AN IMPROVED SYSTEM FOR SHORT-TERM PREDICTION OF EXTREME PRECIPITATION EVENTS
# Deliverable 3.2

**An improved system for short-term prediction of extreme precipitation events.**

<table>
<thead>
<tr>
<th>Related Work Package:</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deliverable lead:</td>
<td>Magnus Lindskog</td>
</tr>
<tr>
<td>Author(s):</td>
<td>Magnus Lindskog and Tomas Landelius</td>
</tr>
<tr>
<td>Contact for queries</td>
<td><a href="mailto:Magnus.Lindskog@smhi.se">Magnus.Lindskog@smhi.se</a></td>
</tr>
<tr>
<td>Grant Agreement Number:</td>
<td>n° 641811</td>
</tr>
<tr>
<td>Instrument:</td>
<td>HORIZON 2020</td>
</tr>
<tr>
<td>Start date of the project:</td>
<td>01.10.2015</td>
</tr>
<tr>
<td>Duration of the project:</td>
<td>48 months</td>
</tr>
<tr>
<td>Website:</td>
<td><a href="http://www.imprex.eu">www.imprex.eu</a></td>
</tr>
<tr>
<td>Abstract</td>
<td>A limited-area forecasting system has been used to investigate the impact of enhanced surface data assimilation on the short-term prediction of extreme precipitation events. In this report the enhancements are described and it is demonstrated that they can lead to improved short-term prediction of extreme precipitation events over a Southern European domain.</td>
</tr>
</tbody>
</table>

---

## Dissemination level of this document

<table>
<thead>
<tr>
<th>X</th>
<th>PU</th>
<th>Public</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PP</td>
<td>Restricted to other programme participants (including the Commission Services)</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>Restricted to a group specified by the consortium (including the European Commission Services)</td>
</tr>
<tr>
<td></td>
<td>CO</td>
<td>Confidential, only for members of the consortium (including the European Commission Services)</td>
</tr>
</tbody>
</table>

---

## Versioning and Contribution History

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Modified by</th>
<th>Modification reasons</th>
</tr>
</thead>
<tbody>
<tr>
<td>v.01</td>
<td>20180228</td>
<td>Magnus Lindskog</td>
<td>Original version by SMHI.</td>
</tr>
<tr>
<td>v.02</td>
<td>20180313</td>
<td>Magnus Lindskog and Tomas Landelius</td>
<td>Revised version by SMHI based on comments from work package leader Erik Kjellström.</td>
</tr>
</tbody>
</table>
Content

List of figures .......................................................................................................................... 4
List of tables ............................................................................................................................. 6
Executive summary ..................................................................................................................... 7

1 Introduction ............................................................................................................................ 9

2 The NWP system and its default surface data assimilation ................................................... 9
   General overview .................................................................................................................... 9
   Default surface data assimilation of temperature and soil moisture .................................. 11

3 Model setup and extreme precipitation events ..................................................................... 12

4 Enhancements in surface data assimilation of temperature and soil moisture ................. 14
   Overview .............................................................................................................................. 14
   Horizontal distribution ........................................................................................................ 14
   Vertical distribution ............................................................................................................. 15
   Use of satellite-based soil moisture observations ............................................................. 15

5 Experimental design ............................................................................................................ 18

6 Results .................................................................................................................................. 19
   Demonstration of functionality of enhancements ................................................................. 19
   Verification scores ................................................................................................................. 22
   Verification for local areas ................................................................................................... 25
   Subjective verification of a case study .................................................................................. 28

7 Conclusions ........................................................................................................................... 33

8 References ............................................................................................................................. 35
List of figures

Figure 1: The HARMONIE-AROME Modelling domain (red frame) and area of studied extreme precipitation events (blue circle).

Figure 2: Channel 9 infrared brightness temperature at 27 July 03.30 UTC derived from the SEVIRI instrument. The low values over Northern France are associated with a convective precipitation cell over Northern France.

Figure 3: Illustration of ASCAT pre-processing for 20130614 09 UTC. ASCAT observation based soil moisture products (unit: %) are converted to $w_s^{ascat}$ soil moisture pseudo-observations (unit: m$^3$/m$^3$).

Figure 4: Merging of ASCAT satellite data from METOP-A (green) and METOP-B (red) for 20140614 09 UTC. The merged data are marked blue.

Figure 5: Impact on the two-metre relative humidity field (unit in plot: 0-1) of one single two-metre relative humidity observation located at a position (marked with a black dot) close to the west-coast of France. The observed relative humidity is approximately 0.15 less than the corresponding model value. Left is with standard structure functions and middle with MESCAN structure functions. Right part shows model orography (unit: m).

Figure 6: Impact on the two-metre temperature field (unit: K) of one single two-metre temperature observation located at a position (marked with a black dot) close to the west-coast of France. The observed relative humidity is approximately 1.5 K larger than the corresponding model value. Left is with standard structure functions and middle with MESCAN structure functions. Right part shows model orography (unit: m).

Figure 7: Histogram of Jacobians for surface grid-points within the model domain for the period 20140621 to 2013062830. Histograms to the left are for $\frac{\partial RH_{2m}}{\partial w_2}$ (unit: 1/m$^3$/m$^3$) and histograms to the right are for $\frac{\partial T_{2m}}{\partial T_2}$ (unit: K/K). Red histograms are for 00 UTC and blue histograms are for 12 UTC.

Figure 8: Data assimilation increments of Soil $w_s$ soil moisture (unit: % of change of the corresponding background value) for experiment EKF-MESC-SCAT (left) and OI (right) at 20140625 18 UTC.

Figure 9: Bias and standard deviation of temperature (left, unit: K) and relative humidity (right, unit: %) forecasts for verification against radiosonde observations within the entire model domain. Verification scores are accumulated for +12 and +24 h forecasts and shown for different vertical levels. Different rows are for different cases and different colours are for different experiments.
Figure 10: Bias and standard deviation of 12 h accumulated precipitation (unit: mm/12 h) for verification against Synop gauge measurements within the entire model domain. Verification scores are shown as function of forecast length from for +12 to +24 h. Different rows are for different events and different colours are for different experiments.

Figure 11: Location of radiosondes used for local verification of the studied cases. Radiosondes marked with red dots are used for the two cases of precipitation event observed over the Pyrenees area and radiosondes marked with blue dots are used for the precipitation event observed over Northern France.

Figure 12: Bias and standard deviation of temperature (left, unit: K) and relative humidity (right, unit: %) forecasts for verification against radiosonde observations over local areas. Verification scores are accumulated for +12 and +24 h forecasts and shown for different vertical levels. Different rows are for different cases and different colours are for different experiments.

Figure 13: Bias and standard deviation of 12 h accumulated precipitation (unit: mm/12 h) for verification against Synop gauge measurements over local areas. Verification scores are shown as function of forecast length from for +12 to +24 h. Different rows are for different events and different colours are for different experiments.

