Data Assimilation of Satellite Soil Moisture Retrievals in the Land Surface Model SURFEX

Åsmund Bakketun

Thesis submitted for the degree of
Master in Geosciences
60 credits

Department of Geosciences
Faculty of mathematics and natural sciences

UNIVERSITY OF OSLO

Autumn 2017

Data Assimilation of Satellite Soil Moisture Retrievals in the Land Surface Model SURFEX

Data Assimilation of Satellite Soil Moisture Retrievals in the Land Surface Model SURFEX

http://www.duo.uio.no/

Printed: Reprosentralen, University of Oslo

UNIVERSITY OF OSLO

Abstract

The Faculty of Mathematics and Natural Sciences
Department of Geosciences

Master in Geosciences

Data Assimilation of Satellite Soil Moisture Retrievals in the Land Surface Model SURFEX

by Åsmund Bakketun

Soil moisture plays an essential role in the land surface energy balance, apportioning the available energy into latent and sensible heat. As an important boundary for the atmosphere, accurate initialization of the land surface can help numerical weather prediction (NWP)-models increase their skill. The aim of this study is to identify the potential of integrating satellite derived soil moisture products in a land surface model over Norway. Offline simulations with the SURFace EXternalisée (SURFEX) land surface model are performed, forced by output from an operational NWP-model. Experiments cover the three summer months June, July, August (JJA) 2016, for a domain in south-eastern Norway. Level 3 data from the Soil Moisture and Ocean Salinity (SMOS) and Soil Moisture Active Passive (SMAP) satellites are, separately, integrated in soil moisture analyses using a simplified extended Kalman filter (SEKF). In order to explicitly represent the retrieved soil moisture variable, the multilayer diffusion version of Interaction Sol-Biosphère-Atmosphère (ISBA) is used. Analysis is performed in the upper seven model soil layers to cover the root zone.

The satellite data are found to be of questionable quality for the area of interest, with poor spatial coverage and noisy signal. Investigation of the linearized observation operator reveals that more weight is given to surface layers in dry conditions, and deep layers in wet conditions. Over all water is added to the surface layers, and removed below $10\,\mathrm{cm}$. Layers below $40\,\mathrm{cm}$ have limited communication with the observed layer during a 6 hour assimilation window, and are suggested to be removed from the control vector. In comparison with observations of 2 meter temperature, simulations with data assimilation (DA) of soil moisture show limited improvement over the control. However, the change in surface fluxes is more pronounced, mainly adjusting the Bowen ratio. Since the simulations are performed offline, the full potential of soil moisture increments are not believed to be reached.

Acknowledgements

First of all I would like to thank my supervisors. Jørn for suggesting land data assimilation, a hot topic in numerical weather prediction, and letting me participate in a land data assimilation workshop in Zagreb, introducing me to the community. William for helping with the motivation of the work, providing literature, and introducing me to the people at NILU and cake Friday. Frode for numerous discussions and great help with the structuring of my thesis. Then Jostein for all advice and tips along the way, commenting on my thesis and also being a great traveling companion.

Mariken, Trygve and Hanneke at Met Norway have also been of great help through frequent discussions, helping with technical issues, and introducing me to the operational aspect of this topic. Thanks to my sisters, Ylva and Andrea, for reading through my thesis and correcting my language, Kristian for sharpening my table tennis skills, and my sweet Ellen for keeping our home and family on track while I have been struggeling with satellites, surface models and data assimilation.

Contents

Chapter 1

Introduction

Early day weather forecasts were purely based on observations. In the twentieth century, Vilhelm Bjerkens postulated that the weather could be predicted if: the state of the system is known with sufficient accuracy at a given time, and the laws describing the development from one state to another are known with sufficient accuracy (Bjerknes, 1904). A weather forecast is thus an initial value problem, known as prediction of the first kind. Edward N. Lorenz later found that the equations describing fluid motion, hence the weather, were chaotic, and that prediction of the weather is limited to a certain lead time (Lorenz, 1963). The predictability limit is scale and situation dependent. Large scale systems are limited to a couple of weeks and afternoon showers to a couple of hours. With increasing lead time, climatic time scales, the boundary conditions play a bigger role, where e.g. distance to the sun and green house gas concentrations influence the outcome. These are predictions of the second kind. In decadal prediction, both initial and boundary conditions are important (Hawkins and Sutton, 2009).

Over the years, descriptions of the system processes have become increasingly detailed and super computers allow the governing equations to be solved[1] with fine resolution. Since the system is chaotic, hence of limited predictability, frequent initializations of the model state are required to not diverge from the true state of the system. The initial state should be as close to the real state as possible to maximize the predictability limit.

This thesis will center around the first kind of prediction, namely the initial value problem. Short and medium range weather predictions, half to two days and ten to fifteen days respectively, will be the focus area.

Observations of the earth system, including temperature, pressure, radiation, precipitation, and wind velocity, have the goal to represent the current state as accurate as possible. Ranging from local in situ stations to satellites covering the entire globe, numerous quantities are measured. These observations are integrated in NWP systems, through DA. Improvement of the European Centre for Medium-Range Weather Forecasts (ECMWF)-model system is shown in Fig. 1.1, where the impact of satellite observations is pronounced in the beginning of the twenty-first century.

[1]The Navier-Stokes equations are yet to be solved analytically, but numerical schemes can approximate a solution given initial and boundary conditions.

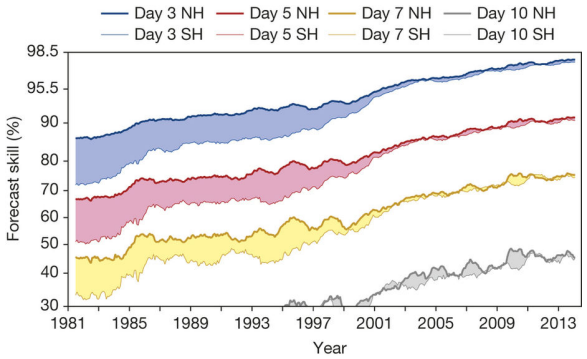


FIGURE 1.1: Improvement in forecast skill of ECMWF (Fig. 1 in Bauer, Thorpe, and Brunet (2015)). Thick line indicates the northern hemisphere, thin line is southern hemisphere. Differences between the hemispheres are caused by the uneven distribution of in situ stations. The revolution of satellite observations after 1999, closes the gap between southern and northern hemispheres.

Soil moisture, plays an important role in the hydrological cycle, influencing the surface energy balance, and is thus an important boundary condition for the atmosphere (Mueller and Seneviratne, 2012; Entekhabi et al., 1999). Soil moisture has received a lot of attention by earth system scientists the last decades, and it is recognized as an essential climate variable. Soil moisture deficit contributes to positive feedback mechanisms for surface temperature extremes on climatic time scales (Seneviratne et al., 2010), and affect surface temperature, and convective cloud formation in short to medium range weather forecasts (Beljaars et al., 1996; Chen and Avissar, 1994). Land surface models have gone from simple conceptual models to complex schemes, representing soil texture heterogeneity, evolution of vegetation, and soil moisture profiles (Balsamo et al., 2009; Rosnay et al., 2002).

Along with better models, initialization of surface parameters have shown to improve forecast skill in both short and medium range weather forecasts (De Rosnay et al., 2014; Hurk, Ettema, and Viterbo, 2008; Drusch and Viterbo, 2007; Douville et al., 2000) and at seasonal scale (Weisheimer et al., 2011; Koster et al., 2011; Koster et al., 2004). In recent years, soil moisture research satellites have been launched, and they provide daily soil moisture maps covering most of the globe.

The main motivation for this work is to improve the land surface boundary conditions used in an NWP-model. By combining model and satellite soil moisture, a more accurate description of the surface state is achieved, which in theory will improve the forecast skill. However, an improved soil moisture estimation can also be valuable for flood warning, agricultural purposes, and for power companies.

Meteorological Co-operation on Operational NWP (MetCoOp) produces operational short-range weather forecasts four times daily, covering northern Europe with the NWP-model Applications of Research to Operations at MEsoscale (AROME)-MetCoOp (Müller et al., 2017). The surface boundary is governed by the externalized surface model SURFEX (Le Moigne, 2012). In the land surface DA, soil moisture and temperature are updated based on observations of screen level variables, temperature and humidity.

In this study, soil moisture retrievals from the SMOS and SMAP satellites are assimilated into decoupled SURFEX simulations, during the 2016 summer months in south-east Norway. The quality of satellite products over high latitudes is investigated, the SURFEX Offline Data Assimilation (SODA) system is tested in conjunction with the ISBA-DIFfusion (DIF) model, and initialization of the surface model is analyzed with respect to model improvement of screen level variables, 2 meter temperature and humidity.

