

L-BAND DATA ASSIMILATION FOR IMPROVED SURFACE MODELLING

Alexander Loew⁽¹⁾, Irena Hajnsek⁽²⁾, Dirk Hoekman^(3,4), Martin Vissers⁽⁴⁾, Malcolm Davidson⁽⁵⁾

⁽¹⁾ University of Munich, Department Geography, Luisenstr. 37, 80333 Munich / GERMANY, Email: a.loew@lmu.de

⁽²⁾ German Aerospace Center (DLR), Muenchner Str. 20, 82234 Wessling / GERMANY, Email: irena.hajnsek@dlr.de

⁽³⁾ Wageningen University, Nieuwe Kanaal 11, 6709 PA Wageningen, NETHERLANDS, Email: dirk.hoekman@wur.nl

⁽⁴⁾ SARVISION, Agro Business Park 10, 6708 PW Wageningen / NETHERLANDS, Email: vissers@sarvision.nl

⁽⁵⁾ European Space Agency, Mission Experts Division, ESTEC / NETHERLANDS, Email: Malcolm.Davidson@esa.int

ABSTRACT

Accurate land surface process modelling might be limited due to lack of reliable model input data. Key surface variables as land cover information or soil moisture conditions have been proven to be observable by remote sensing systems. The integration of remote sensing data into land surface process models might therefore help to improve their simulations results. Longer wavelength SAR data has a higher sensitivity to soil moisture content than higher frequency systems. Recent (ALOS) and planned (e.g. TerraSAR-L) SAR systems are therefore expected to provide valuable information about soil moisture dynamics.

The present study investigates the potential to retrieve land cover information and geophysical parameters from L-band SAR data. The retrieval results are assimilated into a state-of-the-art land surface model to evaluate the merit of L-band SAR data assimilation.

1 INTRODUCTION

Land surface models (LSM) are widely used to describe processes at the interface layer within the soil-vegetation-atmosphere continuum. They provide time series of spatially distributed simulations of land surface variables as e.g. soil moisture, evapotranspiration, surface runoff, vegetation biomass or leaf area index. Simulations are based on different kinds of input parameters, as soil texture, land use pattern, meteorological data, which are highly variable in space and time and are available with different accuracies at different scales.

The applicability of land surface process models might be limited due to the availability of necessary input data sets, the quality of these datasets and the level of complexity and accuracy of the process description within the model.

Remote sensing is not only the most efficient, in many aspects it is also the only way to quantitatively assess spatial distributions of input parameters, which are needed to model surface dynamics. This is especially the case for regions which lack of regular measurements

of the needed input parameters. Thus, the derivation of land surface parameters from remote sensing data might improve land surface model (LSM) results. LSMs have evolved from the demand of the provision of spatially distributed time series of geophysical parameters for numerous applications like e.g. weather forecasting, climatological studies, flood forecasting, runoff prediction and crop yield prediction [1]

Data assimilation, as the science of merging observations and models, might help to improve LSM simulations, by the integration of remote sensing observations [2],[3]. In general, there are several options to integrate remote sensing data into a LSM. These comprise the determination of model parameters, the update of the model state during run-time, the parameter adjustment through model recalibration and inverse modelling of remote sensing observations using e.g. radiative transfer models [4]-[6].

For the first, land cover is the most important static input parameter for a LSM, as it basically determines how the land surface interacts with the atmosphere. A regular update of model states might be achieved by regular observations of quantitative parameters, as e.g. vegetation biomass or soil water content. Especially sequential data assimilation techniques are widely used to update the model state when quantitative observations are available. These approaches comprise variational data assimilation schemes as well as Monte Carlo based methods like the Ensemble Kalman Filter [6] and have been successfully applied for land surface applications in the past.

The present study investigates the potential of L-band SAR data to derive land cover and geophysical parameters. The requirements and trade-offs in terms of accuracies and system configuration are investigated. The impact of integration L-band products into a LSM are investigated.