Figure 14: Multi-sensor precipitation rain-rate (mm/hr) for a large convective system passing over France. The estimated rain-rates are for 20130627 19.00 UTC (upper left), 20130628 03.15 UTC (upper right) and 20130628 08.30 UTC (lower left).

Figure 15: Synop gauge 6 h accumulated precipitation (unit: mm/6 h) between 20130626 18 UTC and 2013062700 00 UTC (left) and between 20130627 06 UTC and 2013062700 12 UTC (right).

Figure 16: Short-range forecast of 6h accumulated precipitation with reference system (OI, left) and system with enhanced surface data assimilation (EKF-MESC-SCAT, right). The forecasts are launched on 20130726 06 UTC and the accumulation period is between 20130727 00 and 20130727 06 UTC, which means that the forecast are at a range of +24 h.

Figure 17: Initial state differences in w2 soil moisture (unit: m3/m3) and T2 soil temperature (right: unit: K) between experiments EKF-MESC-SCAT and OI. The differences are for the initial state 20130726 06 UTC.

Figure 18: Differences in Convective Available Potential Energy (unit: J/kg) between 6 h forecasts of experiments EKF-MESC-SCAT and OI. Forecasts are launched from 20130726 06 UTC and valid at 20130726 12 UTC.
List of tables

Table 1: Description of extreme precipitation events studied with HARMONIE-AROME modelling system.

Table 2: Description the four of parallel experiments applied to each precipitation event studied.
Executive summary

A state-of-the-art km-scale limited-area forecasting system has been used to investigate the impact of enhanced surface data assimilation on the short-term prediction of extreme precipitation events. Data assimilation concerned with finding the best possible initial state by combining observations with a model, is a key aspect for short-term weather prediction. The enhancements comprise an improved representation of horizontal background error statistics, the introduction of a Kalman-filter based data assimilation methodology and the utilization of satellite based soil moisture information. The functionality of the enhancements of some key aspects of the surface data assimilation was demonstrated. Awaiting a semi-operational systematic evaluation over a longer time period, we here checked that our adjustments did not negatively impact a main purpose of our NWP system in IMPREX, which is predicting extreme precipitation events. In this study we confirm that our surface data assimilation adjustments do not negatively affect extreme precipitation forecasts, and even tend to slightly improve them, although not in a robust statistical sense.
Introduction

The Earth’s surface plays a central role both for hydrological and meteorological forecasting. On the one hand, most socio-economic processes that are affected by weather and water are situated on the Earth’s surface. On the other hand, the exchange of heat, water and water vapour between the surface, its deeper layers and the atmosphere are important drivers for the meteorological and hydrological states of the Earth’s system. In this work we focus on the effect of an accurate representation of surface properties for short term (up to a few days) numerical weather prediction (NWP). Data assimilation is a key aspect of NWP, which optimally combines observations with a model in order to spread the observational information and to produce the best possible initial state for the model. Enhanced surface data assimilation in NWP can lead to improved prediction capabilities of extreme precipitation events, which is highly important, among others, for hydrological run-off forecasts. A number of scientific studies have earlier shown a significant impact of soil moisture conditions on weather forecast skill at short and medium range (van den Hurk et al. 2008; Drusch and Viterbo 2007; Douville et al. 2000; Mahfouf et al. 2000; Beljaars et al. 1996) as well as at seasonal range (Weisheimer et al. 2011; Koster et al. 2011, 2004).

In numerical weather prediction (NWP) models the atmospheric state is typically represented by the variables: surface pressure, temperature, specific humidity, wind components, clouds and surface variables. The model variables are defined on discrete grid points. Starting from an initial state, the forecast models then integrate this state forward in time with the aim to simulate the atmospheric evolution on the temporal and spatial scales of the model. Non resolved sub-grid processes are parameterised to represent their average effect at the scales resolved by the model.

Data assimilation is the process where the most likely initial state is estimated, given the available observations and a so-called first guess or model background in terms of a previous forecast. Usually an intermittent data assimilation cycle is used. In case of a 3-hour intermittent data assimilation cycle this means that a 3-hour forecast launched on 00 UTC (from an initial state produced by the data assimilation for 00 UTC) will 3 hours later be used as background state in the data assimilation at 03 UTC. The 3-hour forecast launched on 03 UTC will then be used as background state at 06 UTC and so forth. Also forecasts of longer range than 3 hours are produced but it is the 3-hour forecast that is used for data assimilation within the assimilation cycle procedure.

It was early realised (Lorenz, 1965) that the forecast quality is strongly dependent on an accurate description of the initial state and hence on the abilities of the assimilation system. The reasons for this is that the atmospheric system is strongly non-linear and that small errors, inevitable inherent in the modelling system due to limited model resolution, simplifying assumptions and limited observations, sometimes grow large and deteriorate the forecast. However, the sensitivity to errors in the initial state depends on the atmospheric conditions. In general, the quality of the forecasted precipitation during extreme precipitation events is sensitive to small errors in the model state during and before the event. During other atmospheric conditions, like for example in a high-pressure system with subsiding air-mass the forecasts of precipitation are less sensitive to errors in the model state.
The numerical weather prediction system used here is developed in the framework of the shared “Aire Limitée Adaptation Dynamique Développement InterNational” (ALADIN) - High-Resolution Limited-Area Model (HIRLAM) NWP system. This system can be run with different configurations and here the so-called HIRLAM ALADIN Regional Meso-scale Operational NWP In Europe-Application of Research to Operations at Mesoscale (HARMONIE-AROME) is used (Bengtsson et al., 2017). The main components of the system are: surface data assimilation, upper-air data assimilation and the forecast model for the forward time integration. The data assimilation and modelling components are described in section 2, with emphasis on surface data assimilation.

We aim to improve the short-term (up to a few days) prediction of extreme precipitation and hydrological processes by improving the representation of the surface initial states through the introduction of refinements in the surface data assimilation used in the reference HARMONIE-AROME state-of-the art km-scale forecasting system. The refinements concern both improved data assimilation methodologies and enhanced observation usage. A Kalman-filter based data assimilation technique has been introduced to obtain more situation dependent data assimilation for soil moisture and surface temperature. In the horizontal direction, the background error variation used for the soil moisture and the surface temperature data assimilation has been enhanced to better derive small scale variations in surface conditions. Observation usage has been enhanced by assimilating surface remote sensing data related to soil moisture. Based on the impact of soil moisture initialisation on forecast skill found within a number of scientific studies and on operational experiences from the HARMONIE-AROME modelling system we see a potential for improvement of short-range km-scale NWP forecast by enhancing the surface data assimilation. Three cases associated with heavy precipitation studied within IMPREX were used to evaluate the enhanced surface data assimilation. The cases cover both synoptically driven precipitation in mountainous areas and convective precipitation in flat terrain. The cases have the potential to be influenced by surface fluxes influenced by the surface state and the surface data assimilation. Awaiting a semi-operational systematic evaluation of the enhancements over a longer time period, we check that, for these three selected cases, our adjustments do no negatively impact a main purpose of our NWP system in IMPREX, which is predicting extreme precipitation events. The extreme precipitation events and the general model setup are described in section 3. Section 4 deals with a demonstration of the functionality of the data assimilation improvements while section 6 and 7 deals with the design of the experiments and the corresponding results. Finally conclusions are presented in section 7.