The DA is performed using an Extended Kalman Filter (EKF) (Gelb, 1974) implemented in the SODA environment, updating soil moisture in surface soil and root zone.

This thesis consists of three parts. In part 1, the theoretical basis for modeling, observing, and DA of land surface parameters is presented. This include land surface energy balance and surface hydrology, L-band radiometry of soil moisture, and variations of the sequential Kalman filter (KF).

Part 2 provides description of methodology used in the work. A presentation of the SURFEX model is given, followed by an introduction of satellite products and how they are processed prior to the DA. The SODA environment is presented and modifications to it are described.

In Part 3, results are presented and discussed. Model performance is validated, the quality of satellite retrievals and normalization process are evaluated, soil moisture analysis with ISBA-DIF is verified, followed by evaluation of model improvements due to soil moisture increments with respect to screen level variables.

Part I

Theory

Chapter 2

Land Surface Processes

2.1 Land Surface Energy Balance

The land surface interacts with the atmosphere, and is crucial for the evolution of the planetary boundary layer. In NWP the surface is particularly important, since most of its users are located on the surface and are interested in quantities located at screen level, 2 meters above the surface. These parameters are to a large degree affected by the surface conditions. The surface state is shown to be important for convective cloud formation and precipitation (Beljaars et al., 1996; Chen and Avissar, 1994). Surface conditions have also shown to be important for large scale circulation (Miyakoda, Hembree, and Strickler, 1979) and extreme heat waves (Jaeger and Seneviratne, 2011).

In this chapter, the theory of the land surface energy balance is covered, formulations and notation are adopted from John M. Wallace (2006).

The land surface receives long-wave radiation from the atmosphere, $F_L \downarrow$, and shortwave solar radiation, $F_s \downarrow$. The latter only during the day. Long-wave radiation is emitted from the ground, $F_L \uparrow$, following Stefan Boltzman law [1], and some of the incoming solar radiation is reflected back, $F_s \uparrow$, depending on the surface type and condition. The net radiation is the sum of all these formulated in Eq. (2.1), and illustrated in figure 2.1.

$$F_{net} = F_s \downarrow - F_s \uparrow + F_L \downarrow - F_L \uparrow \tag{2.1}$$

In clear sky weather, the net radiation is dominated by solar radiation and is thus positive. The upward shortwave radiation is proportional to the downward shortwave radiation, and the fraction is known as the surface albedo with values between 0 and 1. The net long-wave radiation is downward for cooler surface than air and upward for the opposite.

If we consider the land surface as very thin, we can assume a heat capacity of zero. The net radiation is then balanced by sensible, F_{Hs}, latent, F_{Es}, heat flux and conductive ground flux, F_{Gs}.

$$F_{net} = F_{Hs} + F_{Es} + F_{Gs} \tag{2.2}$$

When two bodies are in contact and have different temperatures, heat is conducted from the warmer body to the cooler. This process is formulated by Fourier [2], and is a diffusive process. If e.g. the land surface is warmer, the air will be heated and increase temperature while the surface undergoes cooling. If the air does not move, it will continue to warm until equilibrium

[1] black body radiation: $B = \sigma T^4$, σ is Boltzmans constant, and T is the temperature.
[2] Fourier's law: $\mathbf{q} = -k\nabla T$, k is a constant, and T the temperature.

is reached. However, air is moving and turbulent eddies, vertical circulations, are constantly mixing the air within the boundary layer, sustaining the temperature gradient between surface and air, and thus the sensible heat flux. In fair weather conditions, blue skies, the sensible heat flux is usually positive, upwards, during day, where solar radiation heats the surface more rapid than air. During night, outgoing long-wave radiation is emitted more efficiently by the ground than air, and the resulting heat flux is downward.

If liquid water is available, a portion of the available energy will cause evaporation rather than heating. The energy used to evaporate liquid water is contained in the water vapor and is not released until the vapor condenses in a cloud or as dew on vegetation and ground. This is known as latent heat, and the energy flux as latent heat flux. The necessary conditions for evaporation/condensation are slightly more complex than sensible heat flux. The amount of water in the ground/vegetation, and the relative humidity of the air, also have to be included in the process, together with temperature and radiative forcing. The specific heat capacities of dry air C_{air}, water C_{wat} and the specific latent heat of vaporization L_v is given in table 2.1. Notice that the amount of energy needed to increase the temperature of one kilo water by one Kelvin degree, is four times what is required for dry air. Further, vaporizing one kilogram of water requires as much as two thousand times the energy required to raise the temperature of one kilo dry air by one Kelvin degree. Different regimes are identified with the Bowen ratio, $B = F_{Hs}/F_{Es}$, e.g. being large over desert and low over ocean.

The last term of Eq. (2.2) is the ground flux, which is heat conducted to/from deeper soils, positive value is downward. Table 2.1 also includes a selection of heat capacities for soil related materials.

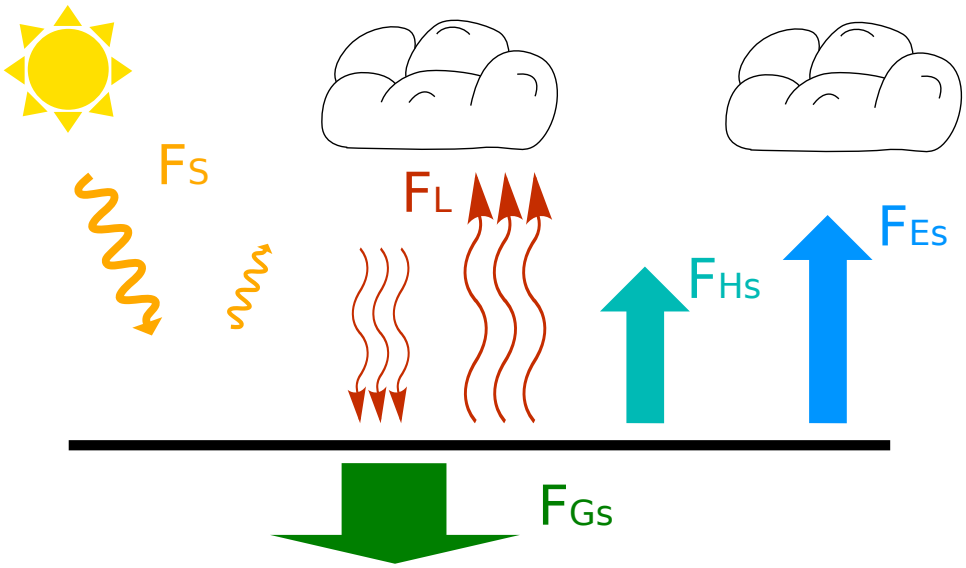


FIGURE 2.1: Schematics of land surface energy balance. F_S and F_L are shortwave and long wave radiation, F_{Hs} is sensible heat flux, F_{Es} is latent heat of evaporation, and F_{Gs} is ground heat flux.

C_{air}	1.004×10^3	$\mathrm{J\,K^{-1}\,kg^{-1}}$	specific heat capacity of dry air
C_{wat}	4.218×10^3	$\mathrm{J\,K^{-1}\,kg^{-1}}$	specific heat capacity of liquid water at $0\,^{\circ}\mathrm{C}$
L_v	2.50×10^6	$\mathrm{J\,kg^{-1}}$	specific latent heat of vaporization $0\,^{\circ}\mathrm{C}$
L_f	3.34×10^5	$\mathrm{J\,kg^{-1}}$	specific latent heat of fusion, water
C_{rock}	0.75×10^3	$\mathrm{J\,K^{-1}\,kg^{-1}}$	specific heat capacity of rock
C_{clay}	0.89×10^3	$\mathrm{J\,K^{-1}\,kg^{-1}}$	specific heat capacity of clay soil
C_{sand}	0.80×10^3	$\mathrm{J\,K^{-1}\,kg^{-1}}$	specific heat capacity of sandy soil
C_{peat}	1.92×10^3	$\mathrm{J\,K^{-1}\,kg^{-1}}$	specific heat capacity of peat soil

TABLE 2.1: Specific heat capacities. First four values from John M. Wallace (2006), ground related values from Arya (2001)

2.2 Surface Hydrology

As seen in Sect. 2.1, the presence of water dramatically influences the energy budget for the land surface. Soil water also influence the surface albedo (table 4.2 in Hartmann (1994)). A wet surface absorbs a larger portion of incoming solar radiation. Vegetation transpiration and growth are also constrained by the water available at the roots. Soil moisture is thus a key component in the energy budget. The most important elements of the hydrological cycle over land, during summer where there is no ice or snow present, are illustrated in Fig. 2.2.