2 MODELS AND DATA SETS

2.1 PROMET land surface model

The Process Oriented Multiscale EvapoTranspiration model (PROMET) is a family of land-surface-process-models which describe the actual evapotranspiration and water balance at different scales, ranging from point scale, to microscale and mesoscale modelling [7]. The model consists of a kernel model which is based on five sub-modules (radiation balance, soil model, vegetation model, aerodynamic model, snow model) to simulate the actual water and energy fluxes and a spatial data modeller, which provides and organizes the spatial input data on the field-, micro and macroscale. The simulations are made on an hourly basis. For a detailed description of the model see [7].

2.2 AGRISAR 2006 campaign

The SAR data and ground measurements, used within this study were collected in the frame of the ESA AGRISAR 2006 campaign [8]. The campaign aimed to collect in-situ, airborne SAR and optical measurements as well as satellite measurements to address specific programmatic needs for the planned SENTINEL 1 and 2 missions of ESA.

The campaign was conducted in the period from 18. April to 2nd of August 2006 at the DEMMIN (Durable Environmental Multidisciplinary Monitoring Information Network) test site in Northern Germany. Main crop types in the area are winter wheat, barley, maize, rape and sugar beet.

2.2.1 Ground measurements

Comprehensive ground data was collected during the AGRISAR campaign, comprising vegetation, soil and atmospheric characteristics. For details see [8].

The energy and water fluxes were measured, using micrometeorological stations, a Bowen-Ratio Energy balance station (BRS) as well as a Large Area Scintillometer (LAS). Soil water fluxes were measured at different depths using TDR probes.

2.2.2 Intercomparison of model and measurements

The PROMET model was used to simulate the soil moisture evolution and evapotranspiration rate during the investigation period. The model simulations are compared against the soil moisture measurements as well as the results obtained from the BRS and LAS. A very good agreement ($R^2=0.97$) is obtained for the net radiation, which determines the available energy for latent and sensible heat flux. The simulated soil moisture fluxes are also in considerable agreement with the measurements as can be seen in Figure 1. The soil moisture dynamics of the upper and lower soil layer are well reproduced by the PROMET model.

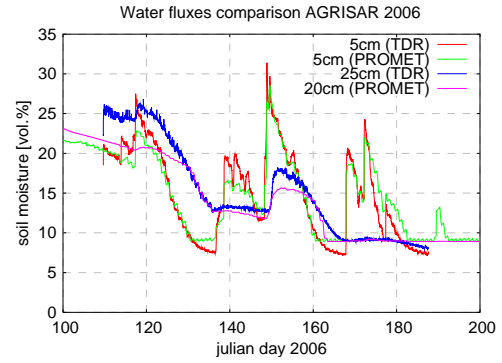


Figure 1: Comparison of simulated and measured SM

2.2.3 Remote sensing data

SAR data was collected at the Demmin test site in the period from April to July 2006 by the E-SAR system of the German Aerospace Center (DLR) [8], which is a polarimetric, multifrequency SAR system. Only the L-band (quad-pol) data is used for the investigations within this study.

3 LAND COVER CLASSIFICATION

Versatile, robust and computational efficient methods for radar image segmentation, which preserve the full polarimetric information content, are of importance as research tools as well as for practical applications in land surface monitoring. The method introduced in [15] consists of six steps. The first step is a transform of the full polarimetric information into nine intensities per band [16]. The next steps relate to unsupervised clustering (i.e. deriving class statistics) encompassing a simple region-growing segmentation (incomplete and over-segmented) followed by model-based agglomerative clustering and expectation-maximization on the pixels of these segments. Classification is achieved by Markov Random Field filtering on the original data [17]. The result is a series of segmented maps, which mainly differ in the number of (unsupervised) classes. Results for the DEMMIN agricultural area is shown. The applications include the use of groundtruth for legend development, the check for groundtruth completeness and the construction of a bottom-up hierarchy of the characteristics that can be distinguished in the radar data. The latter gives important insights in physics of radar backscattering mechanisms. Moreover, the relative importance of crop differences, (full-polarimetric) incidence angle effects and sub-classes (related to factors such as crop varieties or row direction) may be assessed. Figure 2 shows the multi-temporal supervised classification result for DEMMIN with the associated error map in Figure 3. The total error in this classification is low. The main error sources are the class urban, which is not a homogeneous class, and which is often misclassified as

forest. For the purpose of this project this is no problem since the urban areas are already known, only the crop types have to be classified since they rotate each year. The main errors are found for the classes field grass, cutting pasture and grassland. All these grass classes are classified as wheat. Apparently they are very similar in this period of observation when using these 3 polarizations. Another problem is that half of the class set-aside rape is classified as rape.

