





Preliminary study for the use of soil moisture information for Nowcasting-Short Range NWP Forecast

Valerio Cardinali

first-year EUMETSAT fellow, COMET

Supervised by: Lucio Torrisi

in collaboration with German Weather Service (DWD)

EUMETSAT Fellow Day, 7 March 2016





Outline

- Ensemble-Based Data Assimilation: LETKF (Local Ensemble Transform Kalman Filter)
- COSMO Priority Project KENDA (Km-scale Ensemble-Based Data Assimilation)
- COMET NWP system
- H-Saf ASCAT Soil Moisture products
- Description of the work done and future developments



- Ensemble-Based Data Assimilation: LETKF (Local Ensemble Transform Kalman Filter)
- COSMO Priority Project KENDA (Km-scale Ensemble-Based Data Assimilation)
- COMET NWP system
- H-Saf ASCAT Soil Moisture products
- Description of the work done and future developments

EUMETSAT Fellow Day, 7 March 2016



Ensemble-Based Data Assimilation (1)

NWP is an initial/boundary value problem: given an estimate of the present state of the atmosphere (<u>initial conditions</u>), and appropriate surface and boundary conditions, the model simulates (forecasts) the atmospheric evolution.

Currently, operational NWP centers produce initial conditions (analysis) through a statistical combination of observations and short-range forecast, approach known as DATA ASSIMILATION



ensemble data assimilation

Main characteristics

- Monte Carlo techniques
- starting point: ensemble of forecasts
- forecast ensemble perturbations used to represent the forecast error

$$\mathbf{P}^{b} = \frac{1}{m-1} \mathbf{X}^{b} \left(\mathbf{X}^{b} \right)^{T} \qquad X_{b} = x_{b} - \overline{x}_{b}$$

analysis ensemble produced

Ensemble-based data assimilation delivers the best estimate and a representation of the probability density function for the atmospheric state



Ensemble-Based Data Assimilation (2)



- flow-dependent error structures
- No Adjoint operator needed
- intrinsically parallel

LIMITATIONS

- sample size (sampling errors)
- model error representation (filter divergence)

$$\mathbf{P}_{t+1}^{b} = \mathbf{M} \mathbf{P}_{t}^{a} \mathbf{M}^{T} + \mathbf{Q}$$



How to count this behavior in practice

Covariance localization techniquesInflation techniques



Hunt et al. (2007)

Operational at COMET (Italian Air Force Operational Met. Center) since 2011 First Met. Center which uses operationally a pure EnKF DA to initialize a deterministic NWP model

Main characteristics

- analysis done in the space of ensemble perturbations
- analysis computed separately for each grid point selecting only the observations in the vicinity. The observation error covariance R elements are modified by distance-dependent localization factors so that far-away observations have large errors. (explicit localization)
- Analysis ensemble members are locally linear combinations of background ensemble members

► do analysis in the *k*-dimensional ensemble space

$$\mathbf{\bar{w}}^{a} = \mathbf{\tilde{P}}^{a} (\mathbf{Y}^{b})^{T} \mathbf{R}^{-1} (\mathbf{y} - \mathbf{\bar{y}}^{b})$$
$$\mathbf{\tilde{P}}^{a} = [(k-1)\mathbf{I} + (\mathbf{Y}^{b})^{T} \mathbf{R}^{-1} \mathbf{Y}^{b}]^{-1}$$

in model space we have

$$ar{\mathbf{x}}^a = ar{\mathbf{x}}^b + \mathbf{X}^b ar{\mathbf{w}}^a$$
 $\mathbf{P}^a = \mathbf{X}^b ar{\mathbf{P}}^a (\mathbf{X}^b)^T$

Now the analysis ensemble perturbations - with P^a given above - are obtained via

$$\mathbf{X}^{a}=\mathbf{X}^{b}\mathbf{W}^{a},$$

where $\mathbf{W}^a = [(k-1) \widetilde{\mathbf{P}}^a]^{1/2}$



• Ensemble-Based Data Assimilation: LETKF (Local Ensemble Transform Kalman Filter)

- COSMO Priority Project KENDA (Km-scale Ensemble-Based Data Assimilation)
- COMET NWP system
- H-Saf ASCAT Soil Moisture products
- Description of the work done and future developments



COSMO Priority Project KENDA

COSMO: <u>Co</u>nsortium for <u>S</u>mall-scale <u>Mo</u>deling (Germany, Switzerland, Italy, Greece, Poland, Romania and Russia)

<u>KENDA</u> (<u>K</u>m-Scale <u>En</u>semble-Based <u>D</u>ata <u>A</u>ssimilation)

TASK: To develop a separate DA scheme for the convective scale (in which conditions such as non gaussianity, strong non linearity, flow dependent and poorly know balance are much more dominant), and to use a similar approach for a generalized system for global and regional modelling.