Precipitation propagate the chaotic nature of the atmosphere into the soil, causing temporal variability. On short time scale, erroneous precipitation can thus introduce large error in soil water content. The other processes are much slower and easier to model, but their evolution is strongly dependent on static parameters such as soil texture and vegetation cover (Korres, Reichenau, and Schneider, 2013). Incorrect specification of these parameters will introduce a bias in soil moisture.

In the following, basic hydrological concepts are discussed in order to understand model performance and limitations. Notation and formalism are adopted from John M. Wallace (2006) and Dingman (2008)

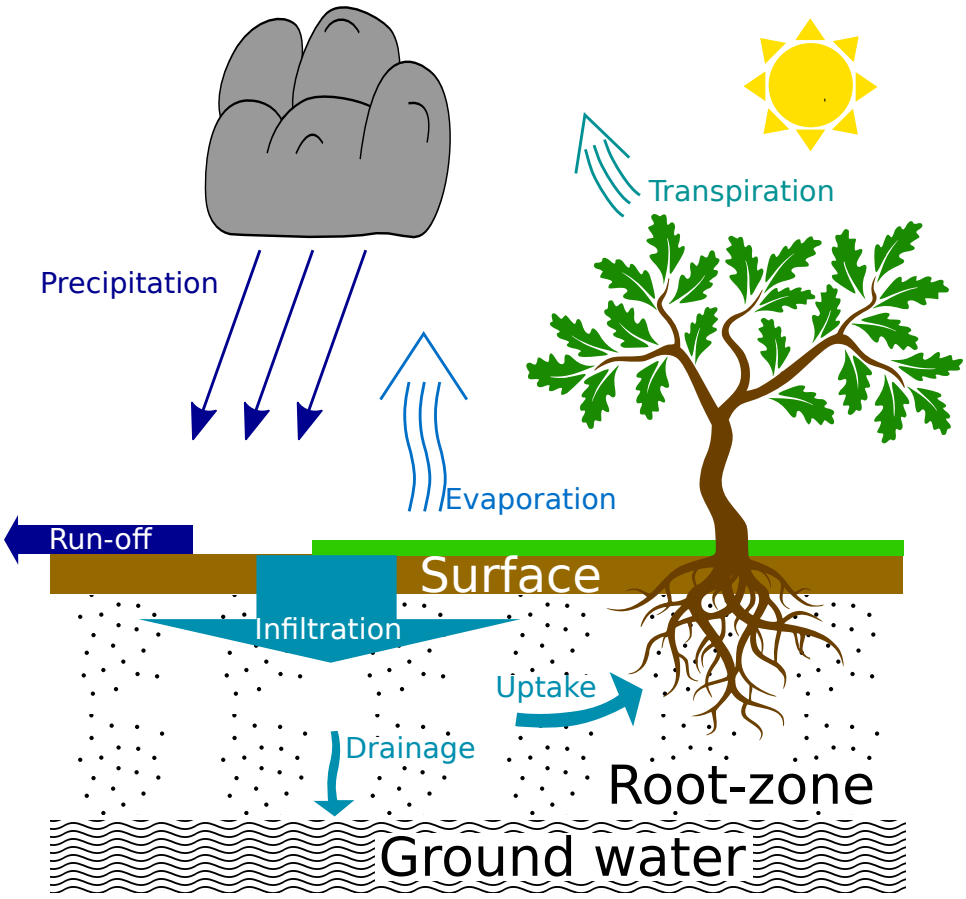


FIGURE 2.2: Elements of the hydrological cycle over land

2.2.1 Available Water

Soil porosity, the volume of air per total volume of soil is a key parameter when studying soil moisture together with field capacity and wilting point. Porosity is equal to the volumetric water content at saturation. When soil is saturated with liquid water after a heavy rainfall, the water content decays due to runoff and drainage. After some time the runoff and drainage rate decrease, and the water content stabilizes. At this point the water content is said to be at field capacity. If soil water is removed through evaporation from surface and transpiration from roots through vegetation, at some point the capillary force of the roots is no longer strong enough to absorb water from the soil, and the plants start to wilt. The wilting point is the soil water content at this point. The range between field capacity and wilting point is called available water. Water contents below wilting point are unavailable with respect to evapotranspiration, but can still decrease to hygroscopic water content that is the minimum water content in natural soil.

2.2.2 Water Transport in Soil

The main source of water in the soil is precipitation. However, to account for the soil moisture we also need to know other sources and sinks, and their respective magnitudes. One important component is the water flow, which is over all a sink for the surface water content. To obtain an equation

for the soil water content and its temporal evolution, Darcy's law is combined with conservation of mass within a volume. For simplicity, only the vertical component is considered.

Darcy's Law

Flow rate in unsaturated porous media can be described with Eq. (2.3)

$$q_z = -K_h(\theta)\left[1 + \frac{d\psi(\theta)}{dz}\right] \tag{2.3}$$

q_z is the vertical volumetric flow rate per cross-section $\mathrm{m\,s^{-1}}$, K_h is the hydraulic conductivity $\mathrm{m\,s^{-1}}$, θ the volumetric water content $\mathrm{m^3\,m^{-3}}$, $\psi = \frac{p}{\gamma_w}$ is the pressure head where p is the water pressure and γ_w the weight density of water, and z is the height above, which point the flow is evaluated. The hydraulic conductivity is a measure of how efficient water flows in the media, and as the notation suggests it is dependent on the water content. K_h depends, together with soil moisture, on the porosity or the grain size of the soil. Soil with large grains like sand, allows larger space between the particles, and thus more space for the water to move freely, compared to fine grained soils like clay. The capillary force of clay soil is larger, and thus holds more water than large grained soil.

Richards Equation

Using mass conservation for a volume $\delta x \delta y \delta z$ and assuming constant density of water ρ_w, we can describe the change of water content in time as the flow rate into the volume minus the flow rate out of it. We introduce z', increasing with depth, opposite of z.

$$\rho_w \delta x \delta y \delta t \delta z \frac{\partial \theta}{\partial t} = \underbrace{\rho_w \delta x \delta y \delta t q_{z'}}_{\text{flux in}} - \underbrace{\rho_w \delta x \delta y \delta t (q_{z'} + \frac{\partial q_{z'}}{\partial z'} \delta z)}_{\text{flux out}}$$

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q_{z'}}{\partial z'}$$

then differentiating Eq. (2.3) with respect to z',

$$\frac{\partial \theta}{\partial t} = -\frac{\partial K_h(\theta)}{\partial z'} + \frac{\partial}{\partial z'}\left(K_h(\theta)\frac{\partial \psi(\theta)}{\partial z'}\right) \tag{2.4}$$

We can define a hydraulic diffusivity $D_h(\theta) \equiv K_h(\theta)\frac{\partial \psi(\theta)}{\partial \theta}$, and rewrite Eq. (2.3) and (2.4),

$$q_z = -K_h(\theta) - D_h(\theta)\frac{\mathrm{d}\theta}{\mathrm{d}z} \tag{2.5}$$

$$\frac{\partial \theta}{\partial t} = -\frac{\partial K_h(\theta)}{\partial z'} + \frac{\partial}{\partial z'}\left(D_h(\theta)\frac{\partial \theta}{\partial z'}\right) \tag{2.6}$$

The hydraulic conductivity and diffusivity are dependent of soil moisture and soil texture. A prediction of soil water content thus requires a soil moisture profile, and information about the soil properties and how they vary

with depth. Richards equation has no analytic solution and has to be approximated. In modern surface models this is usually done using finite differences method (Dingman, 2008, Chapter 6).

2.2.3 Evapotranspiration

Evapotranspiration is a term for liquid water ending up as atmospheric vapor, including evaporation from the surface, transpiration from plants, Wet and dry soil moisture events are extracted from the simulation output to identify differences. and sublimation from ice and snow. About 62 % of the global precipitated water over land is evapotranspired back to the atmosphere (Dingman, 2008, p. 272). Evapotranspiration is not only an important sink for soil moisture, but also a key component in the energy balance discussed in Sect. 2.1.

Evaporation of soil water can be separated into two difference regimes: energy limited and soil moisture limited. In an energy limited regime, the soil contains a sufficient amount of water, so that the evaporation rate is driven by the amount of energy available. If the soil water content is below a critical limit, the soil moisture provides a first order constraint on the evaporation rate (Seneviratne et al., 2010). In soil moisture limited regimes a positive feedback mechanism is possible. If the soil moisture is decreased, it causes decreased evaporation, hence increased heating. Increased temperature reduces the relative humidity, and thus induce stronger evaporation despite the low soil moisture content. This mechanism can cause warm extremes over continental areas, soil moisture is thus important for predicting such events (Weisheimer et al., 2011).

