 180 days, and
- WL1 (yellow): return period of 60 days.

No alert (green) was given when none of the models (neither precipitation nor runoff) exceeded any of the WLs. Warnings from both atmospheric and hydrological EPSs were issued if 33% of the ensemble members exceeded the corresponding WL. Return levels were determined for each region and catchment separately on the basis of statistics of annual maxima of daily precipitation derived from the Frei and Schär (1998) precipitation climatology and scaling assumptions with respect to duration and area.

These rather low warning levels were employed to get at least some "events" during the 6-month DOP. Another boundary condition for defining warning levels was that D-PHASE catchments belong to different countries with their different operational alert levels and systems. A joint (comparable) definition for D-PHASE would have required choosing one of the alert definitions, leaving one privileged country versus all the other countries. Definition using return periods allowed the use of a joint approach with consideration of regional variability. Note that this definition of warning levels is linked with different absolute numbers in different catchments, which is at odds with the current practice in many operational services. In fact, the definition of warning thresholds based on return levels was an issue of major concern among the atmospheric forecasters and also the "end users." This probably best illustrates the inherent difficulty of an (international) FDP: on the one hand it must necessarily be experimental (if the demonstration is positive, then the results may become operational), but it deals with a very serious issue (severe weather), in which the stakeholders need clear procedures and boundary conditions to find optimal solutions in case of an emergency.

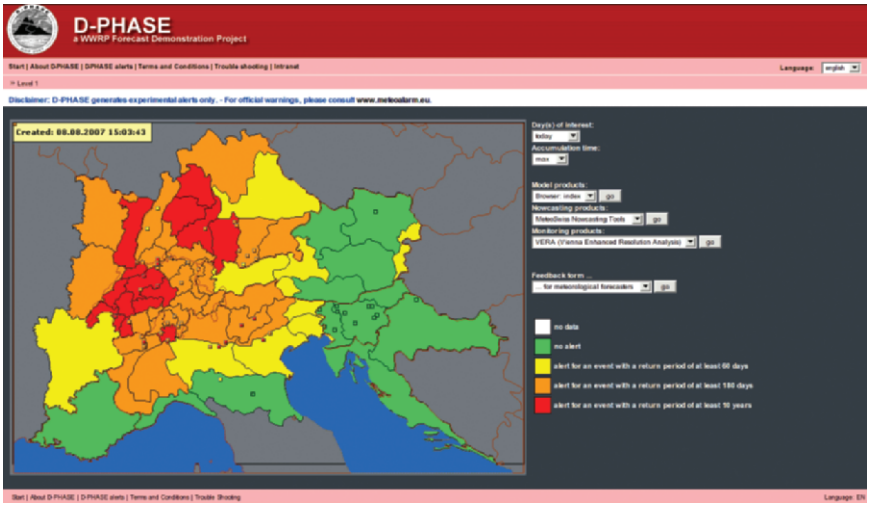

Fig. 3. Screenshot of the D-PHASE VP for 8 Aug 2007, level 1 (alpine-wide view). Blue areas denote the Ligurian (west) and Adriatic (east) Seas, respectively. Green, yellow, orange, and red areas correspond to the defined warning levels (refer to text). Adopted from Zappa et al. (2008).

Beyond the colored warning maps, the platform also featured the corresponding detailed information. For each target region and catchment, the duration and level of warning were graphically displayed for each model in parallel (atmospheric forecasts for target regions and catchments, hydro-

logical forecasts for runoff stations). Even more detail was available by clicking on the desired property. Again, the D-PHASE philosophy with respect to graphical display was to use *identical graphics* for all the different models, produced by the same scripts applied to all the models. Figure 4 highlights this as an example for the 28/29 August 2007 event. The same scales and colors and same spatial representation allow for concentrating on the essential differences in the precipitation fields. Apart from 2D plots, various cross sections or meteograms can be selected. EPSs can display probability of exceedance for a given (selectable) threshold or ensemble averages.

From the VP, users can directly reach the various nowcasting applications through an Internet link. Table 3 lists all of the available information.

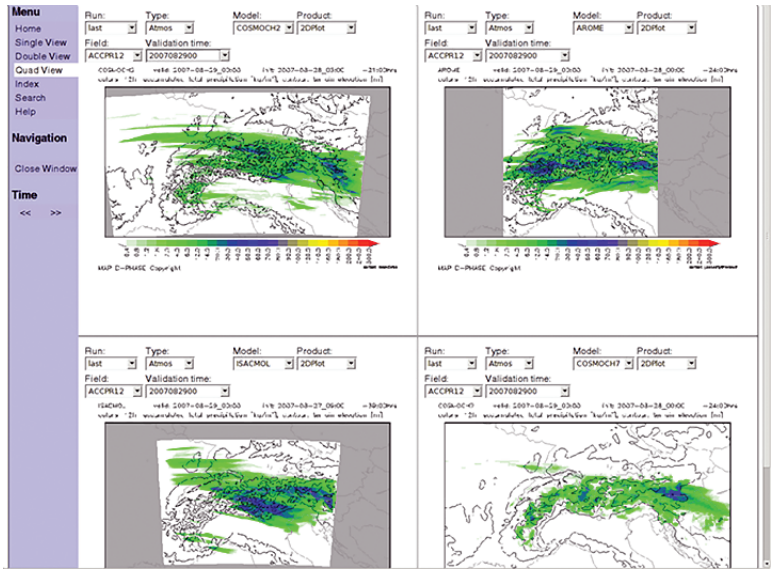

Fig. 4. Screenshot of the graphical display facility on the D-PHASE VP. Example of 12-h accumulated precipitation, valid 0000 UTC 29 Aug 2007, from (top left) COSMO-2, (top right) Application de la Rechereche à l'Operationnel pour la Méso-Echelle (Arome), (bottom left) MOLOCH, and (bottom right) COSMO-7. See Table 1 for model names.