The results in Figure 4 shows an unsupervised classification (a model with 20 classes) with several classes associated with rapeseed: class 0 (red), class 4 (purple), class 10 (orchid), class 14 (coral) and class 17 (yellow). Class 5 (cyan) is associated with set-aside rape. Since the different sub-classes of rapeseed mostly follow field boundaries it may be concluded there are differences between these fields which could be associated with differences such as growth stage, biomass, crop phenology or crop varieties.



Figure 2: Demmin test site supervised classification of multi-temporal ESAR images using 3 dates with 3 polarizations (HH, VV and PL)

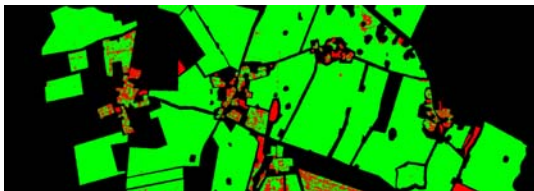


Figure 3: Demmin test site error map for supervised classification result (Green: correct; Red: false; Black: no ground truth)

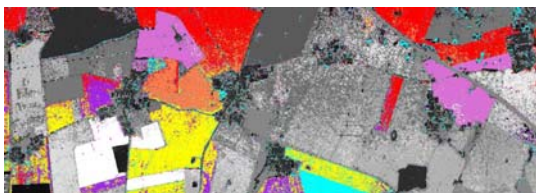


Figure 4: Demmin test site unsupervised classification of multi-temporal ESAR images using 3 dates and 3 polarizations (HH, VV and PR).

4 GEOPHYSICAL PARAMETER RETRIEVAL

Different approaches for the retrieval of geophysical parameters have are reviewed within the scope of the present study. As L-band SAR data has a better penetration into the surface soil layer and through the canopy as higher frequency systems, the study will focus on the retrieval of soil moisture information for bare soil as well as vegetated surfaces. The microwave signatures from the AGRISAR data set have been analysed and the sensitivity of different parameter retrieval techniques have been investigated [8]. The impact of different sensor configurations on the accuracy of the soil moisture retrievals from L-band SAR data will be assessed. The impact of the uncertainties in the SAR derived soil moisture on land surface model simulations will be assessed in section 6.

5 L-BAND DATA ASSIMILATION

Data assimilation is the science of merging observations with a model. Assimilating external observations into a model might either result in a different (better) parameterization of the model static parameters (e.g. land cover, soil parameters) or might improve the model simulations during the model evolution (e.g. update of soil moisture) to improve the model predictions. Within the present study, we deal with both cases. The SAR derived land cover information is used to evaluate the potential of L-band SAR land cover classification for hydrological simulations on the one hand side and the potential of integrating geophysical parameters as e.g. soil moisture is assessed on the other hand.

5.1 Integration of SAR based land cover data

To assess the impact of SAR based land cover classification on PROMET model simulations, a sensitivity analysis was conducted. As the land cover type is a very important parameter for the SVAT model, a misclassified land cover will directly reflect on the estimated soil moisture fluxes and evapotranspiration rates.

The PROMET model is used to simulate land surface processes in the period from 1st January 2006 to 31th July 2006 using a reference land cover map. This open-loop run is then compared against the results obtained using the SAR based land cover classification as input to the SVAT model.

The difference between both simulations is estimated for each grid node individually, comparing the soil moisture for three soil layers and the evapotranspiration rate. The rms error, model efficiency and coefficient of determination is estimated for each grid cell, whereas the open loop run is considered as the reference.