The main FOCUS of the KENDA project has been on the algorithmic development of the LETKF

Assimilation of conventional observations and (work in progress) high resolution remote sensing data (radiances, RADAR data, screen level observations, ground based GNSS slant path delay, **ASCAT soil moisture**)



• Ensemble-Based Data Assimilation: LETKF (Local Ensemble Transform Kalman Filter)

- COSMO Priority Project KENDA (Km-scale Ensemble-Based Data Assimilation)
- COMET NWP system
- H-Saf ASCAT Soil Moisture products
- Description of the work done and future developments

EUMETSAT Fellow Day, 7 March 2016



COMET NWP system





- Ensemble-Based Data Assimilation: LETKF (Local Ensemble Transform Kalman Filter)
- Priority Project KENDA (Km-scale Ensemble-Based Data Assimilation)
- COMET NWP system
- H-Saf ASCAT Soil Moisture products
- Description of the work done and future developments

EUMETSAT Fellow Day, 7 March 2016



H-SAF ASCAT Soil moisture products (1)

ASCAT soil moisture Data provided by EUMETSAT within the H-SAF project, one of the 8 EUMETSAT SAFs, lead by the Italian Air Force Met Service



From backscattering coefficient measurements it is possible to retrieve the soil moisture content in the first 2 cm below the soil by mean of microwave technique thanks to the high sensitivity of microwaves to the water content in the soil surface layer (for microwave frequencies in the C-band (< 10 GHz) the addition of liquid water to the soil strongly increases the soil dielectric constant, and so the backscattering coefficients).



H-SAF ASCAT Soil moisture products (2)

- soil moisture content
- incidence angle
- σ_0 affected by: land cover (vegetation)
 - surface roughness

CHANGE DETECTION ALGORITHM (TU Wien):

Backscatter measurements are extrapolated to a reference incidence angle (40°) and corrected for the influence of vegetation; then they are compared to equivalent existing wet and dry backscatter reference, also defined at 40°. As a result, time series of the topsoil (2 cm) moisture content are obtained in relative units (degree of saturation)

Basic assumptions:

- 1. The relationship between the backscattering coefficients and the surface soil moisture content is linear.
- 2. The backscattering coefficient depends strongly on the incidence angle.



3. An increase in soil moisture simply shift the backscattering-angle of incidence curve upwards, while a change in vegetation affects its shape. For sparse vegetation, the curve tends to drop off rapidly, while for fully grown vegetation it becomes less steep.

 σ'



- Ensemble-Based Data Assimilation: LETKF (Local Ensemble Transform Kalman Filter)
- COSMO Priority Project KENDA (Km-scale Ensemble-Based Data Assimilation)
- COMET NWP system
- H-Saf ASCAT Soil Moisture products
- Description of the work done and future developments

EUMETSAT Fellow Day, 7 March 2016



DESCRIPTION OF THE WORK DONE ...

- Adaptation of COMET NWP system to KENDA (implementation of KENDA-LETKF code in the COMET NWP system)
- Processing of available satellite soil moisture products: computation and monitoring of ASCAT soil moisture observation increments
- Quality control of ASCAT soil moisture DATA

... AND FUTURE DEVELOPMENTS

• implementation of the soil moisture data assimilation into the KENDA-LETKF analysis algorithm



- Adaptation of COMET NWP system to KENDA (implementation of KENDA-LETKF code in the COMET NWP system)
- Processing of available satellite soil moisture products: computation and monitoring of ASCAT soil moisture observation increments
- Quality control of ASCAT soil moisture DATA
- implementation of the soil moisture data assimilation into the KENDA-LETKF analysis algorithm



Adaptation of COMET NWP system to KENDA

September 2015: Start of the migration from COMET-LETKF code to the KENDA-LETKF code **PRE-OPERATIONAL**

WHAT HAS BEEN DONE ?

- Creation of feedback files (containing observations and observation increments) as KENDA-LETKF input. Both for conventional and satellite observations
- Introduction of a different reference atmosphere
- Introduction of a variable horizontal localization with different length scales for different vertical levels
- Some modifications to run the KENDA-LETKF code in full resolution
- Bug errors fixed

NEXT STEP

Implementation of soil moisture observations assimilation within the KENDA-LETKF code



Adaptation of COMET NWP system to KENDA (implementation of KENDA-LETKF code in the COMET NWP system)

- Processing of available satellite soil moisture products: computation and monitoring of ASCAT soil moisture observation increments
- Quality control of ASCAT soil moisture DATA
- implementation of the soil moisture data assimilation into the KENDA-LETKF analysis algorithm



Transformed SOIL MOISTURE

- ASCAT derived Soil Moisture: <u>degree of saturation</u> (%) in the first 2 cm
- COSMO TERRA_ML model soil moisture: <u>liquid water content</u> (m H₂O) in the various model layers



layer structure of the hydrological part of the COSMO TERRA_ML soil model



CDF matching

To scale the ASCAT derived soil moisture to the model climatology so that the <u>cumulative distribution functions</u> (CDF) of satellite and model soil moisture match.