EXAMPLES OF APPLICATIONS. In this section a selection of applications and products, which were designed or tested during the DOP, is presented. Four selected examples are discussed here in some detail. In the online supplement (http://dx.doi.org/10.1175/2009BAMS2776.2), three additional applications (studies) of the D-PHASE datasets and concepts are presented. The first study shows that the COSMO-LEPS 24-h precipitation forecast can be substantially improved by calibration based on reforecasts (Fundel et al. 2009). The second study addresses the operational assimilation of GPS data within the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5) four-dimensional variational data assimilation (4DVAR) system (Zus et al. 2008). The third study explores the potential of a high-resolution poor man's ensemble (Micro-PEPS). Additional applications can be found in Arpagaus et al. (2009) and Zappa et al. (2008) for applications from D-PHASE of a more hydrological nature.

Verification of precipitation alerts using operational data. The D-PHASE dataset offers ample possibilities to objectively verify model data against observations. One of the first such efforts was performed using the Swiss radar composite for JJA (Ament and Arpagaus 2008). Results have already been valuable in assessing model deficiencies and have even led to the detection of a major bug in one of the participating models. Here, we concentrate on warnings—the core product of D-PHASE—issued by all models in the same format during the DOP. The present evaluation is based on some 18 target regions in Switzerland because of high-quality data, but it will be performed for various model parameters and the entire D-PHASE domain once the observational data are ready and available. Reference alerts for the evaluation are based on a combination of radar (Germann et al. 2006) and rain gauge observations of daily precipitation sums interpolated onto a 1-km grid according to Frei et al. (2006) and Frei and Schär (1998). Radar data are spatially and temporally averaged to obtain hourly time series for each target region. Then, systematic errors in radar observations are corrected by a daily, multiplicative calibration to enforce an exact match with corresponding daily sums of the gauge analysis.

Here, we focus on the short-range time scale, which is well covered by the convection-permitting models. Consequently, short accumulation periods (3, 6, and 12 h) and the most recent forecasts are considered (the first three forecast hours are discarded to account for production time). To avoid double penalty effects, both model and observations are aggregated on 6-h intervals by analyzing whether there was any alert within each interval.

Table 3. Nowcasting and online monitoring tools participating in D-PHASE. Coverage refers to the main region. Because of the extension of the underlying observing systems (e.g., radar), some neighboring regions are also covered but possibly with a reduced data quality.

Name	Purpose	Coverage	Institution and contact person(s)
MeteoSwiss NASS	Quantitative precipitation estimate based on radar	Switzerland plus boarder area	MeteoSwiss, Urs Germann
Piemonte-MeteoSwiss composite	Quantitative precipitation estimate based on radar	Piemonte and Switzerland	ARPA Piemonte/MeteoSwiss, Roberto Cremonini/ Urs Germann
MeteoSwiss ENASS	Ensemble quantitative precipitation estimate based on radar	Switzerland plus boarder area	MeteoSwiss, Urs Germann
MeteoSwiss TRT	Thunderstorms Radar Tracking	Switzerland plus boarder area	MeteoSwiss, Alessandro Hering
DLR Cb-TRAM	Tracking and monitoring severe convection using multichannel *Meteosat-8* SEVIRI data	D-PHASE and COPS domain	DLR, Arnold Tafferner
ARPA-SIM radar products		(northern) Italy	ARPA-Emilia Romagna, Andrea Montani
Météo-France radar products		France	Météo-France, Philippe Frayssinet
VERA	Analysis of surface fields for online monitoring	D-PHASE and COPS domain	University of Vienna, Manfred Dorninger
NWP minus VERA	Online monitoring for some of the NWP models; surface fields	D-PHASE and COPS domain	University of Vienna, Theresa Gorgas
CLEPS versus satellite	Online monitoring of CLEPS versus satellite observations	D-PHASE and COPS domain	DLR, Christian Keil

This temporally coarsened information still satisfies the needs of most users.

Model performance is quantified by the "relative value" (Richardson 2000), which varies between 0 (no skill) and 1 (perfect model) and indicates what fraction of economic savings can be achieved relative to the maximum possible (costs for "no forecast" minus costs for "perfect forecast"). The relative value strongly depends on the ratio of costs C for protective actions to losses L in case of no protection. The cost–loss ratio C/L reflects the user's sensitivity against the two types of erroneous forecasts: missed events are critical for low C/L (e.g., alerting local rescue organization) and false alarms for high C/L (e.g., evacuating an entire town).

The validation of all deterministic models (Fig. 5a) with hourly resolution and covering Switzerland indicates that today's models have a positive economic value over a wide range of C/L ratios. The new high-resolution, convection-permitting models are beneficial in this respect. Further analysis will reveal whether this advantage can mainly be attributed to better representation of convection, improved orographic forcing, or faster reinitialization.

All model alerts can be combined to yield a multimodel alert system (MMAS) by assuming that the forecasted probability of an event equals the fraction of models issuing an alert. The decision alert "yes" or "no" can be optimized to the user's C/L by defining an appropriate number of models that are required "yes" to issue an alert (Fig. 5b). Because of this optimization, a MMAS can cover a wider range of C/L ratios than a deterministic forecast; for the present case it reaches higher relative values, especially for low C/L. Similar improvement can be achieved by varying the alert threshold used by a single deterministic forecast (i.e., the best model in the present analysis); for this, predicted precipitation amounts are multiplied by a constant "factor," which is equivalent to varying the warning level. Note that the factor used to obtain