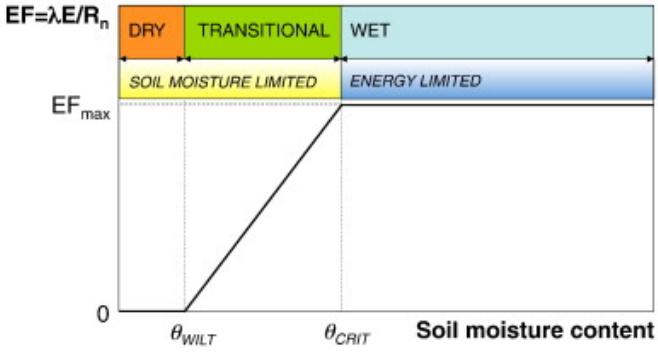


FIGURE 2.3: Schematics of energy limited and soil moisture limited evaporation regimes. Fig. 5 from Seneviratne et al. (2010)

Chapter 3

Observations

3.1 The Role of Observations in Numerical Weather Prediction

In the 17th century, thermometers were developed and the first meteorological measurements recorded. In 1860 Admiral Robert FitzRoy collected meteorological observations through the telegraph, made synoptic charts, and provided forecasts for England (BBC, 2015). Into the 20th century Vilhelm Bjerknes stated that the weather could be predicted by using a set of mathematical equations (Bjerknes, 1904). The equations were shown to be chaotic in nature by Lorenz (1963). When the equations are integrated in time, an initialization, a point in phase space, is required. The chaotic nature implies that an arbitrary small error in the initial condition, will blow up when integrated forward in time. Eventually, the imprecise initialization will cause the trajectory to have a total different solution than the true state. This behavior limits the predictability of the weather to a couple of hours for small scale systems such as convective clouds, and up to one to two weeks for synoptic scale systems. In operational NWP the model is initialized regularly when fresh observations are available. The amount of observations used in the initialization increase continually. To begin with, only observation stations on land were used, then reports from ships, and with time, radiosondes, buoys and a variety of satellites have been launched and utilized in the process. Earth observations are also of great benefit for researchers when studying processes, since they output the real system processes rather than limited model estimations.

3.2 In Situ Observations

We distinguish between direct and indirect measurements. A direct measurement or in situ measurement is a measurement, in direct contact, where the observed quantity is the quantity under study, e.g. temperature measured with a thermometer. In Norway hundreds of stations, manual and automatic, provide temperature, pressure, wind speed and precipitation amounts around the clock, that are used to initialize the forecasting model. Regarding soil moisture the situation is different. In situ measurements of soil moisture does exist in Norway, but quite few and they are not easily available. The International Soil Moisture Network (ISMN) (Dorigo et al., 2011) has managed to gather soil moisture observations from many countries and networks, unfortunately no stations in Norway are included at this point.

The true direct way to measure soil moisture is to weigh a soil sample before and after it has dried, then calculating the mass of water. However, this method is cumbersome, and soil moisture is usually deduced from measurements of other quantities. One method is to measure the electrical resistance between two electrodes placed in the soil. Because water has a different electrical conductivity than dry soil, the soil moisture content can be calculated. Other kinds of measurements use the close link between soil moisture and the dielectric constant, which is also utilized in remote methods.

3.3 Satellite Remote Sensing

In situ measurements are usually very accurate, but the major weakness is the spatial coverage. For how large area could a single point soil moisture measurement represent? For a large flat homogeneous area e.g. a desert, this area could be large. However, in areas with rough topography, mixed forest, and patches of wetland, the spatial variability can be large in few meters.

For a grid cell in a NWP model ranging from $2.5\,\mathrm{km}$ to $20\,\mathrm{km}$, one single soil moisture measurement would be of very limited value, but there is an alternative. Satellites observe the earth from space, giving them a wider view. Satellites have been crucial for improving weather forecast (Bauer, Thorpe, and Brunet, 2015), and a large number of sensors are included in assimilation systems at the operational centers.

Satellites can orbit the earth in different ways. Geostationary satellites are always located over the same spot at the equator. These satellites have an altitude of $35\,786\,\mathrm{km}$ above earth and can cover a large portion of the hemisphere below. Satellites can also be in low orbits including polar orbit with altitudes around $700\,\mathrm{km}$ and an orbit period of about $100\,\mathrm{min}$, where they cover the high latitudes as well.

Satellites cannot observe the studied variable directly, but measure electromagnetic radiation by counting photons, called data level 0, which is then translated to randiances (level 1). Different sensors that operate in different bands are used to observe different quantities. A band is a frequency interval on the electromagnetic spectrum. Different frequencies require different sensors and techniques. We distinguish between passive and active sensors. A passive sensor, or radiometer, observes radiation emitted from the earth or from molecules in the atmosphere. The atmosphere is more transparent to radiation in certain bands, these are utilized when monitoring the surface. The surface is not a perfect black body, but it absorbs and emits according to its emissivity. The radiation temperature observed from a sensor is thus lower than the real kinetic temperature and is known as the brightness temperature. Alternatively the satellite emit radiation itself and then observe the radiation reflected back, back scatter, known as a radar. Lasers are also used, e.g. with the ability to measure terrain in very high detail. When the level 1 data is observed, a retrieval algorithm is used to obtain the variable of interest, termed level 2 data. Level 3 is level 2 data interpolated onto a grid or field. Level 4 data is a model derived state by assimilation of level 1 or 2 data (Lahoz and De Lannoy, 2014).

3.4 Soil Moisture Retrievals with L-band Radiometry

Table 1 in Lahoz and De Lannoy (2014) gives an overview of satellites that observe quantities related to the hydrological cycle, including surface soil moisture. The two most recent soil moisture dedicated satellites, SMOS and SMAP both operate as passive sensors in the L-band (long wave), which corresponds to frequencies in the range $1\text{-}2\,\mathrm{GHz}$. The atmosphere is almost transparent for these frequencies and surface brightness temperature can be observed from space almost unaffected by the atmospheric conditions including clouds (Schmugge, 1998) and (Jones and Vaughan, 2010, Chapter 2). At long wavelengths where the Raylgeigh-jeans approximation holds, the brightness temperature, T_B, is related to the real kinetic temperature T_k as $T_B = \epsilon T_k$, where ϵ is the emissivity.

The emissivity is closely related to the dielectric constant of the material, shown in Fig. 3.1. The dielectric constant of water is $\sim\ 80$ compared to $\sim\ 3.5$ for dry soils. The emissivity is thus highly dependent on the water content of the material/soil, which makes it suitable for remote sensing of soil moisture. The step from observed brightness temperature to soil moisture requires additional information about temperature, surface roughness and vegetation types etc., these may come from databases, models or parallel observations (Jones and Vaughan (2010) their Sect.(5.5.1)).

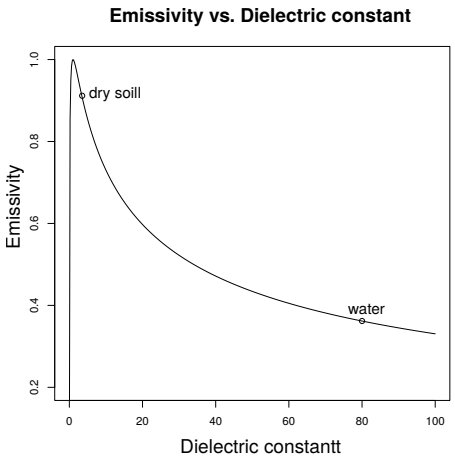


FIGURE 3.1: Emissivity as function of the dielectric constant based on Eq.(1.7) in Behari (2006)

The penetration depth into the soil is dependent of wavelength, where longer waves penetrate, and thus are emitted from, deeper than shorter waves. In the L-band the brightness temperature retrieved is valid for about $5\,\mathrm{cm}$ below the surface. These wavelengths are also transmitted through vegetation less affected than shorter waves. In sum these properties have made L-band radiometry the choice for the two most recent satellites dedicated to soil moisture (Kerr, 2007).

Chapter 4

Data Assimilation

Physical processes described in Ch. 2, make a basis for modeling of the land surface. Observations and satellite products are explained in Ch. 3. Observations are valuable because they provide the best information about the true state of a system, but have the limitations in their spatial and temporal coverage. Models on the other hand, have good spatial and temporal coverage only limited by computer power, but incomplete knowledge of the processes imply limited parameterizations, which are also resolution dependent (Lahoz and De Lannoy, 2014). Models and observations are crucial for our understanding of the system under study. We can learn about our model by comparing it with observations, or we can verify measurements by using models. How we use observations and models together can vary. Most of the time, comparison is done in plots, figures, and statistical measures. If we want to improve a model by using observations or interpolate observations using a model, we need objective tools to combine them properly. Data assimilation (DA) is the field of combining information, optimally, based on their respective uncertainties.

4.1 Model and Observation Uncertainty

Models and observations have errors, and it is important to know them in a DA system. In principle we are combining two values with associated uncertainties, regardless of where they come from, model or observation. However, the nature of the errors is not identical, and the approach we use to estimate them differ.