2 The NWP system and its default surface data assimilation

General overview

The main components of the HARMONIE-AROME configuration of the HIRLAM ALADIN NWP system are: the surface data assimilation, the upper-air data assimilation and the forecast
model. Here we used a version of the HARMONIE-AROME configuration referred to as cy38h1.2. The forecast model configuration, e.g. dynamical core and physical parameterizations, are described in detail in Seity et al. (2011) and Bengtsson et al. (2017). It has a spectral representation with a non-hydrostatic formulation. Stratiform and deep convective clouds are explicitly represented, while for shallow convection a sub-grid parameterization is applied using the EDMF (Eddy Diffusivity Mass Flux) scheme. The representation of the turbulence is based on a prognostic Turbulent Kinetic Energy (TKE) equation combined with a diagnostic mixing length (Cuxart et al., 2000). The radiative transfer of the short-wave spectrum is described with six spectral bands (Fouquart and Bonnel., 1980) and the long-wave radiation is modelled using 16 spectral bands in accordance with Mlawer et al. (1997). Surface processes are modelled using SURFEX (Masson et al, 2013) together with a three layer ISBA scheme (Noilhan and Planton, 1989). Snow effects are parameterized using a one layer snow scheme in accordance with Douville et al. (1995).

Lateral boundary conditions are provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) global forecast model. Operational forecasts launched each 6 hours with a 1 hour output frequency are used. In addition, a spectral large scale mixing of the background state, the 3 hour HARMONIE forecast fields with the lateral boundary ECMWF fields is applied (Müller et al., 2017). In this way we hoped to benefit from the high-quality large scale information from the ECMWF global forecasts in the regional HARMONIE-AROME data assimilation.

In the 3-dimensional variational upper-air data assimilation applied (Fischer et al., 2005), conventional types of in-situ measurements are used, including: radiosonde, pilot-balloon wind, synop, ship, and aircraft measurements. These in-situ measurements concern surface pressure, temperatures, winds and relative humidity. In addition, radiances from the AMSU-A, AMSU-B/MHS instruments were used. Climatological background error statistics are used and these are derived from an ensemble of HARMONIE-AROME forecast differences obtained through downscaling of ECMWF Ensemble Data Assimilation (EDA)-based forecast fields over the limited area model domain. Then the evolved high-resolution ensemble was scaled to be consistent with the amplitude of the 3-hour forecast error for HARMONIE-AROME.

The default HARMONIE-AROME surface data assimilation is based on optimal interpolation of Synop observations utilizing horizontally homogenous and isotropic background error statistics. Synop measurements of two metre temperature and relative humidity as well as snow observations are then used to initialise the surface temperature, soil temperature, soil moisture and snow field of the surface part of the grid-points. There is however room for improvement regarding both methodology and observation usage for the surface data assimilation. Given the impact of soil moisture initialisation on forecast skill found within a number of scientific studies and the rather basic current system it is believed that enhancement of some key aspects of the HARMONIE-AROME surface data assimilation will lead to improved short-range numerical weather prediction. In particular we think that spatial inhomogeneities and flow dependency can be better represented and we see a potential in utilizing satellite based information. The current default system and enhancement carried out within this project are described in more detail below.
The HARMONIE-AROME system was applied in a deterministic mode, where only one initial state is estimated for each situation and only one forecast is launched from that initial state.

**Default surface data assimilation of temperature and soil moisture**

The surface data assimilation is to initialise the land area part of the grid-points. Variables to be initialised are: surface temperature, soil temperature, soil moisture and snow. The initialisation is done applying an optimal interpolation technique and by utilizing Synop observations of two-metre temperature and relative humidity as well as snow observations.

For temperatures and moisture, the surface data assimilation procedure is comprised of two steps. In the first step, a two-dimensional data assimilation based on the CANARI (Code d'Analyse Nécessaire à ARPEGE pour ses Rejets et son Initialisation) optimal interpolation scheme (Taillefer, 2002) is carried out to horizontally distribute the information from two-metre temperature and relative humidity observations and to get updated two-metre temperature and two-metre relative humidity values over all grid-points containing land parts, as described in equation (1):

\[ x_a = x_b + B H^T (H B H^T + R)^{-1} (y - H x_b) \]  

(1)

where \( x_a \) and \( x_b \) represents respectively the analysed and background values of grid-point two-metre temperature and relative humidity. \( B \) and \( R \) are the matrices containing the covariances of background errors and observation errors, respectively. Furthermore, \( y \) denotes the vector of the observations and \( H \) is the observation operator, projecting the model state on the observations. Finally, \( H^T \) denotes transpose of matrix and \( R^{-1} \) denotes inverse of matrix. The observation errors are assumed to be uncorrelated while the background errors are assumed to have an isotropic and horizontally homogenous background error statistics with a horizontal correlation function given by equation (2):

\[ corr_{i,j} = e^{-\frac{r_{i,j}}{2a}} \]  

(2)

Here \( corr_{i,j} \) represents the horizontal background error correlation between points \( i \) and \( j \) separated apart by a distance \( r_{i,j} \). The parameter \( a \) characterises the correlation length scale and has a value of 85 km for humidity and 80 km for temperature. The observation error standard deviations for temperature and relative humidity are assumed to be 1.4 K and 10%, respectively. The corresponding background error standard deviations are assumed to be 1.6 K and 18%, respectively. The correlations are sometimes referred to as structure functions. These numbers are applied generally irrespective of location and season based on Taillefer (2002).
In a second step, after the two-dimensional horizontal distribution just described, the two-metre temperature and relative humidity information is vertically distributed to update the surface temperature and soil moisture applying another optimal interpolation scheme as described in equations (3-6). This is done independently for each grid-point using the horizontally distributed two-metre temperature and relative humidity information to update the surface and soil temperatures as well as soil moisture:

\[
\begin{align*}
    w_g^a &= w_g^b + \alpha_1(T_{2m}^a - T_{2m}^b) + \alpha_2(RH_{2m}^a - RH_{2m}^b) \\
    w_2^a &= w_2^b + \beta_1(T_{2m}^a - T_{2m}^b) + \beta_2(RH_{2m}^a - RH_{2m}^b) \\
    T_s^a &= T_s^b + \mu_1(T_{2m}^a - T_{2m}^b) + \mu_2(RH_{2m}^a - RH_{2m}^b) \\
    T_2^a &= T_2^b + \nu_1(T_{2m}^a - T_{2m}^b) + \nu_2(RH_{2m}^a - RH_{2m}^b)
\end{align*}
\]