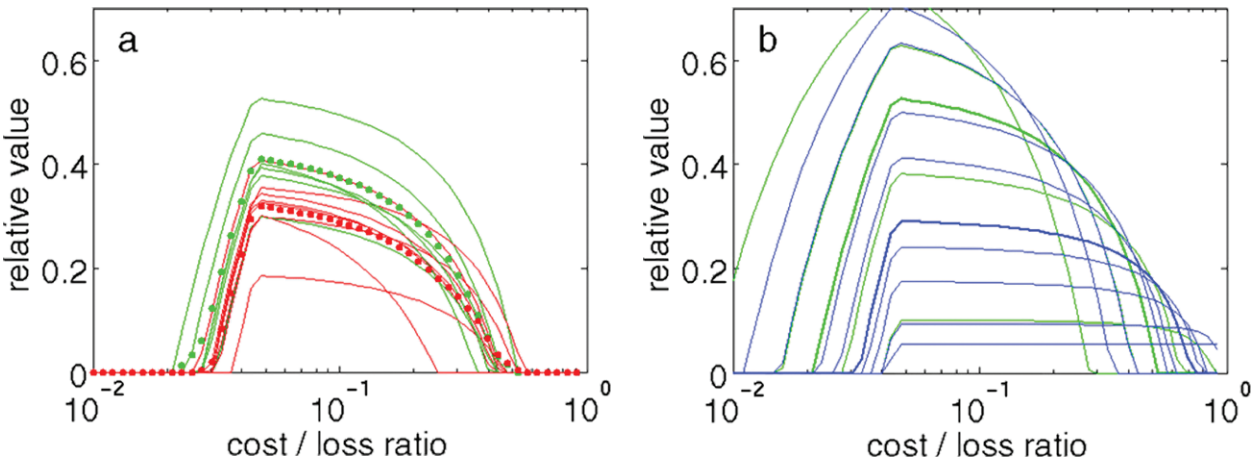

Fig. 5. Relative value depending on *C*/*L* ratio for alert level yellow (return period of six events per year) for 18 Swiss warning regions during summer 2007. (a) Deterministic models with parameterized deep convection (red) and high-resolution models resolving deep convection (green) with dots indicating the mean relative values, respectively. (b) Performance of an uncalibrated MMAS (blue lines) made up from all model alerts displayed in (a) with varying probability threshold to issue an alert (10%–90%) in contrast to statically calibrated forecasts (factors of 0.5, 0.8, 1.0, 1.25, and 2.0; green lines) of the best model in (a). Thick (green and blue) lines indicate the result for percentage of models that issue an alert of 50% and a calibration factor 1.

the present result is rather substantial and does not reflect the true model uncertainty. These results indicate that it is worthwhile to consider multimodel forecasts. However, multimodels are not the only possible approach; other methods such as model recalibration might be competitive, at least in terms of computing resources. It is ongoing work to extend the present analysis to the whole D-PHASE domain and period and to include the participating single-model ensemble systems.

The COSMO short-range ensemble prediction system. The COSMO-SREPS (Marsigli et al. 2006) has taken advantage of the DOP as an extensive test bed, over which to perform a robust statistical evaluation of its behavior and address the issue of the role of applied perturbations in determining its spread and skill. In COSMO-SREPS, initial and boundary conditions are taken from the short-range EPS of the Spanish meteorological office (AEMET), denoted as INM-SREPS in Table 1). This ensemble combines four different operational global models [Integrated Forecast System (IFS), Global Model Europe (GME), Global Forecast System (GFS), and Unified Model (UM)] with four limited-area models to yield a 16-member multimodel EPS (Garcia-Moya et al. 2009). For COSMO-SREPS the four runs of the COSMO limited-area model (25-km resolution), nested on the four different global models, are employed. Then, four COSMO runs at 10-km resolution are nested on each of them, differentiated by adopting four different configurations of the physics parameterization, resulting in a 16-member with combined perturbations from the initial conditions and model physics.

As an example of this application, some results obtained for the first months of the DOP are shown. COSMO-SREPS was run daily at 0000 UTC during this period and integrated over 72 h. Observations are taken from a high-resolution dataset encompassing about 1,400 stations in northern Italy and Switzerland.

In Fig. 6, the relative operating characteristics (ROC) area values for 24-h accumulated precipitation (6–30-h forecast range) are shown as a function of the precipitation threshold. The ROC is a measure of the likelihood that probability forecasts for an event are higher for occurrences than for nonoccurrences of the event (Casati et al. 2008). If the ensemble system predicts all events (probability of detection going to one) at zero false alarm rate then the ROC area is one. Random guessing yields a ROC area of 0.5. The light-blue lines in Fig. 6 represent the ROC area of the full 16-member ensemble, which gives an indication of the COSMO-SREPS skill in forecasting precipitation for that period and in that particular area. In the left panel, the additional lines show the ROC area values of the four-member ensembles made up by the four members nested on one particular global model. These four members are differentiated only with respect to the physical parameterizations. They represent the skill of ensembles, which are only model perturbed but have the same initial and boundary conditions.

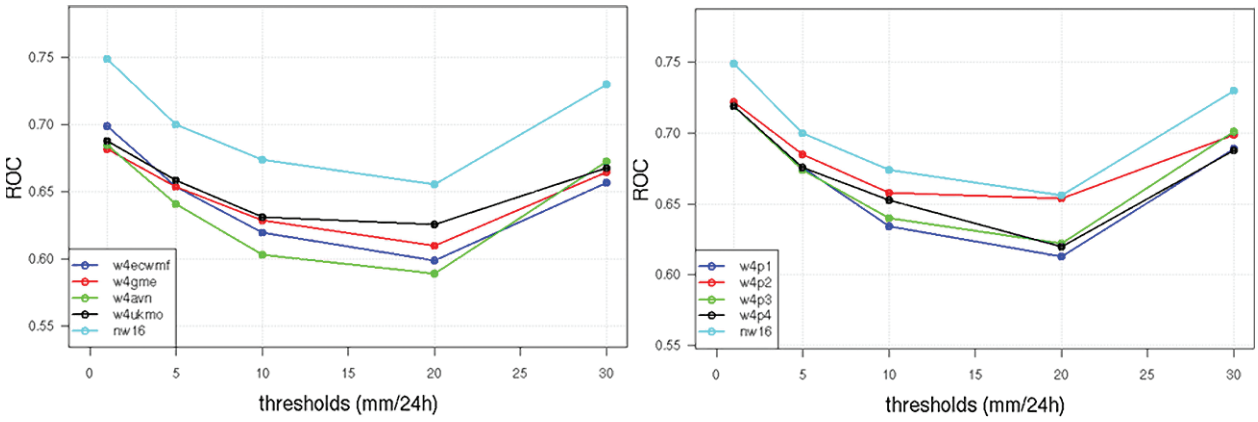

Fig. 6. ROC area as a function of threshold for 24-h accumulated precipitation in the Alpine region (high-resolution precipitation dataset in northern Italy and Switzerland as a reference). (left) Full 16-member COSMO-SREPS (light-blue) vs 4-member ensembles with identical "mother run" (blue: ECMWF, red: GME, green: GFS, black: UM). (right) Same as (left) but for identical physical perturbation (blue: p1, red: p2, green: p3, black: p4).