We usually distinguish between model error and background error, where the model error comes from to finite resolution, parameterizations, and lack of knowledge about the processes. The background error is related to the error of the previous analysis and the predictability of the system.

The error of a measurement, observation error, can be separated into components, instrument error and error of representativeness. The instrument error is, as the name suggests, the uncertainty of the measuring device, and should be given by the manufacturer. The error of representativeness arises when we use a measurement, of one scale, to represent another scale, e.g. when we compare 2 meter temperatures from a synop station with a model grid cell of several kilometers, or using a 30 by 30 km resolution soil moisture product together with a 2.5 by 2.5 km model. In remote sensing, or indirect measurements, a retrieval algorithm is used to obtain the variable of interest, which is often based on approximations and models. These errors are often estimated and provided along the retrieval product as described later in Ch. 10.

It is common to distinguish between random and systematic errors. A signal has some degree of random noise. Averaging a signal will remove

this noise, meaning it has zero mean. The precision of an observation is related to this kind of noise. Systematic error, or bias, in a signal cannot be removed by averaging, but if the systematic error is known it can be subtracted from the signal.

4.2 Optimal Combination of Data

As stated above, we need an objective tool to combine two observations, the model can be treated as an observation, that takes care of the respective uncertainties of each observation. A simple 1D case is presented in Lahoz, Khattatov, and Ménard (2010). Assume we have two observations, x_1 and x_2 with the corresponding standard deviations σ_1 and σ_2. The optimal combination of these two is given by Eq. (4.1).

$$x = \frac{\frac{x_1}{\sigma_1^2} + \frac{x_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} \tag{4.1}$$

where the denominator is the inverse of the variance of the new value x. If we let σ_2 become very large, x_2 contributes very little, and the variance of the analysis approaches the variance of x_1, $\sigma_2^2 \leq \min(\sigma_1^2, \sigma_2^2)$. The analysis is thus more certain than the most certain observation. This is true if the observation errors are bias free, uncorrelated, and normally distributed.

We can generalize this into a multidimensional case, assuming a perfect model, and obtain the Best Linear Unbiased Estimate (BLUE) (Sect. 2.4 in Lahoz, Khattatov, and Ménard (2010)). Let $\mathbf{x}_k \in \mathbb{R}^n$ be the true state in model space at time t_k, and $\mathcal{M}_{k,k+1} : \mathbb{R}^n \to \mathbb{R}^n$ be the non-linear operator propagating the model state from t_k to t_{k+1} so that,

$$\mathbf{x}_{k+1} = \mathcal{M}_{k,k+1}(\mathbf{x}_k), \quad k = 0, \ldots, N-1. \tag{4.2}$$

$\mathbf{x}_0^b$ is a prior estimate of the state $\mathbf{x}_k$, also known as the background state, with errors $\epsilon_0^b = \mathbf{x}_0 - \mathbf{x}_0^b$. The observation $\mathbf{y}_k \in \mathbb{R}^{p_k}$ is the observation vector, with the following relation to model state,

$$\mathbf{y}_k = \mathcal{H}_k(\mathbf{x}_k) + \epsilon_k^o, \tag{4.3}$$

where $\mathcal{H}_k$ is a non-linear operator transforming model space to observation space, including grid interpolations and conversions. ϵ_k^o is the observation error in observation space.

A linear estimate of the analysis $\mathbf{x}^a$ can be found using $\mathbf{M}_{0,k}$ and $\mathbf{H}_k$ as the linearized versions of $\mathcal{M}_{0,k}$ and $\mathcal{H}_k$. Assuming that the errors ϵ_0^b and ϵ_k^o are randomly distributed, have zero means, and are uncorrelated, with corresponding covariance matrices $\mathbf{B}_0 \in \mathbb{R}^{n \times n}$ and $\mathbf{R}_k = \mathbb{R}^{p_k \times p_k}$ respectively. The following equations can then be derived using either *minimum variance estimate* or *maximum a posteriori Bayesian estimate*.

$$\mathbf{x}_0^a = \mathbf{x}_0^b + \mathbf{K}\mathbf{d} \tag{4.4}$$

$$\mathbf{K} = \mathbf{B}_0\mathbf{H}^T(\mathbf{H}\mathbf{B}_0\mathbf{H}^T + \mathbf{R})^{-1} \tag{4.5}$$

$\mathbf{K} \in \mathbb{R}^{n \times p}$ is called the gain matrix, and $\mathbf{d}$ is the innovation vector

$$\mathbf{d} = \mathbf{y}_k - \mathcal{H}_k(\mathbf{x}_0^b) \tag{4.6}$$

4.3 Sequential Data Assimilation

Operational weather forecasts are produced with intervals in time, e.g. the Scandinavian AROME-MetCoOp model is initiated four times daily at 00, 06, 12, and 18 (Müller et al., 2017). Observations are available at different times, depending on the type, e.g. automatic station measurements every hour/minute or satellites once/twice a day. It is thus a question, which observations that should be used for the next analysis. In sequential DA, observations within a certain time window around the analysis time are used, as observations for that time. The model is then propagated to the next analysis time, where a new estimate is made when observations become available. In contrast to sequential DA schemes, the 4D variational approach uses observations when they are available during the width of the assimilation window. The two types are illustrated in Fig. 4.1.

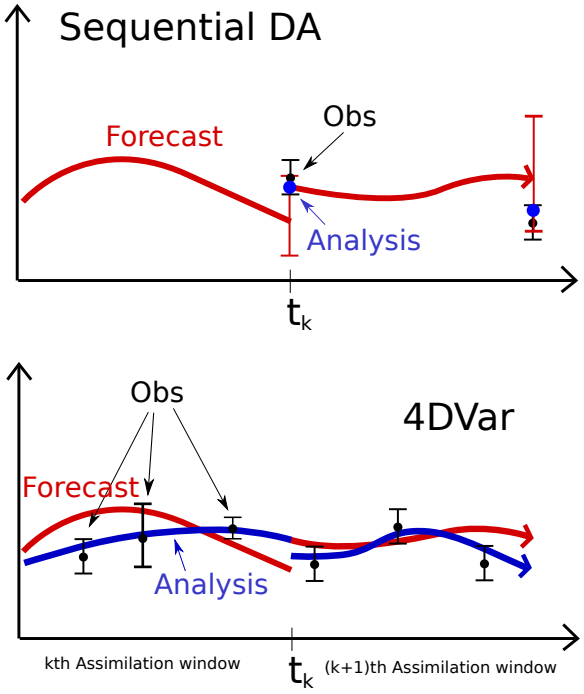


FIGURE 4.1: Model trajectories for sequential and variational DA. Red lines are model trajectories, blue lines or point indicate the resulting analysis after DA, black dots are observations.

4.4 Kalman Filter

The KF is a sequential DA scheme formulated by Kalman (1960).

At a time t_k a prior estimate, background, of the current state is known, $\mathbf{x_k^b}$, and observations, $\mathbf{y_k}$, are available. The analysis is on the form,

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k^b) \tag{4.7}$$

with the following analysis error covariance matrix,

$$\mathbf{A}_k = (\mathbf{I} - \mathbf{K_k}\mathbf{H_k})\mathbf{B_k}. \tag{4.8}$$

$\mathbf{K_k}$ is known as the Kalman gain matrix and gives the weight to the observations.

$$\mathbf{K_k} = \mathbf{B_k H_k^T}(\mathbf{H_k B_k H_k^T} + \mathbf{R_k})^{-\mathbf{1}} \tag{4.9}$$

$\mathbf{H_k}$ is the observation operator, transforming model space to observation space.

The analysis state is then propagated forward to the next analysis time,

$$\mathbf{x}_{k+1}^b = \mathbf{M}\mathbf{x}_k^a, \tag{4.10}$$

where $\mathbf{M}$ is a linear model. The new background error covariance matrix $\mathbf{B_{k+1}}$, used in the next analysis, is

$$\mathbf{B_{k+1}} = \mathbf{M A_k M^T} + \mathbf{Q} \tag{4.11}$$

given a model error covariance matrix $\mathbf{Q}$.