where \(T_{2m}\) and \(RH_{2m}\) denotes two-meter temperature and relative humidity, respectively. \(T_s\) and \(w_g\) denote surface temperature and ground layer moisture and \(T_2\) and \(w_2\) denote second layer temperature and soil moisture, respectively. The surface temperature layer is associated with a time scale of less than one hour while the second layer time scale is represented by the mean of \(T_s\) over one day. The surface soil moisture layer is on the order of 1 cm thick and the thickness of the second and third layers are defined by the root depth and soil depth, respectively, which are vegetation type dependent. Here superscript \(^a\), \(^b\) and \(^o\) denotes analyzed, background and observed value respectively, where information from the observed value has actually been horizontally distributed to all surface grid-point using the horizontal optimal interpolation in the first step. Furthermore, \(\alpha_1, \alpha_2, \beta_1, \beta_2, \mu_1, \mu_2, \nu_1\) and \(\nu_2\) are empirically derived coefficients described in (Giard and Bazile, 2000). The lowermost surface layer is associated with time-scales longer than a day and is not updated by the surface data assimilation.

3 Model setup and extreme precipitation events

The HARMONIE-AROME modelling system as described above has been set-up over a south European domain shown in Figure 1 to investigate the model capability to predict three different extreme precipitation events. The cases cover both synoptically driven precipitation in mountainous areas and convective precipitation in flat terrain. The cases have the potential to be influenced by surface fluxes influenced by the surface state and the surface data assimilation. A total of 768 × 648 horizontal grid-points were used with a horizontal model resolution of 2.5 kilometres and 65 vertical levels. In the upper-air data assimilation, conventional types of observations and satellite measurements from the AMSU-A and AMSU-B/MHS instruments, placed on-board polar orbiting satellites, were used. In addition, a spectral large-scale mixing of the background state (the 3 h HARMONIE forecast) with the lateral boundary ECMWF fields was applied. In this way, we hoped to benefit from the high-quality large-scale information from the ECMWF global forecasts in the regional HARMONIE-
AROME upper-air data assimilation. A 3-h data assimilation was applied and longer forecasts up to 48 hours forward in time were launched every sixth hour. The extreme precipitation events all took place on a location within the blue circle within Figure 1 and are described in Table 1. One of the events, studied in more detail, is associated with the huge convective precipitation cell in Northern France, clearly visible in the satellite data shown in Figure 2. At least some of the events were associated with flooding and are as well of importance for a hydrological modelling perspective. Here we focus on the importance of surface data assimilation of the HARMONIE-AROME modelling system for prediction of the precipitation events.

Figure 1: HARMONIE-AROME Modelling domain (red frame) and area of studied extreme precipitation events (blue circle).

Table 1: Description of extreme precipitation events studied with HARMONIE-AROME modelling system.

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-19 June 2013</td>
<td>Total rainfall of 110-180 mm or more in the higher altitudes in less than 48 hours in the Pyrenees area in south-west France.</td>
</tr>
<tr>
<td>21-28 July 2013</td>
<td>A huge convective precipitation cell crosses Northern France during the night between the 26 and 27 July 2013.</td>
</tr>
<tr>
<td>22-25 June 2014</td>
<td>Heavy precipitation event in the in the Pyrenees area in south-west France. Total precipitation of 38 mm in 1 hour at the commune of Mauléon and 70-100 mm over a few hours in the same area.</td>
</tr>
</tbody>
</table>
Figure 2: Channel 9 infrared brightness temperature at 27 July 03.30 UTC derived from the SEVIRI instrument. The low values over Northern France are associated with a convective precipitation cell over Northern France.

4 Enhancements in surface data assimilation of temperature and soil moisture

Overview

The default surface data assimilation of the HARMONIE-AROME modelling system, described in section 2 has been enhanced in three aspects. These three aspects are related to an improved horizontal and vertical distribution of temperature and moisture information in the surface data assimilation, as well as to the utilization of satellite based soil moisture data. The enhancements are described in detail in the three sections below.

Horizontal distribution

The choice of uniform spatial correlation used in the reference setup has clear limitations when terrain is hilly and around coastlines. The enhanced background error statistics for temperature and humidity consist in utilizing the correlation function described in Hägglmark et al. (2000), from now-on referred to as MESCAN, and given by:

\[
\text{corr}(r_{ij}, d_p, d_z) = 0.5 \left[ e^{-\frac{r_{ij}}{L}} + \left( 1 + \frac{2r_{ij}}{L} \right) e^{-\frac{2r_{ij}}{L}} \right] F_p(d_p)F_z(d_z) \tag{7}
\]

where \( F_p \) and \( F_z \) are empirical functions describing the difference due to the difference of land-fraction \( (d_e) \) and the difference of height \( (d_z) \) respectively, between points \( i \) and \( j \). As described in Hägglmark et al. (2000) both functions are linear and vary from 1 for \( d_p = d_e = 0 \) to 0.5 for \( d_p = 1 \) and \( d_z \) larger than 500 m, respectively. The functions are described in Hägglmark et al. (2000). The parameter \( L \) characterizes the correlation length scale and has a value of 190 km for both relative humidity.
and temperature at the two metre level. Given the difference in correlation function this corresponds roughly to the values 80 km and 85 km used together with the correlation function specified in equation (2).

**Vertical distribution**

The current reference surface scheme is unable to represent the flow-dependency inherent in the coupled atmospheric/surface system. The enhanced vertical distribution of two metre temperatures and two metre temperature information is based on an extended Kalman-filter (EKF) technique which is considered standard in the theory of nonlinear state estimation (Julier and Uhlmann, 2004). Formally the equation of the analysed state of the linear Kalman-filter is given by equation (1), as for optimal interpolation. For the EKF, however, the components of the Jacobian of the observation operator, $H_{ij}$ are approximated by equation (8).

$$H_{ij} = \frac{\partial y_i}{\partial x_j} \approx \frac{y_i(x+\delta x_j)-y_i(x)}{\delta x_j} \tag{8}$$

Here $y_i$ represents different observations, $x_i$ represents different model state variables and $\delta x_j$ represents a small perturbation of $x_j$. Following Mahfouf et al. (2009) the Jacobians were approximated using equation (8) by carrying out one perturbed run of the surface model for each control variable of the surface data assimilation, which in our case is in the number of four. So that for our system equation (8) requires four extra runs to estimate the state dependent Jacobian. This procedure results in a dynamic, flow-dependent, relation between model state variables and observations, rather than the static values imposed by equations (3-6). The relation between the observations and the model state and the observations is thus flow-dependent in the case of EKF, which is not the case for the default OI system. Like for the case with optimal interpolation, a 3-hour data assimilation cycle was applied for the state variables. With a Kalman-Filter approach one also has the possibility to estimate the analysis error covariance matrix and to propagate it to obtain a flow-dependent background error covariance matrix. For simplicity reasons, we however chose to apply a static background error covariance matrix which implies that our enhanced methodology should be denoted a Simplified Extended Kalman Filter (SEKF). A next step would we to also include a flow-dependent background error covariance matrix and to evaluate its performance. The background errors of the vertical background error covariance matrix are assumed uncorrelated and the standard deviations are given by 0.15 m$^3$/m$^3$ 0.1 m$^3$/m$^3$, 2.0 K and 2.0 K for $w_z$, $w_g$, $T_2$ and $T_s$, using the terminology of equations (3-6). Also, the observation errors are assumed uncorrelated and these are given by 1.0 K and 0.1% for two-metre temperature gridded temperature and two-metre relative humidity, respectively. For satellite-based pseudo-observations of $w_g$, as described below, the observation error standard deviations are set to 0.4 m$^3$/m$^3$. The error specification is taken from Mahfouf et al. (2009).