Apart from the decrease in skill evident when passing from a 16-member to a four-member ensemble, which is expected, the interesting point is that the different 4-member ensembles have different skill, that driven by UM showing higher ROC area values, whereas that driven by GFS being least skillful.

In the right panel of Fig. 6, the 4-member ensembles made up from identical physics perturbations are shown. These represent the skill of ensembles that are perturbed in the initial and boundary conditions only but have the same model setup. Comparing the two panels of Fig. 6 indicates that perturbation of initial conditions generally yields somewhat more skillful performance than physical perturbation. This indicates that the higher degree of diversity among members with different initial and boundary conditions yields better skill as compared to the smaller-scale variability introduced by the physics perturbations.

As for the role of the different parameterizations, the four-member ensemble in which model perturbation "p2" is applied to each member (red line in the right panel of Fig. 6) turns out to be more skillful than other four-member ensembles. For the present purpose of demonstration, it does not matter to which parameter p2 refers, although the performance of different parameters will be analyzed in more detail in the future. As a general conclusion, this result indicates that optimal tuning of a multimodel ensemble bears some potential for improving the skill.

Operational nowcasting of thunderstorms in the Alps using TRT. Local flash floods can cause severe damage in complex terrain and require the tracking of individual, even relatively small thunderstorms. Orography may also lead to the organization of convective cells at the meso-beta scale (20–200 km) in the pre-Alpine and plateau region. A typical problem of nowcasting is thus the identification and tracking of storm cells as in the Thunderstorm Identification, Tracking, and Nowcasting (TITAN; Dixon and Wiener 1993) or the Storm Classification Identification and Tracking (SCIT; Johnson et al. 1998) algorithms.

The DOP was an excellent opportunity to demonstrate the performance of the operational, object-oriented nowcasting system Thunderstorms Radar Tracking (TRT). TRT is a multiple-radar, multiple-sensor system that uses heuristic- and centroid-based methods for the automatic detection, tracking, characterization, and extrapolation of intense convective cells. It fully exploits volumetric reflectivity data of multiple-radar composites to describe the 3D storm structure and properties (Hering et al. 2006) and is tuned to identify individual cells rather than storm systems. Hence, the evolution of cell-based attributes, like vertically integrated liquid (VIL) 15/45-dBZ echo tops, the altitude of maximum storm reflectivity, and cloud-to-ground lightning flashes, are available to forecasters in real time as well as the respective gridded fields.

TRT is based on a dynamic threshold scheme applied to the reflectivity data of multiple-radar composites with a time resolution of 5 min and a spatial resolution of 2 km (Hering et al. 2004). It is able to identify each storm object at individual thresholds, depending on the stage of the storm's life cycle. With the current thresholds, thunderstorm cells as small as 16 km^2 (four pixels) can be identified. A detected storm cell is tracked in successive images using the

method of the geographical overlapping of cells. Complex cases with several cells, splits, and mergers are also taken into account.

The latest improvement in TRT is the cell severity ranking (CSR), developed for D-PHASE and tested during the DOP. Its goal is to find and highlight the most dangerous and strongest cells by combining the most significant cell severity attributes into one single parameter. For this purpose cells are classified into four distinct categories of severity and represented by a color-coded ellipse (Fig. 7). The severity categories are computed by integrating the three cell-based attributes VIL, the mean of 45-dB*Z* echo-top altitude, and the maximum cell reflectivity, with a different weighting (Hering et al. 2008).

TRT also includes a 1-h position forecast. The thunderstorm's estimated future position is computed using the individual cell's weighted displacement velocity. The expected position is displayed with an ellipse filled with parallel lines (Fig. 7), and it takes into account the spread (standard deviation) of the velocity vectors from the last three 5-min time steps. The size of the ellipse is proportional to the uncertainty of the position forecast (i.e., the larger the ellipse, the greater the uncertainty).

CSR was successfully used by forecasters during the DOP, and it allowed forecasters to focus on the most severe cells maintaining situational awareness and to speed up the decision process of thunderstorm warnings. Likewise, CSR has been a very popular and easy-to-use nowcasting product for the end users. A systematic evaluation of the extrapolation routine will be performed using the D-PHASE data.

VERA. VERA is a high-resolution, real-time analysis tool for applications over complex topography. Its purpose is to provide the best possible automated atmospheric analysis to assess the model forecasts for the ongoing situation and thus to decide the most trustworthy model. VERA uses physical a priori knowledge (so-called fingerprints) of typical meteorological patterns occurring over complex terrain. For example, the patterns of mesoscale cold highs or heat lows over complex terrain in undisturbed weather situations are strongly connected to topography. The fingerprint approach uses this information to refine the analysis. Details of the technical implementation for VERA can be found in Steinacker et al. (2000, 2006) or Bica et al. (2007).

A comprehensive data quality control (QC) scheme is used *upstream* of the analysis to exclude erroneous data. This module detects and filters unrealistic single measurements, gross errors, and systematic errors, thus eliminating in the spinup of the analysis erroneous patterns due to data errors (Häberli et al. 2004). If the domain under consideration is large enough, then VERA estimation of mean area precipitation yields satisfying results (Dorninger et al. 2008).

VERA was set up for the D-PHASE domain (Fig. 2) and for the parameters mean sea level pressure, (equivalent) potential temperature, 10-m wind, precipitation, and moisture flux divergence. Data from the World Meteorological Organization's (WMO's) Global Telecommunications Systems (GTS) were used to produce model-independent analyses on an hourly basis in real time during the DOP. Twenty-two minutes after observations, the graphics were available on the VP. The location and strength of the meteorological phenomenon under consideration were compared to model forecasts on the VERA grid in real time for selected models. VERA-to-model differences guided the forecaster concerning optimal model choice (different models or initialization times).