5.1.1 Impact on evapotranspiration

As the land cover was correctly classified for large parts of the image data set (Figure 3), no impact on the PROMET simulations can be observed. Only for ~5% of the image pixels, a difference in the evaporation is observed. The cumulated evaporation difference is estimated from the entire period for each grid cell. The observed deviations in ET range between 18 mm and 20 mm, which corresponds to ~25 % of the precipitation within the observation period. Figure 5 shows the frequency distribution of the cumulated ET difference. The rms error of the evapotranspiration is less than 1 mm for all grid cells.

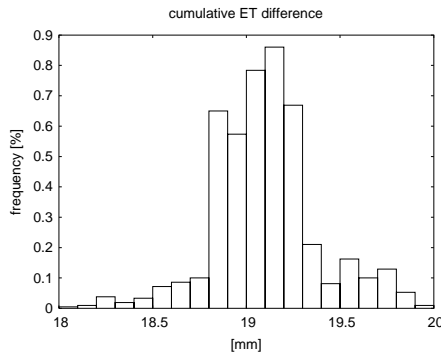


Figure 5: Difference in evapotranspiration

5.1.2 Impact on soil moisture

Soil moisture rms error is between 1 and 4 vol.% for those pixels, where the SAR based land cover was wrong.

5.2 Assimilation of geophysical parameters

Various studies have investigated the potential to assimilate hydrological variables into land surface process models, including the assimilation of soil moisture and crop properties [9].

Given a time variant (physical) model

$$\frac{\partial \psi}{\partial t} = f(\psi) + \varepsilon \quad (1)$$

it is the objective of a data assimilation technique to make “best” use of the observation d_t (e.g. soil moisture, vegetation biomass, reflectance, backscattering coefficient) to improve the model simulation. The model state vector ψ comprises the entire physical state of the model (e.g. soil moisture, canopy temperature, snow depth ...) at a given time. As both, model simulations and observations might be prone to uncertainties, the model results and observations have to be weighted in a certain way. Different data assimilation techniques have been developed, ranging from statistical rescaling methods to Monte Carlo based methods as e.g. the Ensemble

Kalman Filter (EnKF). The EnKF approach was implemented for the present study. Its theoretical foundation is described in detail in e.g. [10]-[12]. The EnKF has been applied successfully to numerous data assimilation problems in hydrological sciences.

It allows to estimate a best update of a model forecast, given an observation d , whereas uncertainties of the observations as well as of the model forecast are considered appropriately. The model simulation uncertainties are derived from its error covariance matrix which is estimated from perturbed model trajectories which form a distribution of different model predictions. All members of this ensemble are updated when an observation is available.

6 SYNTHETIC EXPERIMENT

Assimilating surface soil moisture into a land surface model (LSM) is expected to improve the simulation of soil moisture fluxes and latent and sensible heat fluxes. A synthetic study was conducted to evaluate the potential to assimilate *surface soil moisture* information into the PROMET model and to assess impact of uncertain soil moisture observations as well as the need for the temporal repetition frequency of the observations.

6.1 Generating reference conditions

Two baseline simulations are generated using the PROMET model. The first is an open-loop run where the SVAT model results rely on the meteorological forcing and model parameterisation only. Thus the open-loop represents the model results as they are obtained without any noise or exterior information added to the model state. This open-loop is considered as a first guess, compared to a true surface condition.

As it is assumed that the open-loop results will typically deviate from real conditions due to uncertainties in the model parameterisation and limitations in model physics, a synthetic *true* surface condition data set, which deviates from the open-loop simulation is generated based on the open loop results. Random model error (gaussian white noise) are added to the model state during the simulations to perturbate the model simulations. The resulting model trajectory is considered as the *true* surface condition in the following and will be the reference benchmark for the evaluation of the merit of assimilating surface soil moisture information into the process model. Figure 6 shows examples of perturbed surface soil moisture compared to the open loop condition. The rms error between the open loop run and the *true* state is 3.5 vol.%.