4.5 Extended Kalman Filter

For non-linear systems, the propagation of a state is given by a non-linear operator $\mathcal{M}$,

$$\mathbf{x_{k+1}} = \mathcal{M}(\mathbf{x_k}). \tag{4.12}$$

The observation operator is also non-linear $\mathcal{H}$. The innovation is thus on the form $\mathbf{y_k} - \mathcal{H}(\mathbf{x_k})$. However, we can approximate the model and observation operators with linearizations within the assimilation window t_k to t_{k+1}.

$$\mathbf{H} = \frac{\partial \mathcal{H}}{\partial \mathbf{x}} \tag{4.13}$$

$$\mathbf{M} = \frac{\partial \mathcal{M}}{\partial \mathbf{x}} \tag{4.14}$$

Calculating the linear model and observation operator, and to propagate the error covariance matrix is computationally expensive. For atmospheric application this method is rarely used due to its large number of calculations. Other methods are developed to suite a practical implementation. However, for land DA, where the length of the state vector is significantly smaller, the technique is well used. A popular version is the SEKF where the background error covariance matrix, $\mathbf{B}$, is constant in time. Draper, Mahfouf, and Walker (2009) found that the SEKF and EKF produced similar soil moisture increments, but recommended to use the EKF for future work. However the SEKF has been preferred in later studies including Mahfouf et al. (2009), Draper, Mahfouf, and Walker (2011), De Rosnay et al. (2013), Barbu et al. (2014), Fairbairn et al. (2015), and Albergel et al. (2017). In this study the SEKF is used, because of its simplicity and that there are some challenges in the implementation described in Sect. 7.1.2

4.6 Ensemble Kalman Filter

Another version of the KF is the Ensemble Kalman Filter (EnKF) described in Evensen (2003). Rather than evolving the background error covariance matrix as in Eq. (4.11), an ensemble forecast is made for the background state, then the background error is approximated with the ensemble spread. $\overline{\mathbf{x^b}}$ denotes the ensemble mean.

$$\mathbf{B} \simeq \mathbf{B_e} = \overline{(\mathbf{x^b} - \overline{\mathbf{x^b}})(\mathbf{x^b} - \overline{\mathbf{x^b}})^\mathbf{T}} \tag{4.15}$$

The computation of the background error covariance matrix is free, but the cost of running a large ensemble has to be taken into account.

One of the main challenges of the EnKF is to perturb the ensemble. The goal is to perturb each ensemble member at the beginning of the assimilation window to produce a realistic spread, hence uncertainty, in the background field. The simplest way is to give the members random perturbations with a certain distribution. For soil moisture, which is limited to the dynamic range of the available water, certain thresholds have to constrain the perturbations. Another problem is that the ensemble collapses during the assimilation window. This can happen when the forcing is dominating the state relative to the initial perturbations, e.g. during heavy rainfall. Because the atmosphere is the most uncertain and in some cases dominating driver of the system, perturbation could be added to the forcing, instead or together with perturbations in the soil parameters (Fairbairn et al., 2015).

Part II

Methods and Implementation

Chapter 5

The Surface Model SURFEX

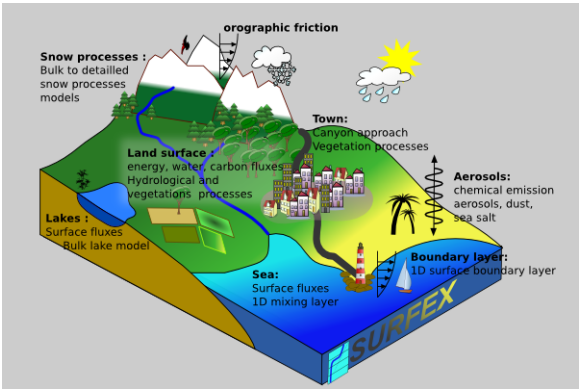


FIGURE 5.1: Schematics of the SURFEX model. `http://www.umr-cnrm.fr/surfex/`

SURFEX is a full surface-atmosphere interaction model, developed by Centre National de Recherches Météorologiques (CNRM). SURFEX is designed to run coupled with an atmospheric model, inline, but can also run decoupled, an offline run. There is no interaction between grid cells, hence single point simulations are possible. For these reasons SURFEX is computationally affordable to run. The model offers a variety of options for carbon cycle, aerosol, snow, and vegetation to mention a few (Le Moigne, 2012).

The SURFEX model serves as the surface scheme in the operational NWP model AROME in the HIgh Resolution Limited Area Model (HIRLAM) Aire Limitée Adaptation Dynamique Développement International (ALADIN) Research on Mesoscale Operational NWP in Europe (HARMONIE) forecasting system (Bengtsson et al., 2017), making it an obvious choice in this study. SURFEX also has the option to use advanced soil physics with the ISBA-DIF scheme. ISBA-DIF represents the ground with multiple, 14 by default, layers. Energy transport between layers is described with the diffusion equation, and water transport by Richards equation. This is a more physical approach than the older force restore method.

5.1 Model Structure and Equations

A SURFEX grid cell is split into four tiles: nature, town, lake, and ocean. Separate schemes handle these tiles individually, and a weighted mean of the energy fluxes etc., is calculated at the end of the simulation time step. We will be focusing on the nature tile, solved with the ISBA model. A grid

cell, 2.5×2.5 km, is most likely to contain multiple vegetation types, land usage, and soil properties. Rather than averaging the properties of the different cover types and then calculating the energy balance for the nature tile, it is possible to calculate the energy transports for each cover type (patch) individually and then averaging the fluxes at the end. Because each patch has its own energy budget, the number of offline simulations is multiplied with the number of patches, plus the assimilation has to be done for each patch. Using multiple patches will thus dramatically increase the computational load. Another problem is how to apportion one soil moisture observation from the satellite to multiple patches. Only one patch is thus used in this work.

A detailed description of the model is found in Le Moigne (2012). An option to the diffusive approach is the force restore (FR), a simpler scheme than ISBA-DIF. The surface and soil are represented with three layers, surface, root zone, and deep soil. It is thus cheaper to run than the diffusion version with 14 layers. Currently the FR is implemented in the operational forecasting system in Scandinavia. The ISBA-DIF is used in this work, because of the better agreement between soil moisture variables and satellite product, model soil water content $1\text{-}4\,\mathrm{cm}$ and $0\text{-}5\,\mathrm{cm}$ respectively.

The prognostic variables of interest in this study, soil temperature and soil moisture are described with equations (5.1) and (5.2).

$$c_h \frac{\partial T_g}{\partial t} = \lambda \frac{\partial^2 T_g}{\partial z^2} + \Phi \tag{5.1}$$

$$\frac{\partial w_l}{\partial t} = -\frac{\partial F}{\partial z} - \frac{\Phi}{L_f \rho_w} - \frac{S_l}{\rho_w} \tag{5.2}$$

here c_h is the total heat capacity, T_g is the soil temperature, λ is the thermal conductivity, Φ is a latent heat source/sink resulting from phase transformation of soil water.

w_l is the volumetric liquid soil water content.F is the vertical flow rate, L_f is the latent heat of fusion, ρ_w is the density of liquid water, S_l is a source/sink of evapotranspiration and lateral inflow.

The soil temperature, Eq.(5.1), is solved with a backward differences implicit time scheme, and the water content, Eq.(5.2), is solved with the Crank-Nicolson scheme (Le Moigne, 2012).

Chapter 6

Observations

6.1 SMOS and SMAP

Soil moisture has crucial role in the earth system and has gained interest in many fields of earth science over the years (Seneviratne et al., 2013; Jaeger and Seneviratne, 2011; Koster et al., 2011; Weisheimer et al., 2011; Seneviratne et al., 2010; Koster et al., 2004; Koster, 2004; Beljaars et al., 1996). Recently two satellite missions dedicated to soil moisture have been launched. The SMOS (2009), and SMAP (2015), both operating in the L-band (1-2 GHz), aim to provide global soil moisture retrieval for the top $\sim 5\,\mathrm{cm}$ of the soil, with a resolution of about $30\,\mathrm{km}$, an uncertainty below $0.04\,\mathrm{m}^3\,\mathrm{m}^{-3}$, and a revisit time of 3 days.

SMOS and SMAP are in sun-synchronous polar orbits. They both pass the equator 6am and 6pm local time, but in opposite modes. SMOS is ascending at the 6am pass, where SMAP is descending. In the retrieval of soil moisture, it is preferred to assume that soil surface and vegetation have the same temperature, and that the soil temperature profile is uniform. These conditions usually occur in the morning around 6 am, hence sun-synchronous orbit is suitable. A phenomenon called Faraday rotation, which is rotation of the polarization plane of electromagnetic waves due to a magnetic field, can disturb the signal and impair the soil moisture retrieval. The effect of Faraday rotation is minimized in the morning (ONeill et al., 2015). The sun-synchronous orbit also has the benefit that the satellite always receives solar radiation, which can supply the unit with power.