Figure 8 shows an example of a frontal movement from 18- and 6-h forecasts, respectively, issued by COSMO-2. Both are compared to the same VERA analysis. The forecasted frontal system (in terms of

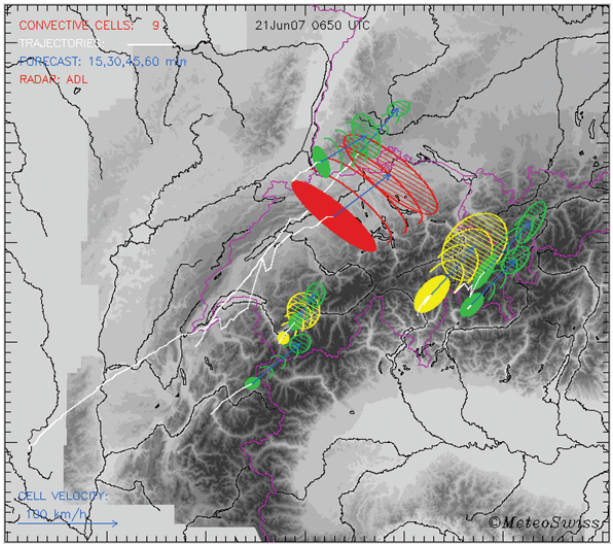

Fig. 7. Operational visualization of the TRT cell ranking product over the Alpine region (0650 UTC 21 Jun 2007). The solid-shaded ellipses indicate the current position and size of the thunderstorm cells. Hatched ellipses show a 60-min forecast of each cell's position, and the size is proportional to the uncertainty of the forecast. The cell objects are classified into the four categories: very weak (not visualized), weak (green), moderate (yellow), and severe (red). Elliptical arcs show the cell's expected position in 15-min steps.

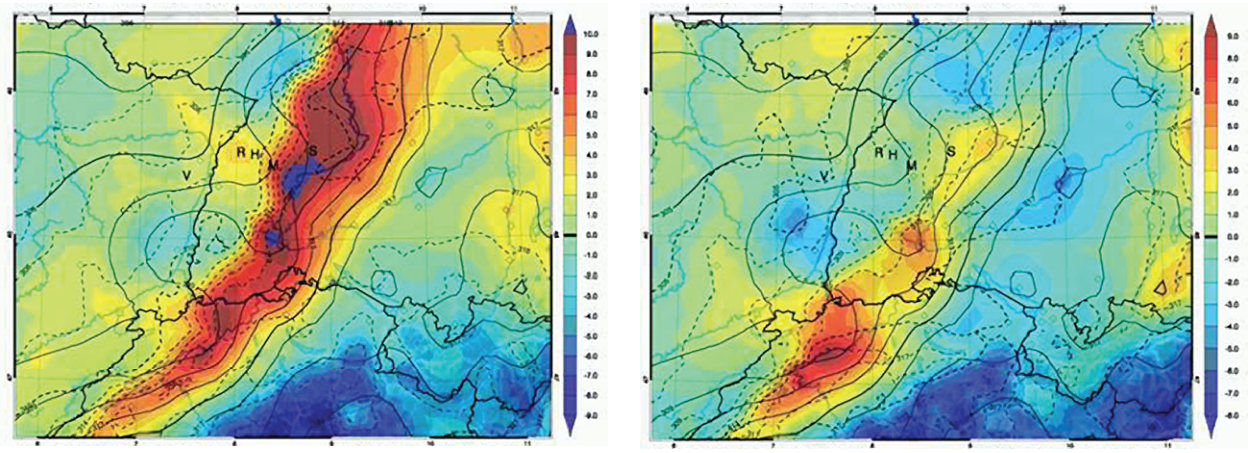

Fig. 8. Real-time COSMO-2-minus-VERA comparison of equivalent potential temperature; valid at 0600 UTC 18 Sep 2007. Shown is the northern part of Switzerland and surrounding countries. (left) 18-h forecast with COSMO-2; (right) 6-h forecast with COSMO-2. The color code depicts the difference (COSMO-2 forecast minus VERA analysis). Red means predicted values higher than analyzed and blue means predicted values lower than analyzed. Solid lines mean VERA analysis and dashed lines mean COSMO-2 predictions. Bold solid lines depict country borders. GTS data are the data source for analyzed fields.

the location of the gradient of the equivalent potential temperature) lacks behind the analysis for the older forecast. In the more recent forecast, the positions of forecasted and analyzed front locations fit quite well. Therefore, the forecaster would rely on the newer forecast in this case.

Online VERA was intended as additional information to the users. Still, a reanalysis of the DOP is planned using additional data from various sources in the D-PHASE domain and high-density datasets in the COPS domain, not available in real time. This reanalysis will again be compared to all the model runs and will result in a comprehensive model intercomparison with an independent reference. Clearly, these analyses will also be valuable for process studies.

USER FEEDBACK. *Social scientific approach.* A central part of evaluation activities for D-PHASE was a social scientific evaluation focusing on the users of the VP. Because one of the goals of D-PHASE was to investigate and improve situation analysis and decision-making processes by users, it was necessary to survey users in addition to objectively evaluating model quality. The approach was to measuring changes instead of simply assessing post hoc judgments about the usefulness or success, as is usually done. Another aim of the DOP was to create improved understanding among users of prognostic information and new approaches such as ensemble modeling. Thus, the goals of the social scientific evaluation were to establish whether the approaches were understood by users, how the quality of the situation analysis improved, and how the mutual perception of two user groups—end users (civil protection, among others) and atmospheric forecasters—had changed during the DOP.

A combination of methods was chosen:

- a quantitative quasi-experimental approach to assess changes during the DOP. [For this, 26 users completed two questionnaires, one in the beginning and one after the DOP.],
- a 1-day workshop with 16 users to discuss the outcomes of the DOP and possible measures for improving the VP [Information on the main problems encountered and on usage of D-PHASE information was collected.], and
- interviews with selected users to investigate how D-PHASE tools were incorporated into daily practice.

The main instrument of the evaluation, the semi-standardized questionnaire, contained questions about

- the VP per se; that is, its performance, accessibility, and general value for the users;
- content on the VP; that is, clarity of information, comprehensibility, complexity, and adequacy of information and perceived quality of model forecasts;
- effect of the platform on the users; that is, rise in comprehension of models, duties, and problems

of other user groups; rise in self-perceived competence (certainty with own decisions); and general judgments on the usefulness of information with regard to situation analysis and decision support.