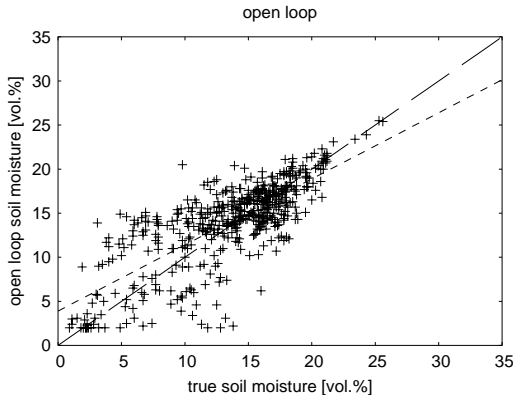


Figure 6: Comparison of true and open loop SM

6.2 Generation of sensor observations

Synthetic sensor observations are simulated based on the true surface state. Remote sensing derived surface soil moisture is expected to have uncertainties in the order of 2.0 – 7.0 vol.% (rms error) [13]-[14].

Random gaussian error of 1, 2.5 and 4 vol.% is therefore added to the true surface soil moisture state, to cope observations, like they are available from remote sensing systems.

6.3 Surface soil moisture assimilation

The synthetic surface soil moisture observations are then assimilated into the PROMET model. The model is used in its generic mode, which is comparable to the open loop case. Surface soil moisture is assimilated using different temporal frequencies: daily, 7-days, 14-days. The simulated soil moisture for the assimilation case is then compared against the true reference soil moisture. Figure 7 shows the comparison between the true soil moisture and the assimilation results obtained for the observations with an rms error of 2.5 vol.%, using daily and 7-day assimilation steps. It is obvious, that the data assimilation improves the model simulations compared to the open-loop simulation. The rms error between the reference data set and the assimilation data set is 1.7, 3.1 and 3.4 vol.% for the daily, 7day and 14day soil moisture assimilations respectively. Thus, the rms error is improved especially for daily and weekly observations, compared to the rms error of 3.5 vol.% for the open-loop run. Similar results were also obtained for the observations with 1 and 4 vol.% rms error.

Table 1 summarizes the rms error and coefficient of determination (R^2), calculated between the reference data set and the assimilation data sets. In all cases, the performance of the assimilation results is better than that of the open loop run. Thus, assimilating surface soil moisture provides an additional merit to the LSM. Nevertheless, the results indicate, that an accuracy of

4 vol.% and nearly weekly observations might be required to improve LSM simulations significantly. The results from the 14-day assimilations and 4 vol.% error do not add an additional value to the LSM.

7 SUMMARY AND CONCLUSIONS

Low frequency SAR systems as e.g. L-band have high potential in the retrieval of geophysical parameters as e.g. soil moisture and derivation of land cover information. The preliminary results of the present study show a good land cover classification performance for supervised and unsupervised approaches. Assimilating surface soil moisture information into LSMs has potential to improve the estimation of land surface energy and water fluxes. The SAR based soil moisture observations are expected to have accuracies better than 4 vol.% (rms error) and a temporal frequency better than a week to improve the LSM results.

Table 1: Assimilation error statistic

Obs. error [vol.-%]	1-day		7-day		14-day	
	rms	R^2	rms	R^2	rms	R^2
Open-loop	3.49	0.58	3.49	0.58	3.49	0.58
1.0	1.21	0.95	2.83	0.73	3.20	0.64
2.5	1.72	0.89	3.14	0.65	3.40	0.60
4.0	2.20	0.81	3.20	0.62	3.47	0.57

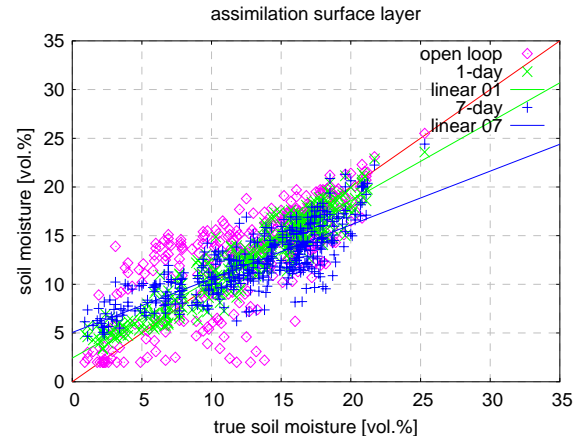


Figure 7: Comparison of true soil moisture vs. assimilation results obtained for different observation frequencies and an observation error of 2.5 vol.%

ACKNOWLEDGEMENTS

The present study was funded under ESA contract 19569. The remote sensing data and ground measurements were conducted within the frame of ESA AGRISAR 2006 campaign (No. 19974/06/I-LG).