SAMPLE CDF definition

Let X_1, \ldots, X_n be independent and identically distributed aleatory variables with distribution function (cdf) F(x).

The sample cumulative distribution function is defined as

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n 1(x_i \le x)$$

The concept of CDF was used in similar studies (Drusch et al 2005, Drusch 2007) to effectively remove biases of soil moisture observations.

This method doesn't allow deriving "correct" soil moisture. Rather it removes differences between satellite observations and model data by ensuring statistical consistency.



CDF matching: our implementation

OUR STUDY:

preliminary test on operational COMET NWP configuration (10 km)

- 1 year time series of ASCAT and model SM data (january 2015 january 2016)
- model data from COMET-LETKF system (10 km grid spacing)
- 2 options investigated for the choice of the soil type to assign to an ASCAT observation
 - the soil type of the grid point closest to the observation
 - the most probable soil type among these of the 9 grid points closest to the observation (ASCAT resolution: 25 km, 10 km grid spacing)
- CDF matching performed for each soil type separately
 - COSMO TERRA_ML soil types: sand, sandy loam, loam, loamy clay, clay, peat
- Piece wise sample CDF for ASCAT and model sm data, using 13 percentiles

0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 1

- Linear regression analysis of ASCAT data plotted against model data
 - 2 options investigated:
 - total regression analysis
 - local regression analysis



CDF matching: example



ASCAT sample CDF, loam, closest grid point



model sample CDF, loam, closest grid point



CDF matching: local regression analysis global regression analysis

$$\omega_{obs} = max \left(0, a + b \frac{\theta_{obs}}{100} \right)$$

b slope, a intercept



Normalization methods

$$\omega_{obs} = \omega_{ADP} + \frac{\theta_{obs}}{100} (\omega_{PV} - \omega_{ADP}) \qquad \omega_{obs} = \omega_{ADP} + \frac{\theta_{obs}}{100} \left(\frac{\omega_{PV} + \omega_{FC}}{2} - \omega_{ADP} \right)$$

soil type	1 ice	2 rock	3 sand	4 sandy	5 loam	6 loamy	7 clay	8 peat
				loam		clay		
volume of voids w_{PV} [1]	_	-	0.364	0.445	0.455	0.475	0.507	0.863
field capacity w_{FC} [1]	-	-	0.196	0.260	0.340	0.370	0.463	0.763
permanent wilting point w_{PWP} [1]	_	-	0.042	0.100	0.110	0.185	0.257	0.265
air dryness point w_{ADP} [1]	-	-	0.012	0.030	0.035	0.060	0.065	0.098

soil parameters values in the COSMO TERRA_ML soil model 8 different soil types: ice, rock, sand, sandy loam, loam, loamy clay, clay, peat

<u>Volume of voids</u>: maximum possible volume of water that the soil can hold
 <u>Field capacity</u>: amount of soil moisture held in the soil after excess water has drained away and the rate of downward movement has decreased.
 <u>Wilting point</u>: the minimal amount of water the plant requires not to wilt
 <u>Air Dryness point</u>: minimum possible amount of water that can remain in the soil

Observation increments statistics

The transformed ASCAT soil moisture data has to be compared to the equivalent model values

OBSERVATION INCREMENTS

quantity directly used in the LETKF algorithm

- difference between the observed value and its model equivalent value (ensemble mean)
- first guess values linearly interpolated in time
- 2 methods for space interpolation:
- nearest grid point
- average on the 9 nearest grid points
- model values calculated using the COMET-LETKF system
- 10 km resolution

Because of the assumption of no bias and gaussianity for the ensemble-based DA, their distribution in terms of bias and symmetry will be analyzed



Evaluation of results (CDF) (1)

observation increments (january 2015 - january 2016)





Evaluation of results (CDF) (2)

observation increments (january 2015 - january 2016)



EUMETSAT Fellow Day, 7 March 2016



Evaluation of the results: normalization methods comparison

observation increments (january 2015 - january 2016)



first formulation, soil type most probable among the 9 closest grid points

> BIAS: 0.0004021 SYMM: -0.2597088



second formulation, soil type most probable among the 9 closest grid points

> BIAS: -0.0251703 SYMM: -0.3667752

The first formulation is better in terms of bias and symmetry



Evaluation of the results (first normalization method)

observation increments (january 2015 - january 2016)