**Use of satellite-based soil moisture observations**

The utilization of satellite based soil moisture products has potential for obtaining an improved model soil moisture estimation through surface data assimilation. One of such satellite-based soil moisture products, which has been investigated here, is the data from the Advanced SCATterometer (ASCAT) instrument placed onboard the EUMETSAT (European Organisation for the Exploitation of Meteorological Satellites) MetOp-A and MetOp-B polar orbiting satellites, launched in 2012 and 2016 respectively. The
ASCAT instrument is an active real aperture radar backscatter instrument which operates in the C-band (5.255 GHz) using six vertically polarized antennas. The spatial resolution of the level 2 (L2) gridded data (produced using the Discrete Global Grid, DGG) used in this paper is about 25 km. Soil moisture from ASCAT (version W54) is derived using the change detection approach described by Wagner et al. (1999). The derived surface soil moisture, measured to a depth of approximately 0.5 cm to 2 cm, represents the degree of saturation from 0 % (dry) to 100 % (saturated). It is derived by scaling the normalized backscattering coefficients between the lowest / highest values corresponding to the driest / wettest soil conditions. The approach is based on the assumption that over a long period of time (in our case two months) the highest (lowest) observed reflectivity corresponds to the maximum (minimum) soil moisture.

In order to relate the ASCAT level 2 surface top layer soil moisture product to the HARMONIE-AROME model characteristics a pre-processing is applied. This is needed since in the present HARMONIE version we do not have an observation operator that relates the surface conditions to the raw satellite measurement. The pre-processing consists in, for each horizontal position of the model domain with observation coverage, finding the maximum and minimum observed ASCAT surface top layer soil moisture values \((\text{obs}^{\text{prod max}}_{\text{ascat}}, \text{obs}^{\text{prod min}}_{\text{ascat}})\) from data over an extended period of two months. It should be noted that this period was constrained by simultaneously available satellite and model data and is on the shorter side to assure sufficient correspondence between the dynamic ranges of the satellite data and the model. After semi-operational runs have started a more extended data set can be built. A future alternative could also be to directly relate the model soil moisture to the raw backscattered signal from the SCATTEROMETER instrument onboard the satellite. For the corresponding positions and period also the maximum and minimum soil moisture values \((w^{\text{mod max}}_g, w^{\text{mod min}}_g)\) of the uppermost surface level from HARMONIE is calculated. The model values are taken from a two-month model run carried out for the period corresponding to the satellite observations. The maximum and minimum value of the ASCAT level 2 surface top layer soil moisture product and the uppermost surface level soil moisture values from the learning period \((\text{obs}^{\text{prod max}}_{\text{ascat}}, \text{obs}^{\text{prod min}}_{\text{ascat}})\) are then used to convert the surface top layer soil moisture product for a particular time to uppermost level soil moisture pseudo-observation \((w^{\text{ascat}}_g)\). The conversion is according to equation (12) and is graphically illustrated in Figure 3.

\[
w^{\text{ascat}}_g = w^{\text{mod min}}_g + \frac{w^{\text{mod max}}_g-w^{\text{mod min}}_g}{\text{obs}^{\text{prod max}}_{\text{ascat}}-\text{obs}^{\text{prod min}}_{\text{ascat}}} (\text{obs}^{\text{prod max}}_{\text{ascat}} - \text{obs}^{\text{prod min}}_{\text{ascat}}) \quad (12)
\]

Here the model values are taken from the grid-point closest to the position of the observations. The procedure is carried out individually for the ASCAT instruments on the METOP-A and METOP-B satellites, which have different satellites passes. Potential differences in characteristics between ascending and descending satellite passes are implicitly handled by the fact that the conversion is done individually for each horizontal position and each position is generally covered by a satellite observation by either an ascending or a descending satellite passage. A proper statistical relation between the ASCAT level 2 surface top layer soil moisture product and the model relies on a long enough period to cover all realistic soil moisture conditions. The two months used here is based on what was practically feasible and is on the shorter side. In the future, it will be extended based on archived model data.
Deliverable 3.2

Figure 3: Illustration of ASCAT pre-processing for 20130614 09 UTC. ASCAT observation based soil moisture products (unit: %) are converted to $w_{g}^{ascat}$ soil moisture pseudo-observations (unit: m$^3$/m$^3$).

The surface data-assimilation in HARMONIE-AROME is prepared to handle ASCAT-based soil moisture pseudo-observations over each model grid-point from one data source only. This means that merging satellite-based soil moisture pseudo-observations from METOP-A and METOP-B is required prior to data assimilation. In case both METOP-A and METOP-B data exist over the same grid-point only METOP-A data are kept over that grid-point. The merging of satellite data is illustrated in Figure 4, for one particular case.

Figure 4: Merging of ASCAT satellite data from METOP-A (green) and METOP-B (red) for 20140614 09 UTC. The merged data are marked blue.
5 Experimental design

As presented in Section 3, the HARMONIE-AROME configuration of the HIRLAM-ALADIN NWP system has been setup over a Southern European domain to investigate the prediction of four heavy precipitation events as described in Table 1. For each event, four parallel experiments were run, as described in Table 2. For each precipitation event, the experiment named OI was run for a two-week period in a 3-hour data assimilation cycle mode, prior to the experimental periods. The purpose was to spin-up various aspects of the modelling system, such as surface state and bias correction coefficients. Thereafter this spun-up state was copied to the other three experiments (named OI-MESC, EKF-MESC and EKF-MESC-SCAT) which were run in cycled data assimilation mode for a couple of days prior to the period of the experiment. The idea was to have a spun-up initial state consistent with the respective different data assimilation methods of the different experiments. In addition we allow for differences between the three experiments already from the very beginning of the experimental periods.