Preliminary results indicate that D-PHASE was regularly used and that the information was incorporated into decision procedures. D-PHASE was mainly used before events, for example, in cases of first evidence for possible events. It was used less often during events for estimating the severity or the event's evolution. End users rated the D-PHASE platform as valuable, trustworthy, and easy to navigate. Room for improvement was noted concerning the technical performance of the platform (speed and availability of services).

Among the elements on the platform, regional and local maps were most often used. No particular information was missed on the VP. On the contrary, the amount of information was considered "rather too large." Most users indicated that they had been able to interpret the information but not all models were (subjectively) rated equally trustworthy. Users found the information beneficial to support situation analysis and decision making but could not decide if decisions had actually improved in specific cases. The availability of ensemble models and probabilistic information, which was largely unknown to most of end users prior to the DOP, was perceived as added value.

Feedback from atmospheric forecasters. A subjective verification was performed daily during the DOP by the on-duty forecasters of MeteoSwiss. This evaluation was targeted to assess the true benefits of the warning process for the forecaster and thus directly responds to the requirements of a FDP.

Several questions were addressed with an online multiple choice questionnaire. Questions ranged from "countable facts" (concerning models employed, flow situation, and so on) to subjective judgments (e.g., determining which VP component helped the most in the situation analysis?). Some preliminary results that can be deduced from a quick survey of the returned forms, are expressed in the following forecaster statements:

- Limited-area ensemble prediction systems have a significant positive effect (as compared to having only deterministic models available) for the estimation of the precipitation amount and for building confidence into the forecast.
- For precipitation events in the Alps, high-resolution models provide added value in about half of the cases. (In most other cases they have no added value, and sometimes these models even provide a poorer guidance than the coarser models.) The benefit of the higher resolution is more pronounced for ensemble systems than for deterministic models. Convection-resolving deterministic models sometimes failed to produce any precipitation at all, particularly in convective situations. [Note that this statement to some extent contrasts to the objective verification results (e.g., Fig. 5). Clearly, the reasons for the differences in subjective and objective judgment from different user groups will be the subject of deeper analysis of the results from this project.]
- Concerning alerts and their presentation on the VP, the large variety of models was appreciated by many forecasters. As for traditional model products (plots, meteograms, among others), however, there is little advantage in having (too) many models of the same type at disposition to increase forecast confidence. The number of available models during D-PHASE was too large to cope with and forecasters had to restrict their attention to well-known models. This holds true for normal as well as for high-impact weather situations.
- A suitable visualization system is essential for the forecasters to effectively use the vast amount of data and to extract the essential facts without losing relevant information. The VP, designed with the contribution of the forecasters, was a success for the duty. The automatic alerts allowed for a rapid overview of the relevant information and easy access to the data required in the warning process. However, the added value of the VP depends on the spread (both temporal and spatial) among the visualized models.
- There was no special preference among the newly available systems (i.e., models or tools introduced at the occasion of D-PHASE) in best supporting the forecasters in their decision-making process.
- Atmospheric forecasters appreciated the availability of hydrological information to better address end user needs and improve their own understanding of the hydrological processes.
- Last, but not least: the acceptance of the new generation of NWP model systems as well as the whole D-PHASE forecasting system and VP differed strongly among the individual forecasters.

SUMMARY AND CONCLUSIONS. The WWRP Forecast Demonstration Project D-PHASE successfully demonstrated recent advances in forecasting heavy precipitation events and related floods in the Alpine region. Many of these advances

have been initiated during the Mesoscale Alpine Programme (MAP). In this contribution we have presented the overall background of the project, its elements, and a number of applications that were triggered by D-PHASE.

At a bare minimum, D-PHASE has produced an unprecedented dataset that brings together results from more than 30 atmospheric modeling systems with a common domain, a common output format, joint warning procedures, and plots using the same coding (that can be compared without adaptation). At the same time, hydrological runoff simulations were performed in more than 40 catchments all over the Alps, and nowcasting information was provided to the numerous end users. The collaboration with the WWRP RDP COPS furthermore brought a unique opportunity to have high-quality observational data in a subdomain and subperiod of the DOP.

The available dataset will allow for

- systematically demonstrating the additional value of convection-permitting high-resolution atmospheric modeling (encouraging preliminary results, such as in Fig. 5);
- investigating the properties and performance of EPSs, both for atmospheric and hydrological models. Examples can be found in the "Examples of applications" section and in the online supplement [http://dx.doi.org/10.1175/2009BAMS2776.2; hydrological examples are reported in Zappa et al. (2008) and Germann et al. (2009)];
- studying predictability of convection processes and convective initiation using the present model results in connection with the observational results of COPS;
- identify benchmark models of all types by comparing them with a range of other models of the same category, or even other model types;
- systematically evaluating nowcasting tools such as the position forecast of convective systems in the TRT (see "Operational nowcasting of thunderstorms in the Alps using TRT" section) using the available data and possibly extending their functionality by introducing model products; and
- judging the end user feedback on its own grounds (and accept the consequences) and comparing it to the "objective" verification results, thus learning even more concerning the improvement of the overall forecasting/warning system.

Just as MAP proved the feasibility of atmospheric/hydrological coupling (e.g., Ranzi et al. 2007), D-PHASE successfully demonstrated its operational use and extension to ensemble techniques. Judging from the forecaster's preliminary conclusion (see the "User feedback" section), this is not only an advance in technical terms but it also helps the respective communities to take into account the other's sphere (hydrosphere versus atmosphere) to improve the decisions and forecasts in one's own models. When "expanding" into other areas such as air quality or health factor forecasting, the atmospheric community should keep such an experience well in mind.

The single most important factor of success for D-PHASE was probably the interoperability of all of the models. Common formats, common warning levels, and common routines to actually determine the warnings from the model outputs rendered the results comparable and therefore highly valuable.

The present FDP has provided a precious dataset that has to be further exploited. In conjunction with the reference observational dataset due to COPS in parts of the domain and during parts of the DOP, these data are now available as a test bed for atmospheric convection, in combination with orographic precipitation and coupled to hydrological modeling. At the time of writing, such plans are being considered within working groups of WWRP and the Hydrologic Ensemble Prediction Experiment (HEPEX).