The level 3 products used in this study, SMOS form Centre Aval de Traitement des Données SMOS (CATDS) (Jacquette et al., 2010) and SMAP from National Snow and Ice Data Center (NSIDC) (ONeill, 2016), are retrieved with a physical approach. A forward radiative transfer (RT) model, like the Community Microwave Emission Modeling Platform (CMEM) (De Rosnay et al., 2009) or the Land Parameter Retrieval Model (LPRM) (Schalie et al., 2016). translates the state, atmosphere, vegetation and soil moisture etc. into radiation equivalent to what the sensor receives at the top of the atmosphere. By an iterative process, these parameters are tuned to minimize a cost function, the distance to the observed signal, and thus find the best value for the soil moisture. The forward model requires static and dynamic input data. Static fields, constant in time, include soil texture, vegetation cover and topography. These data sets come from ancillary sources including Food and Agricultural Organization (United nations) (FAO) soil sand and clay database (Batjes, 1997) and ECOCLIMAP (Le Moigne, 2012) for vegetation types and land use data. Meteorological, dynamic, data are obtained from the global NWP models ECMWF and Global Modeling and Assimilation Office (GMAO) for SMOS and SMAP respectively.

The fully detailed retrieval algorithm of SMOS is described in Kerr et al. (2012).

In addition to the passive sensor, SMAP is equipped with an active unit, a radar. The radar measurements should have provided soil moisture retrievals at much finer resolution than the radiometer, 3 by $3\,\mathrm{km}$ and 30 by $30\,\mathrm{km}$ respectively. A composite product with a resolution of 9 by $9\,\mathrm{km}$ was also planned. Unfortunately the radar stopped working already a few months after launch, (Cole and Buis, 2015), but the passive instrument still provides retrievals to days date. The retrieval algorithm for SMAP is described in the algorithm theoretical basis document (ATBD) (ONeill et al., 2015).

6.2 Scaling Observations

L-band soil moisture products represent the first few cm of the soil, $5\,\mathrm{cm}$ is often used as a measure. However, the penetration depth is dependent on the texture of the soil, the vegetation cover, and the soil moisture itself. Using the satellite product directly to represent a certain variable in a model is thus not appropriate. The level 2 and 3 soil moisture products are, as described above, calculated with the use of some surface model, different from the model being used in this study. Bias is expected and the data need preprocessing.

In this study, a linear re-scaling method, Eq. (6.1), is applied to the satellite data prior to the simulation. The goal is to have a set of observations w'_{obs} with the same climatology as the model. SURFEX is run, with no assimilation, for the extent of the experiment period, and model mean $\overline{w}_{mod}$ and standard deviation σ_{mod} are used to normalize the original satellite data w_{obs}.

$$w'_{obs} = (w_{obs} - \overline{w}_{obs})\frac{\sigma_{mod}}{\sigma_{obs}} + \overline{w}_{mod} \tag{6.1}$$

Where $\overline{w}_{obs}$ and σ_{obs} are the satellite mean and standard deviation respectively. This corresponds the first modes of the cumulative density function (CDF). If a sufficient amount of data is available, multiple years (Reichle, 2004), a full CDF matching can be an alternative to the linear version. The time span of the experiments in this thesis is to short and the linear approach is most suitable.

The seasonal variation is assumed to be negligible during the summer months (JJA), and a constant mean is used for the entire period. Alternatively one can use a moving average to account for the seasonal variations. The mean and standard deviation are calculated individually for each model grid cell. Only time points where satellite observations are available, for the specific location, are used. A minimum of 15 data points during the three month period is required for each grid cell, cells with fewer are removed.

Since the resolution of the satellites is much coarser than in the model, the nearest neighboring satellite observation is used. No smoothing is performed, and neighboring model grid cells are assigned the same satellite observation. After the scaling, variations of soil properties will lead to different climatology and thus different soil moisture values in neighboring cells.

Chapter 7

SURFEX Offline Data Assimilation

SURFEX has the option for DA, also for offline simulations. This system is called SODA. The methods currently implemented are Optimal interpolation (OI), EKF, and EnKF in different versions. In SURFEX v8.0 used in this work, the following variables can be used as observations: 2 meter temperature (T2M), 2 meter relative humidity (HU2M), superficial soil moisture (WG1) and leaf area index (LAI). The possible control variables are: superficial soil temperature (TG1), Root zone temperature (TG2), WG1 (as above) and root zone soil moisture (WG2).

The observable variables 2 meter temperature and relative humidity are diagnostic variables in SURFEX.

7.1 The Extended Kalman Filter

The implementation of the EKF in SODA is described in Le Moigne (2012), and covered briefly below.

7.1.1 The Linearized Observation Operator

Assuming linear behavior during the assimilation window, a finite difference approximation provides the linearized version, $\mathbf{H}$, of the observation operator $\mathcal{H}$.

$$H_{ij} = \frac{\partial \mathbf{y_i}}{\partial \mathbf{x_j}} \tag{7.1}$$

ISBA is propagated from time t_0 to t_1, $n+1$ times where n is the length of the control vector. Each simulation is given a small perturbation to one of the control variables, except the control. $\partial \mathbf{x_j}$ corresponds to the initial jth perturbation and $\partial \mathbf{y_i}$ corresponds to the ith control variable after a model run. The sensitivity of the observed variable to a small perturbation of another variable is calculated, and is used when giving weight to each of the control variables.

7.1.2 The Background Error Covariance Matrix

Except for the initial run, where $\mathbf{B}$ is prescribed, the background error covariance matrix $\mathbf{B}$ is propagated from the previous analysis according to Eq. (4.11). The state is propagated $\mathbf{x_b^{t_1}} = \mathcal{M}(\mathbf{x_a^{t_0}})$ where $\mathcal{M}$ is the ISBA

model, and $\mathbf{B}$ is updated by

$$\mathbf{B}^{\mathbf{t_1}} = \mathbf{M}\mathbf{A}^{\mathbf{t_0}}\mathbf{M}^{\mathbf{T}} + \mathbf{Q}, \tag{7.2}$$

where $\mathbf{M}$ is the Jacobian to $\mathcal{M}$, $\mathbf{A}^{\mathbf{t_0}} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}^{\mathbf{t_0}}$ is the analysis error covariance matrix, and $\mathbf{Q}$ is a model error covariance matrix.

Draper, Mahfouf, and Walker (2009) have found that using a SEKF, where the $\mathbf{B}$ matrix is constant in time and purely diagonal will produce similar useful increments as with an evolving background error covariance matrix. Due to its simpler implementation and lighter computational load, it has become popular in land DA, (Mahfouf et al., 2009; Barbu et al., 2011; De Rosnay et al., 2013; Barbu et al., 2014; Fairbairn et al., 2015; Albergel et al., 2017) and will also be the practice in this study. There are also some issues in the implementation of the evolving $\mathbf{B}$ matrix, where singular matrices cause the program to crash. To avoid singular matrices, a constraint is currently set on the Jacobian to not exceed the range -0.1-1. As this constraint was implemented to avoid program failure, the latest version of SURFEX (v8.1) the limitation of Jacobian values are removed, and new constraint is set on analysis increments (Albergel and Munier, 2017).

7.1.3 Modifications

The valid observation and control variables listed above are outdated with regard to the ISBA-DIF model. The ISBA force restore scheme only has two soil layers with water content, whereas the diffusion version has up to 14. In the diffusion version, the first and second soil layer cover the first $4\,\mathrm{cm}$ of the soil, where in the force restore version they cover the entire root zone. The control vector thus has to be increased in order to include root zone soil layers in the ISBA-DIF version.

SURFEX v.8.0 has in this study been modified to use soil water content in all seven first upper soil layers (WG1-7) as possible control variables, reaching to a depth of $80\,\mathrm{cm}$ below the surface. To have the best correspondence with L-band satellite products, observation of water content in the second soil layer, 1-$4\,\mathrm{cm}$ (WG2), is also included as an option.

In the latest version of SURFEX, v8.1 released during this work, similar changes are included to make DA with ISBA-DIF possible.

7.2 Background Errors

The result of a DA is directly dependent on the prescribed observation and background/forecast errors. These values usually have to be tuned to the specific system and set up (Desroziers and Ivanov, 2001), and they cannot be tuned separately. In studies, with SEKF, Balsamo et al. (2007), Mahfouf et al. (2009) and Draper, Mahfouf, and Walker (2011), the background error is set to 0.1-0.2 times the dynamic range ($w_{sat} - w_{wilt}$). However, all the above studies used the force restore version of ISBA with only 2 soil layers, whereas in this study the ISBA-DIF scheme, with 7 control variables, is used. Albergel et al. (2017) used the ISBA-DIF model in a similar DA study and prescribed a background standard deviation of 0.04 for the second layer

soil moisture WG2 and 0.02 for deeper layers. Despite that the experimental setup is different, these values for background error are adopted in this study.