Table 2: Description of the four experiments applied to each precipitation event studied.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OI</td>
<td>HARMONIE-AROME reference system with surface data assimilation for temperature and relative humidity based on optimal interpolation with isotropic and homogenous structure functions in the horizontal and optimal interpolation in the vertical. Observations of two metre temperature and two metre relative humidity are assimilated.</td>
</tr>
<tr>
<td>OI-MESC</td>
<td>Settings as in experiment OI, except for MESCAN horizontally varying structure functions replacing the isotropic and homogenous structure functions in the horizontal and optimal interpolation.</td>
</tr>
<tr>
<td>EKF-MESC</td>
<td>Settings as in OI-MESC, except for an SEKF being used in the vertical instead of optimal interpolation.</td>
</tr>
<tr>
<td>EKF-MESC-SCAT</td>
<td>Settings as in experiment EKF-MESC, except for that ASCAT soil moisture pseudo-observations from METOP-A and METOP-B were assimilated with SEKF, in addition to observations of two metre temperature and two metre relative humidity.</td>
</tr>
</tbody>
</table>

The idea of having four different parallel experiments is first of all to investigate the sensitivity of the prediction of various heavy precipitation events to different data assimilation enhancements. In addition, the results can give an indication of the sensitivity of the results to variations of the initial state and of the predictability of each event. As shown in Table 1, the three cases that were used to evaluate the enhanced surface data assimilation are associated with heavy precipitation that is studied within IMPREX. The cases cover both synoptically driven precipitation in mountainous areas and convective precipitation in flat terrain. The cases have the potential to be influenced by surface fluxes influenced by the surface state and the surface data assimilation.
6 Results

Demonstration of functionality of enhancements

The three main enhancements introduced here into the HARMONIE-AROME system are presented in Section 4 and are related to: (1) horizontally dependent background error statistics, (2) application of a SEKF for vertical distribution of temperature and moisture information, and (3) assimilation of ASCAT soil moisture data. The functionality of each of these enhancements can be highlighted by studying idealised experiments or by studying individual assimilation cycles.

Figure 5: Impact on the two-metre relative humidity field (unit in plot: 0-1) of one single two-metre relative humidity observation located at a position (marked with a black dot) close to the west-coast of France. The observed relative humidity is approximately 0.15 less than the corresponding model value. Left is with standard structure functions and middle with MESCO structure functions. Right part shows model orography (unit: m).

In Figure 5 the functionality of applying the MESCO structure functions is demonstrated and compared with utilizing the default structure functions. The demonstration is for one single SYNOP Relative Humidity observation at the 2-meter level (unit: 100%). The observation is located close to the west coast of France and the observed relative humidity is approximately 15% less than the corresponding model value. The position of the observation is marked with a black dot. It can be seen how the land-sea mask and the orography influences the horizontal spread of the observational information when the MESCO structure functions are applied. In Figure 6 the impact of one single two-metre temperature observation located in a valley within the Alps is shown. Clearly the distribution of observational information is influenced by orography when applying the MESCO structure functions. The observation represents more the conditions in low level terrain than at the top of the mountains and that is better handled by the MESCO structure functions than by the original OI structure functions.
Figure 6: Impact on the two-metre temperature field (unit: K) of one single two-metre temperature observation located at a position (marked with a black dot) in the Alps. The observed relative humidity is approximately 2 K larger than the corresponding model value. Left is with standard structure functions and middle with MESCAN structure functions. Right part shows model orography (unit: m).

A major advantage of applying a SEKF scheme in the vertical, as compared to applying an optimal interpolation scheme is the non-static couplings between the observations and the surface model state variables. This functionality of the SEKF scheme is visualized in Figure 7, which shows the distribution of Jacobians for all surface model grid-points for an eight day period, 20140621 to 20130628. The distribution is shown for $\frac{\partial RH_{2m}}{\partial w_2}$ (unit: $1/m^3/m^3$) and $\frac{\partial T_{2m}}{\partial T_2}$ (unit: K/K) for both 00 UTC (red) and 12 UTC (blue). Clearly, both at 00 and 12 UTC there is a significant variation of Jacobians within the surface grid-points of the domain. This variation is due to different characteristics in different parts of the domain and also due to different weather situations and surface properties. In addition one can see that the distributions of the Jacobians are different between night and day. For example, $\frac{\partial T_{2m}}{\partial T_2}$ Jacobians are larger during night-time, indicating a stronger coupling when the solar radiation is small and when the transfer of heat from the deep soil to the surface makes a significant contribution. There is also a variation of Jacobians from one day to another, which is not visualized in Figure 7. In the original OI default settings the values shown in the histograms of Figure 7 would have been represented by a single value instead of a wide distribution.

The functionality of assimilating satellite based soil-moisture information from the ASCAT instrument in the form of $w_g$ soil moisture is demonstrated in Figure 8. The increments are presented as percentage of change of the corresponding background value. Shown is the $w_g$ soil moisture increments at 20140625 18 UTC. The increments are due to assimilation of satellite soil moisture information and to the assimilation of two-metre relative humidity and temperature information from synop stations. For $w_g$ soil moisture, however, the increments are larger along the paths of satellite passes. In Figure 8, two such satellite passage paths can be identified. For reference, we show also in Figure 8 the corresponding increments $w_g$ soil moisture increments from the OI reference system, assimilating humidity and temperature information from synop stations only. Note the difference in scales between the OI increments and the increments containing also satellite information. The structure of the ASCAT based increments corresponds to the ones based on synop measurements only in some areas within the swaths, while also differences can be found. The dominating impact of satellite information
for $w_g$ soil moisture data assimilation increments is due to the error specification of the data assimilation and also due to the Jacobians of the Kalman-filter methodology. For $w_2$ soil moisture and also for soil temperatures satellite information is less dominant in the assimilation increments.

Figure 7: Histogram of Jacobians for surface grid-points within the model domain for the period 20140621 to 20130628. Histograms to the left are for $\frac{\partial RH_{2m}}{\partial W_2}$ (unit: 1/m³/m³) and histograms to the right are for $\frac{\partial T_{2m}}{\partial T_2}$ (unit: K/K). Red histograms are for 00 UTC and blue histograms are for 12 UTC.

Figure 8: Data assimilation increments of Soil $w_g$ soil moisture (unit: % of change of the corresponding background value) for experiment EKF-MESC-SCAT (left) and OI (right) at 20140625 18 UTC.
Verification scores

A common way of measuring forecast quality in the NWP community is to compare the forecasts of weather parameters with the corresponding observations at the time for which the forecast is valid. The error, $\varepsilon$, of the forecast of a particular variable $X_f$, for a particular observation site is given by:

$$\varepsilon = X_f - X_o$$  \hspace{1cm} (13)

where $X_o$ is the corresponding observed variable, treated as the truth. For forecasts of surface parameters synop observations are used for verification and for forecasts of upper-air parameters radiosonde observations are used. Two typical statistical measures used are bias and standard deviation of forecast errors. The bias, $\bar{\varepsilon}$, provides a measure of the systematic error in the forecast, as compared with observations, and is defined by:

$$\bar{\varepsilon} = \frac{1}{N} \sum_{n=1}^{N} \varepsilon_n$$  \hspace{1cm} (14)

where $\varepsilon_n$ is the error for one particular variable, for one particular time and for one particular observation site. $N$ is the number of comparisons between model and observation. The standard deviation, $S$, provides a measure of the dispersion of the forecast errors and is given by:

$$S_t = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N} (\varepsilon_n - \bar{\varepsilon})^2}$$  \hspace{1cm} (15)

Usually the statistics is calculated as function of forecast length and presented as an average, either for all observation sites within the model domain, or for the observations within a specific region of interest.