8 REFERENCES

- [1] Capehart WJ, Carlson TN. (1994): Estimating near-surface soil moisture availability using a meteorologically driven soil-water profile model. *Journal of Hydrology* 160: 1–20
- [2] Bach, H., Mauser, W. (2003): Methods and Examples of Remote Sensing Data Assimilation in Land Surface Process Modelling. *IEEE Trans. Geosci. Remote Sensing*. Vol. 41 (7). pp. 1629-1637.
- [3] Reichle, R.H., Entekhabi, D., Mclaughlin, D.B. (2001): Downscaling of radio brightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach. *Water Resources Research*, Vol. 37 (9), pp. 2353-2364.
- [4] Bach, H., Verhoef, W., Schneider, K. (2000): Coupling remote sensing observation models and a growth model for improved retrieval of (geo)biophysical information from optical remote sensing data. *Remote Sensing for Agriculture, Ecosystems and Hydrology*, SPIE Vol. 4171. pp. 1-11.
- [5] Parde, M., Wigneron, J.P., Chanzy, A., Waldteufel, P. Kerr, Y., Huet, S. (2003): Retrieving surface soil moisture over a wheat field: comparison of different methods. *Remote Sens. Environ.* Vol. 87. pp. 334-344.
- [6] Walker, J. P., and P. R. Houser, 2001: A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations. *J. Geophys. Res.*, 106, 11761–11774.
- [7] Mauser, W., Schädlich, S. (1998): Modeling the Spatial Distribution of Evapotranspiration on Different Scale using Remote Sensing Data, *J. Hydrol. (Special BAHC Issue)* 212-213.
- [8] Hajnsek, Irena; Horn, Ralf; Scheiber, Rolf; Bianchi, Remo; Davidson, Malcolm (2007): The AGRISAR Campaign: Monitoring the Vegetation Cycle using Polarimetric SAR Data. This issue.
- [9] Reichle, R., Walker, J., Koster, R., Houser, P. (2002): Extended versus Ensemble Kalman Filtering for Land Data assimilation. *Journal of Hydrometeorology*. 3. 729-740.
- [10] Evensen, G. (1994): Sequential data assimilation with a non-linear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *J. Geophys. Res.* 99 (C5). 10143-10162.
- [11] Burgers, G., Leeuwen, P., Evensen, G. (1998): Analysis scheme in the ensemble Kalman filter. *Mon. Weath. Rev.* 126. 1719-1724.
- [12] Evensen, G. (2003): The ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*. 53. 343-367.
- [13] S. Le Hégarat-Mascle, M. Zrib, F. Alem, A. Weisse, C. Loumagne, “Soil moisture estimation from ERS/SAR data: Toward an operational methodology. *IEEE Trans. Geosci. Remote Sensing*. vol. 40, no. 12, pp. 2647-2658, 2002.
- [14] Loew, A., Ludwig R., Mauser, W., “Derivation of surface soil moisture from ENVISAT ASAR WideSwath and Image mode data in agricultural areas”. *IEEE Trans. Geosc. Rem. Sens.. IEEE Trans. Geosc. Rem. Sensing*. vol 44(4). 889-899, 2006.
- [15] Hoekman, D.H., T. Tran, and M.A.M. Vissers, 2007, Unsupervised full-polarimetric segmentation for evaluation of backscatter mechanisms of agricultural crops, *Proc. of POLinSAR 2007 Workshop*, , ESA SP-644,22-26 Jan 2007, Frascati , Italy, CD-ROM (8 pages).
- [16] Hoekman, D.H., and M.A.M. Vissers, 2003, A new polarimetric classification approach evaluated for agricultural crops, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.41, No.12, December 2003, pp.2881-2889.
- [17] Tran, N.T., R. Wehrens, D.H. Hoekman, and L.M.C. Buydens, 2005, Initialization of Markov Random Field Clustering of Large Remote Sensing Images, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.43, No.8, August 2005, pp.1912-1919.