ACKNOWLEDGMENTS. Most authors have been funded by their home institutions who made D-PHASE happen: MeteoSwiss, University of Vienna, Swiss Federal Institute for Forest Snow and Landscape Research, ARPA-SIM, University of Brescia, Météo-France, ISAC-CNR, USAM, Met Office, University Paul Sabatier, ARPAV, ETHZ, ZAMG, DLR, Deutscher Wetterdienst (DWD), and University of Hohenheim. D-PHASE profited from research contracts in the framework of COST 731, from the Swiss National Centre for Competence in Research Climate (NCCR Climate) as well as from SBF C05.0105. COPS, by the collaboration with which D-PHASE largely profited, was funded by the German Research Foundation (DFG) within the Priority Program 1167. Some of the verification exercises and modeling studies profited from data exchange through COSMO, in particular the high-resolution precipitation network fed by the Italian regional meteorological services of Bolzano, Emilia-Romagna, Friuli-Venezia-Giulia, Liguria, Marche, Piemonte, Trento, and ARPA Veneto. The Spanish Meteorological service INM [recently changed to Agencia Estatal de Meteorologia (AEMET)] kindly provided the boundary fields for the COSMO-SREPS project, and ARPA Piemonte contributed to the Swiss–Italian radar composit. The efficient and flexible programming of the D-PHASE VP through

next-generation software and its owner, G. Kittel, is very much acknowledged.

REFERENCES

Ament, F., and M. Arpagaus, cited 2008: Precipitation forecasts in Switzerland—An assessment by the forecast demonstration project MAP D-PHASE. [Available online at www.smr.arpa.emr.it/dphase-cost/presentazioni/21-May/Morning/Session_3_am/3.04-amt/200805_ament_dphase_bologna.pdf.]

Arpagaus, M., and Coauthors, 2009: MAP D-PHASE: Demonstrating forecast capabilities for flood events in the Alpine region. MeteoSwiss Publication Series 78, 75 pp. [Available online at www.meteoschweiz.admin.ch/web/de/forschung/publikationen/meteoschweiz_publikationen/veroeffentlichungen.html.]

Bacchi, B., and R. Ranzi, 2003: Hydrological and meteorological aspects of floods in the Alps: An overview. *Hydrol. Earth Syst. Sci.,* **7,** 784–798.

Benoit, R., and Coauthors, 2002: The real-time ultrafinescale forecast support during the special observing period of the MAP. *Bull. Amer. Meteor. Soc.,* **83,** 85–109.

——, N. Kouwen, W. Yu, S. Chamberland, and P. Pellerin, 2003: Hydrometeorological aspects of the Real-Time Ultrafinescale Forecast Support during the Special Observing Period of the MAP. *Hydrol. Earth Syst. Sci.,* **7,** 877–889.

Bica, B., and Coauthors, 2007: Thermally and dynamically induced pressure features over complex terrain from high-resolution analyses. *J. Appl. Meteor. Climatol.,* **46,** 50–65.

Binder, P., and C. Schär, 1996: MAP design proposal. Mesoscale Alpine Programme Rep., 79 pp. [Available online at www.map.meteoswiss.ch/map-doc/proposal.htm.]

Bougeault, P., and Coauthors, 2001: The MAP special observing period. *Bull. Amer. Meteor. Soc.,* **82,** 433–462.

Casati, B., and Coauthors, 2008: Forecast verification: Current status and future directions. *Meteor. Appl.,* **15,** 3–18, doi:10.1002/met.52.

Denhard, M., and S. Trepte, 2006: Calibration of the European multi-model ensemble SRNWP-PEPS. *Proc. Second THORPEX Int. Science Symp.,* WMO Tech. Document 1355, WWRP/THORPEX 7, Landshut, Germany, THORPEX, 56–57.

Dixon, M., and G. Wiener, 1993: TITAN: Thunderstorm identification, tracking, analysis, and nowcasting—A radar-based methodology. *J. Atmos. Oceanic Technol.,* **10,** 785–797.

Dorninger, M., S. Schneider, R. Steinacker, 2008: On the interpolation of precipitation data over complex terrain. *Meteor. Atmos. Phys.,* 175–189, doi:10.1007/s00703-008-0287-6.

Frei, C., and C. Schär, 1998: A precipitation climatology of the Alps from high-resolution rain-gauge observations. *Int. J. Climatol.,* **18,** 873–900.

——, R. Schöll, S. Fukutome, J. Schmidli, and P. L. Vidale, 2006: Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models. *J. Geophys. Res.,* **111,** D06105, doi:10.1029/2005JD005965.

Fundel, F., M. A. Liniger, A. Walser, C. Frei, and C. Appenzeller, 2009: Calibrated precipitation forecasts for a limited area ensemble forecast system using reforecasts. *Mon. Wea. Rev.,* in press.

Garcia-Moya, J. A., A. Callado, C. Santos, D. Santos-Munoz, and J. Simarro, 2009: Predictability for short-range forecast: A multi-model approach. AEMET Tech. Rep., in press.

Germann, U., G. Galli, M. Boscacci, and M. Bolliger, 2006: Radar precipitation measurement in a mountainous region. *Quart. J. Roy. Meteor. Soc.,* **132,** 1669–1692.

——, M. Berenguer, D. Sempere-Torres, and M. Zappa, 2009: Ensemble radar precipitation estimation for hydrology in a mountainous region. *Quart. J. Roy. Meteor. Soc.,* **135,** 445–456, doi:10.1002/qj.375.

Häberli, C., I. Groehn, R. Steinacker, W. Pöttschacher, and M. Dorninger, 2004: Performance of the surface observation network during MAP. *Meteor. Z.,* **13,** 109–122.

Hering, A. M., C. Morel, G. Galli, S. Sénési, P. Ambrosetti, and M. Boscacci, 2004: Now casting thunderstorms in the alpine region using a radar based adaptive thresholding scheme. *Proc. Third European Conf. on Radar in Meteorology and Hydrology (ERAD),* Visby, Sweden, ERAD, 206–211.

——, U. Germann, M. Boscacci, and S. Sénési, 2006: Operational thunderstorm nowcasting in the Alpine region using 3D-radar severe weather parameters and lightning data. *Proc. Fourth European Conf. on Radar in Meteorology and Hydrology (ERAD),* Barcelona, Spain, ERAD, 453–456. [Available online at www.grahi.upc.es/ERAD2006/proceedingsMask/00122.pdf.]