In terms of forecast verification scores, the modifications of surface data assimilation have a considerable impact for the three events studied. In Figure 9 the bias and standard deviation of temperature and relative humidity forecasts for verification against all radiosonde measurements within the entire domain are shown. The scores are accumulated over +12 h and +24 h forecast ranges and are presented for different vertical atmospheric levels, for each event in Table 1. In terms of bias, the differences between the four different experiments are largest at the surface but extend up to a vertical level of 500 hPa. Kalman-filter based experiments result in less biased low level relative humidity forecasts. In terms of standard deviations, the impact of the different surface data assimilation procedures is most evident for relative humidity and it extends up to a vertical level of at least 500 hPa. In terms of bias and standard deviation profiles of temperature and relative humidity, the largest differences between the experiments are for the convective event that took place in Northern Europe between 20130721 and 20130728 (event number 2 in Table 1).

In Figure 10 the corresponding scores for forecasts of 12 h accumulated precipitation verified against all Synop measurements are presented as a function of forecast length between +12 and + 24 hours. Again, the scores are in the form of bias and standard deviation. It can be seen
that the differences in surface data assimilation influence the precipitation scores, in particular for forecast ranges up to +18 hours. In terms of verification scores for precipitation forecasts the experiments with an enhanced surface data assimilation performs better than the default data assimilation.
Figure 9: Bias (left set of curves) and standard deviation (right set of curves) of temperature (left panels, unit: K) and relative humidity (right panels, unit: %) forecasts for verification against all radiosonde observations within the entire model domain. Verification scores are accumulated for +12 and +24 h forecasts and shown for different vertical levels. Different rows are for different cases and different colours are for different experiments.
Figure 10: Bias (upper set of curves) and standard deviation (lower set of curves) of 12 h accumulated precipitation (unit: mm/12 h) for verification against all Synop gauge measurements within the entire model domain. Verification scores are shown as function of forecast length from for +12 to +24 h. Different rows are for different events and different colours are for different experiments.

One should be aware of that the sample size is too small to draw any statistically significant results of whether the surface data assimilation enhancements improve the short-range numerical weather prediction forecasts in general. The purpose here is mainly to demonstrate that the surface data assimilation did have an effect on the verification scores of a number of severe events and has the potential to improve these forecasts. Encouragingly, for the three events studied there was a positive impact of the surface data assimilation enhancements.

Verification for local areas

To really focus on verification of the forecasts within the areas of extreme precipitation two local areas were defined. The Southern local area includes the Pyrenees area and is defined by the box confined by the latitudes 41°N, 47°N and longitudes 5°W, 10°E. The Northern local area includes Northern France and is defined by the box confined by the latitudes 47°N, 56°N and longitudes 5°W, 12°E. Only Synop and radiosonde measurements within the Southern box are used in the local verification of extreme precipitation events in the Pyrenees area and only Synop and radiosonde measurements within the Northern box are used for the local verification of the extreme precipitation event in Northern France. The radiosonde measurements contained within the Northern and Southern boxes are shown in Figure 11.

Figures 12 and 13 correspond to Figure 9 and 10 but using only local data for the verification. By comparing Figures 9 and 12 one can see that differences between experiments in general get larger when verifying in the local area of interest only. This demonstrates that the data assimilation adjustments do have a more noticeable impact in terms of verification scores in the area of the extreme events. Also in the local verification Kalman filter-based forecasts result in less biased relative humidity close to the surface. In terms of relative humidity standard deviation, the benefit of Kalman filter-based forecasts is even more evident in the local verification, except for degradation above 600 hPa in the case for the event observed in June 2014. In Figure 13 one can see that, also for local verification and for the two experiments carried out for the events in June and July 2013 (upper and middle panel), precipitation forecasts were better for all the experiments including the surface data assimilation improvements, as compared with the reference run. However, for the event in June 2014 (lower panel), only the experiment including all the surface data assimilation enhancements showed better scores than the reference experiment.
Figure 11: Location of radiosondes used for local verification of the studied cases. Radiosondes marked with red dots are used for the two cases of precipitation event observed over the Pyrenees area and radiosondes marked with blue dots are used for the precipitation event observed over Northern France.
Figure 12: Bias (left set of curves) and standard deviation (right set of curves) of temperature (left panels, unit: K) and relative humidity (right panels, unit: %) forecasts for verification against radiosonde observations over local areas. Verification scores are accumulated for +12 and +24 h forecasts and shown for different vertical levels. Different rows are for different cases and different colours are for different experiments.
Figure 13: Bias (lower set of curves) and standard deviation (upper set of curves) of 12 h accumulated precipitation (unit: mm/12 h) for verification against Synop gauge measurements over local areas. Verification scores are shown as function of forecast length from for +12 to +24 h. Different rows are for different events and different colours are for different experiments.

Subjective verification of a case study

To get a more detailed insight into how the prediction of a severe precipitation event might be influenced by the surface data assimilation we studied in more detail the event observed during 21-28 July 2013. During this period a huge convective precipitation cell crosses Northern France. Its development in south–western France, development and propagation towards north-eastern France, is shown in Figure 14, which is a satellite multi-sensor precipitation rain/rate estimate. The precipitation rain-rate (mm/hr) is retrieved by combining the infrared brightness temperature derived from the SEVIRI instrument on-board Meteosat-10 satellite and the passive microwave data coming from the SSMIS instrument on board the polar orbiting satellite DMSP F16. The location and movement of the convective precipitation system is visible in Synop gauge 6h accumulated precipitation observations, illustrated in Figure 15.
Figure 14: Multi-sensor precipitation rain-rate (mm/h) for a large convective system passing over France. The estimated rain-rates are for 20130727 19.00 UTC (upper left), 20130628 03.15 UTC (upper right) and 20130728 08.30 UTC (lower left).