Conversely with respect to the CDF matching method, with this normalization the choice of soil type most probable among the 9 closest grid point is better in terms ob bias and symmetry

0.2

0.1



Evaluation of the results (CONCLUSIONS)

In terms of the observation	increments,	the best results	have been
obtained with:			

- CDF matching method soil type of the nearest grid point local regression analysis
- Normalization method
 First formulation
 soil type = most probable among the soil types of the 9 closest grid
 points

BOTH THE METHODS WILL BE TESTED IN THE ASSIMILATION USING THE KENDA-LETKF CODE



- Adaptation of COMET NWP system to KENDA (implementation of KENDA-LETKF code in the COMET NWP system)
- Processing of available satellite soil moisture products: computation and monitoring of ASCAT soil moisture observation increments
- Quality control of ASCAT soil moisture DATA
- implementation of the soil moisture data assimilation into the KENDA-LETKF analysis algorithm



Quality Control before assimilation of ASCAT soil moisture DATA (1)

Soil moisture cannot be estimated if the fraction of dense vegetation, open water, snow/frozen soils, mountains, sand dunes and/or wetlands dominates the scatterometer footprint



ASCAT data is rejected where:

- snow: the analysed snow amount is greater than 0.05 kg/m^2
- *frost*: the 2m Temperature analysis is below 275.15 K
- wetlands: the inundation and wetland amount has a value greater than 15%
- *mountains*: the topographic complexity has a value greater than 20%
- <u>ASCAT estimated error</u>: the error in the ASCAT surface soil wetness is estimated to be greater than 7% (Met Office) or 8% (ECMWF). This check rejects ASCAT data from regions with dense vegetation and sand dunes.



Quality Control before assimilation of ASCAT soil moisture DATA (2)





an observation is discarded if its observation increment is larger (in absolute value) than a value which is typically in a range between 2 and 3 times a typical climatological standard deviation

The standard deviation is calculated considering a long period of data (observation increments) and pulling out the **gaussian distribution** that best fits them

BUT:

The soil moisture's observation increments are **highly non gaussian**, too concentrated around the value 0 (<u>due to the fact that the obs incr are</u> <u>very close to the 0 value in dry and</u> <u>saturated condition</u>)



To avoid discarding good quality observations, a control variable for the soil moisture whose increments have a gaussian behavior could be obtained, so to apply the quality control to this control variable

Construction of a gaussian control variable for the ASCAT soil moisture DATA (1)

<u>method proposed by Holm (2001)</u> <u>to find a variable for humidity with</u> <u>gaussian forecast differences</u>

- Find a variable φ whose forecast difference $\delta \varphi$ follows a gaussian conditional error distribution $P(\delta \varphi | \Phi)$ as a function of some variable Φ ;
- Determine the bias (b(Φ)) and standard deviation (σ(Φ)) of the forecast differences as a function of Φ, with the bias preferably negligible;
- Normalize forecast differences by the bias and standard deviation,

$$\tilde{\delta \varphi} = \frac{\delta \varphi - b(\Phi)}{\sigma(\Phi)};$$
 (1)

• Change the control variable according to equation 1. **IDEA:** to apply the Holm method to the obs increments instead of to the forecast differences

- φ = soil moisture observation
- Φ = average between soil moisture obs and its model equivalent
- Step functions for bias and stdv <u>for each soil</u> <u>type</u>, partitioning the interval between the max and min value of Φ in 40 bins



Construction of a gaussian control variable for the ASCAT soil moisture DATA (2)

obs increments (CDF method) (january 2015 - january 2016)

not normalized obs incr

bias: -0.0012186 stdv: 0.0800698 symm: -0.3240673

normalized obs incr

- Adaptation of COMET NWP system to KENDA (implementation of KENDA-LETKF code in the COMET NWP system)
- Processing of available satellite soil moisture products: computation and monitoring of ASCAT soil moisture observation increments
- Quality control of ASCAT soil moisture DATA
- implementation of the soil moisture data assimilation into the KENDA-LETKF analysis algorithm

Soil moisture assimilation: future developments

1. Use of derived soil moisture increments in the KENDA-LETKF code, to improve the analysis of atmospheric variables in the lowest levels

- Introduction of the soil moisture variable in the state vector
- Use of soil moisture observations to improve the analysis of atmospheric variables in the lowest levels, through a suitable vertical localization
- Implementation of an horizontal localization with a suitable length scale to account for the influence of a soil moisture observation on a close grid point

2. Development of a suitable soil moisture analysis

This step requires further discussions and investigations

3. tuning at higher resolution (2.8 km)

Thanks for your attention!

ACKNOWLEDGEMENTS:

- Christoph Schraff (DWD)
- Andreas Rhodin (DWD)
- Hendrik Reich (DWD)
- Francesca Marcucci (COMET)