——, ——, ——, and ——, 2008: Operational nowcasting of thunderstorms in the Alps during MAP D-PHASE. *Proc. Fifth European Conf. on Radar in Meteorology and Hydrology (ERAD),* Helsinki, Finland, ERAD, 5 pp.

Hohenegger, C., and C. Schär, 2007: Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull. Amer. Meteor. Soc.,* **88,** 1783–1793.

——, D. Lüthi, and C. Schär, 2006: Predictability mysteries in cloud-resolving models. *Mon. Wea. Rev.,* **134,** 2095–2107.

Johnson, J. T., P. L. MacKeen, A. Witt, E. D. Mitchell, G. J. Stumpf, M. D. Eilts, and K. W. Thomas, 1998: The Storm Cell Identification and Tracking algorithm: An enhanced WSR-88D algorithm. *Wea. Forecasting,* **13,** 263–276.

Keenan, T., and Coauthors, 2003: The Sydney 2000 World Weather Research Programme Forecast Demonstration Project: Overview and current status. *Bull. Amer. Meteor. Soc.,* **84,** 1041–1054.

Marsigli, C., A. Montani, F. Nerozzi, T. Paccagnella, S. Tibaldi, F. Molteni, and R. Buizza, 2001: A strategy for high-resolution ensemble prediction. Part II: Limited-area experiments in four Alpine flood events. *Quart. J. Roy. Meteor. Soc.,* **127,** 2095–2115.

——, ——, and T. Paccagnella, cited 2006: The COSMO-SREPS project. [Available online at http://SRNWP.met/hu/annual_meetings/2006/programme2006/html.]

Molteni, F., R. Buizza, C. Marsigli, A. Montani, F. Nerozzi, and T. Paccagnella, 2001: A strategy for high-resolution ensemble prediction. Part I: Definition of representative members and global-model experiments. *Quart. J. Roy. Meteor. Soc.,* **127,** 2069–2094.

Ranzi, R., B. Bacchi, and G. Grossi, 2003: Runoff measurements and hydrological modelling for the estimation of rainfall volumes in an Alpine basin. *Quart. J. Roy. Meteor. Soc.,* **129,** 653–672.

——, M. Zappa, and B. Bacchi, 2007: Hydrological aspects of the Mesoscale Alpine Programme: Findings from field experiments and simulations. *Quart. J. Roy. Meteor. Soc.,* **133,** 867–880.

Richard, E., A. Buzzi, and G. Zängel, 2007: Quantitative precipitation forecast in the Alps: The advances achieved by the Mesoscale Alpine Programme. *Quart. J. Roy. Meteor. Soc.,* **133,** 831–846.

Richardson, D. S., 2000: Skill and relative economic value of the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.,* **126,** 649–667.

Rotach, M. W., and D. Zardi, 2007: On the boundary-layer structure over highly complex terrain: Key findings from MAP. *Quart. J. Roy. Meteor. Soc.,* **133,** 937–948, doi:10.1002/qj.71.

Rotunno, R., and R. A. Houze, 2007: Lessons on orographic precipitation from the Mesoscale Alpine Programme. *Quart. J. Roy. Meteor. Soc.,* **133,** 811–830.

Siccardi, F., G. Boni, L. Ferraris, and R. Rudari, 2005: A hydrometeorological approach for probabilistic flood forecast. *J. Geophys. Res.,* **110,** D05101, doi:10.1029/2004JD005314.

Steinacker, R., C. Häberli, and W. Pöttschacher, 2000: A transparent method for the analysis and quality evoluation of irregularly distributed and noisy observational data. *Mon. Wea. Rev.,* **128,** 2303–2316.

——, and Coauthors, 2006: A mesoscale data analysis and downscaling method over complex terrain. *Mon. Wea. Rev.,* **134,** 2758–2771.

Verbunt, M., A. Walser, J. Gurtz, A. Montani, and C. Schär, 2007: Probabilistic flood forecasting with a limited-area ensemble prediction system: Selected case studies. *J. Hydrometeor.,* **8,** 897–909.

Volkert, H., and T. Gutermann, 2007: Inter-domain cooperation for mesoscale atmospheric laboratories: The Mesoscale Alpine Programme as a rich study case. *Quart. J. Roy. Meteor. Soc.,* **133,** 949–968.

Walser, A., and C. Schär, 2004: Convection-resolving precipitation forecasting and its predictability in Alpine river catchments. *J. Hydrol.,* **288,** 57–73.

——, D. Lüthi, and C. Schär, 2004a: Predictability of precipitation in a cloud-resolving model. *Mon. Wea. Rev.,* **132,** 560–577.

——, M. W. Rotach, M. Arpagaus, C. Appenzeller, C. Marsigli, and A. Montani, 2004b: A Limited-Area Ensemble Prediction System. Preprints, *Int. Workshop on Timely Warnings of Heavy Precipitation Episodes and Flash Floods,* Ljubljana, Solvenia, Slovenia Meteorological Society, 45–50.

Weigel, A. P., F. K. Chow, M. W. Rotach, R. L. Street, and M. Xue, 2006: High-resolution large-eddy simulations of flow in a steep Alpine valley. Part II: Flow structure and heat budgets. *J. Appl. Meteor. Climatol.,* **45,** 87–107.

Wulfmeyer, V., and Coauthors, 2008: Research campaign: The Convective and Orographically Induced Precipitation Study. *Bull. Amer. Meteor. Soc.,* **89,** 1477–1486.

Wunram, C., 2006: Concept of COPS long term data archive. *Proc. First MAP D-PHASE Scientific Meeting,* Vienna, Austria, World Weather Research Programme, 33.

Zappa, M., and Coauthors, 2008: MAP D-PHASE: Real-time demonstration of hydrological ensemble prediction systems. *Atmos. Sci. Lett.,* **9,** 80–87, doi:10.1002/asl.183.

Zus, F., M. Grzeschik, H. S. Bauer, V. Wulfmeyer, G. Dick, and M. Bender, 2008: Development and optimization of the IPM MM5 GPS slant path 4DVAR System. *Meteor. Z.,* **17,** 867–885.