Figure 15: Synop gauge 6 h accumulated precipitation (unit: mm/6 h) between 20130626 18 UTC and 2013062700 00 UTC (left) and between 20130627 06 UTC and 2013062700 12 UTC (right).
As a demonstration of how the improved surface data assimilation may lead to an improved short-range precipitation forecast for a severe precipitation event the forecasted 6 h accumulated precipitation valid from 20130627 00 UTC to  20130627 06 UTC is shown in Figure 16. The forecasts are launched at 20130626 06 UTC, which means that the forecast range is +24 h. The forecasts are based on the reference system (OI, left) and the fully enhanced system (EKF-MESC-SCAT, right). The central time of the accumulation interval of the precipitation corresponds very well with the upper right part of Figure 14. The forecasted position of the cell with largest precipitation amounts agrees well with the satellite derived product for the experiment with improved surface data assimilation. The corresponding forecast produced with the reference system is located too far to the north-east and the extent of the area with large precipitation is too large, as compared with the satellite derived product. One should keep in mind, however, that the satellite product shown in Figure 14 is an estimate of instantaneous rain rate while the forecasted precipitation in Figure 16 is accumulated over 6 h, so the amounts are not directly comparable. However, comparing the forecasted amounts of Figure 16 with the corresponding observed amounts, but 6 h later, in Figure 15 gives an indication that the model overestimated the precipitation amounts, for both the reference and the improved system.

Figure 16: Short-range forecast of 6 h accumulated precipitation with reference system (OI, left) and system with enhanced surface data assimilation (EKF-MESC-SCAT, right). The forecasts are launched on 20130726 06 UTC and the accumulation period is between 20130727 00 and 20130727 06 UTC, which means that the forecast are at a range of +24 h.

The difference between the two experiments producing the forecasts of Figure 16 is the surface data assimilation and secondary effects caused by it. In Figure 17, the differences in $w_2$ soil moisture and $T_2$ soil surface temperature between the experiment with improved surface data assimilation and the experiment with reference surface data assimilation are shown for a forecast launched on 20130726 06 UTC. One should keep in mind that there are
as well differences in upper air fields due to influence of the atmosphere of earlier surface states. However differences in $w_2$ and $T_2$ do have a considerable potential to affect the forecast and from Figure 13 it is evident that such differences exist. The experiment with improved surface data assimilation has larger soil moisture values in the area of the development and propagation of the convective system, in Northern France. Also a considerable soil temperature difference exists between the two experiments and with a typical magnitude of 2 K. Overall the soil temperature is lower in the experiment with improved surface data assimilation in the area of the development and propagation of the convective system, in Northern France. However, in some areas close to the Northern coast of France, soil temperatures are larger in the forecast with improved surface data assimilation. Higher soil moisture values and larger soil temperature will enable for larger latent and sensible heat fluxes from the surface to the atmosphere. As a result one could expect more energy to the atmosphere and more triggering of convective precipitation in these areas. Interestingly one can, by comparing Figures 16 and 17, see that the location of maximum precipitation is in an area with increased soil moisture and surface temperatures in the experiment with improved surface data assimilation as compared with the forecast with reference data assimilation. For the forecast with reference data assimilation, the area of large precipitation amounts is extended more to the east, where the surface temperatures are higher than in the experiment with improved surface data assimilation. Thus, for this particular case, the improved surface data assimilation acted to improve the forecasted location and spatial extent of the area of maximum convective precipitation. Figure 18 illustrate the difference in Convective Available Potential Energy (CAPE) between a 6h forecast launched from 20130726 06 UTC, with improved surface data assimilation and with the reference surface data assimilation, respectively. The differences are thus valid for 20130726 12 UTC, during the evolution of the convective system. CAPE is defined as the amount of potential energy an air parcels acquires when lifted adiabatically from its lifting condensation level to the level of neutral buoyancy. Clear differences can be identified, such as larger CAPE values in the experiment with improved surface data assimilation in the area where 18h later the maximum precipitation amounts appear in the run with the improvements.
Figure 17: Initial state differences in $w_2$ soil moisture (unit: $m^3/m^3$) and $T_2$ soil temperature (right: unit: K) between experiments EKF-MESC-SCAT and OI. The differences are for the initial state 20130726 06 UTC.

Figure 18: Differences in Convective Available Potential Energy (unit: J/kg) between 6h forecasts of experiments EKF-MESC-SCAT and OI. Forecasts are launched on 20130726 06 UTC and valid at 20130726 12 UTC.
7 Conclusions

The HARMONIE-AROME limited area modelling system has been enhanced in several aspects with respect to the surface data assimilation. The enhancements have been described in detail and their functionality has been demonstrated. The enhancements concern both improved methodologies for data assimilation and increased observation usage.

The functionality of the enhancements of some key aspects of the surface data assimilation was demonstrated. Awaiting a semi-operational systematic evaluation over a longer time period, we here checked that our adjustments did not negatively impact a main purpose of our NWP system in IMPREX, which is predicting extreme precipitation events. In this study we confirm that our surface data assimilation adjustments do not negatively affect extreme precipitation forecasts, and even tend to slightly improve them, although not in a robust statistical sense. Future work includes extended semi-operational experiments. Also the mismatch in resolution between model (2.5 km) and satellite data (25 km) should be addressed.

Another issue that needs to be considered is the sharp gradients in the analysis increments that arise along the borders of the satellite swath. Furthermore the methodology introduced here should be applied and evaluated together with hydrological models for extreme precipitation events. Such evaluations take place within some of the case study work packages of IMPREX. The immediate implication of this deliverable for these case studies (WP7.1 and WP8.3) is that additional NWP realisations of the events with an incrementally enhanced system are available. These realisations can support estimating the sensitivity of the hydrological forecasts to the surface data assimilation of the NWP model. In a longer perspective, after that a robust statistical evaluation of extended semi-operational experiments has been carried out, the improvements could be implemented in future high resolution NWP systems to potentially improve prediction skills, preferably combined with an ensemble forecasting system to assess uncertainties.

Results and developments presented in this deliverable would potentially benefit several types of users. First of all, soil moisture affects local precipitation (Arnault et al., 2016). Therefore the method of improved soil-moisture handling presented in this deliverable could lead to improved skill in case of locally induced heavy precipitation, when synoptic-scale forcing is weak. A better prediction of such severe events could support water management at hydropower-companies, authorities in cities and farmers to take actions. This could help prevent disturbances in communities, economic losses and even losses of lives. Such heavy precipitation events can cause problems in areas close to small-scale drainage basins, as well as in flatlands and even in larger scales drainage basins, if hydropower magazines are filled up close to their limits. There is also a direct benefit of an improved knowledge regarding soil moisture. For example soil moisture has effects on underground electric cables (Marshall and Fuhrmann, 2015) and how efficient electric power can be transported. An improved soil moisture product could therefore support optimization of electric power transport. One can also speculate that the enhanced soil moisture analysis product presented in this deliverable can be used in future soil moisture re-analysis projects to produce products showing in detail spatial soil moisture variations and the seasonally variability. Such products could be beneficial for agriculture to support the decision on what kinds of crops to grow in different areas and seasons.
IMPREX has received funding from the European Union Horizon 2020 Research and Innovation Programme under Grant agreement N° 641811
8 References


mesoscale analysis system, Tellus A. Dynamic Meteorology and Oceanography, 52, Iss. 1.


Noilhan, J., and Planton, S., 1989.: A Simple Parameterization of Land Surface Processes for Meteoro-