7.3 Error Diagnostics and Observation Errors

A method to diagnose the consistency of the DA system is described in in Desroziers et al. (2005), making use of innovations, observation minus forecast (O-F), $\mathbf{d}_b^o$, and residuals observation minus analysis (O-A), $\mathbf{d}_a^o$. By manipulating Eq. (4.6) the following relations are obtained. The innovation,

$$\mathbf{d}_b^o \simeq \epsilon_o - \mathbf{H}\epsilon_b \tag{7.3}$$

and the associated covariance below. Note that the errors are assumed to be uncorrelated,

$$E[\mathbf{d}_b^o(\mathbf{d}_b^o)^T] = \mathbf{R} + \mathbf{HBH}^T \tag{7.4}$$

where E denotes the expected value. The residual (O-A) can be written,

$$\mathbf{d}_a^o = \mathbf{R}(\mathbf{HBH}^T + \mathbf{R})^{-1}\mathbf{d}_b^o, \tag{7.5}$$

and its covariance to the innovation,

$$E[\mathbf{d}_a^o(\mathbf{d}_b^o)^T] = \mathbf{R}(\mathbf{HBH}^T + \mathbf{R})^{-1}E[\mathbf{d}_b^o(\mathbf{d}_b^o)^T] = \mathbf{R} \tag{7.6}$$

Innovations $\mathbf{d_b^o}$ and residuals $\mathbf{d_a^o}$ are available after the simulation and we can diagnose the consistency of our system.

$$(\tilde{\sigma^o})^2 = \sum_i^N \frac{(y_i^o - y_i^a)(y_i^o - y_i^b)}{N} \tag{7.7}$$

This method is also applied to tune the observation error. After a simulation is performed, Eq. (7.7) is evaluated, then $(\tilde{\sigma^o})^2$ is divided with the prescribed observation error. The result is called the chi-squared parameter, χ^2, and has value of unity for a consistent DA system. If $\chi^2 > 1$ after the first simulation, we choose a larger observation error, and vice versa. When a second simulation is performed, a new χ^2 is obtained, and a method of choice for finding roots can be used to approximate a new observation error. The system is non-linear and multiple iterations might be necessary to obtain the acquired consistency, $\chi^2 \approx 1$.

Chapter 8

Experimental Setup

8.1 Location

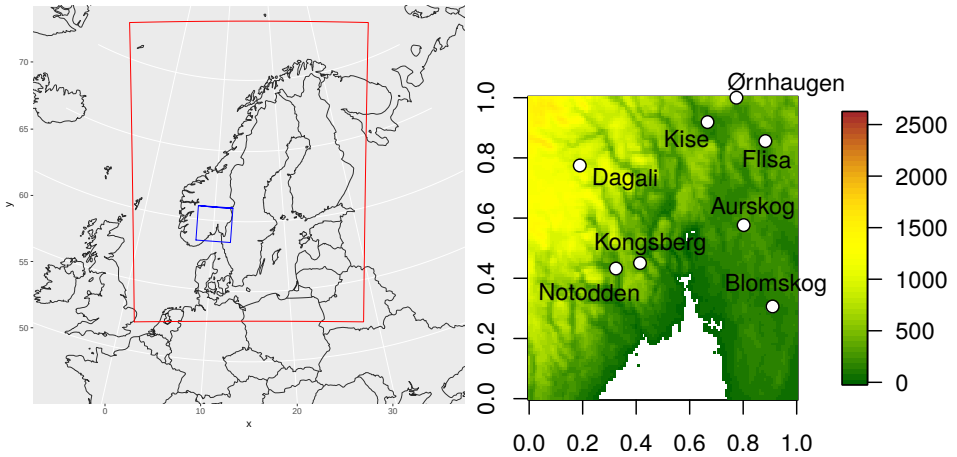


FIGURE 8.1: Left panel shows the domain of AROME-MetCoOp (red) and local SURFEX experiment (blue). Right panel shows the experiment domain with topography. The experiment domain size is 111x111 grid cells, each cell is 2.5km by 2.5km. The axes show the fractional grid number where 0 corresponds to the first grid cell, starting in the south west corner (7.73E 58.49N), and 1 corresponds to the last (111th) grid cell in each direction, that is $111 \times 2.5\,\mathrm{km} = 277.5\,\mathrm{km}$ in each direction toward the north east corner (12.65E 60.96N)

Norway has a high coast to land area ratio and the weather is heavily influenced by the ocean and systems approaching from the west. However, high mountain areas stretching from south to north separate the country with respect to the weather. The eastern side of the mountains has relatively flat topography and large agricultural areas. These are favorable conditions for L-band radiometry for soil moisture retrievals covered in Sect. 3.4. The domain selected for this study is thus on the east side of the mountain ridge, covering south-east Norway shown in Fig. 8.1. The domain is on a Lambert conformal projection with the reference longitude and latitude at $10.1\,\mathrm{deg}$ and $59.75\,\mathrm{deg}$ respectively, to best match the forcing from AROME-MetCoOp. The domain has a south-west corner at (7.73E 58.49N) and a north-east corner at (12.65E 60.96N), each grid cell has size 2.5km by 2.5km as the forcing data.

8.2 Time Period

The quality of soil moisture retrievals from L-band radiometry is reduced when the soil is frozen or snow is present on the ground (Kerr et al., 2012). These weaknesses make Norway a difficult area to fully utilize soil moisture satellites. To minimize frozen soil and snow cover, the three summer months JJA are selected for this study. Further, since data from both SMOS and SMAP were available from 2016 and this study started that year, simulations were thus done with data from 2016.

Figure 8.2 shows difference between monthly averages from 2016 and the monthly normal (upper panels) and the anomalies plotted against each other in lower panel. In the lower panel, the first quadrant is wet and warm, second is dry and warm, third is cold and dry, and fourth is cold and wet. Precipitation is shown in percentage and temperature in degrees. 2016 was close to a normal year especially in June, July was wetter, and August was warmer and drier than the normal for the selected stations. Only Kise, Flisa, Kongsberg, and Aurskog have rain gauges.

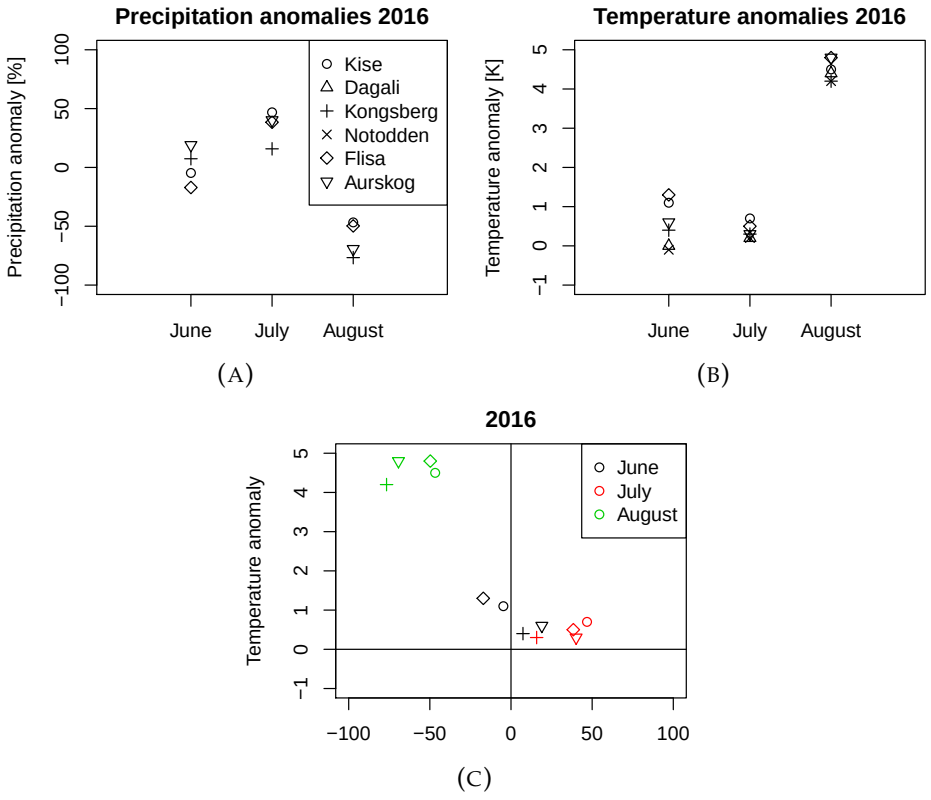


FIGURE 8.2: A) Precipitation anomalies for a selection of stations, B) Temperature anomalies C) temperature vs. precipitation anomalies for June, July, and August.

8.3 Physiography Data

SURFEX requires physiography data (PGD), used to initialize and run the model. In this study, the database used for land cover types is ECOCLIMAP-II (Le Moigne, 2012), which has a spatial resolution of 1km with a global coverage. Different soil texture databases provide the required fields for SURFEX, including Harmonized World Soil Database (HWSD) and FAO. In the operational forecasting model AROME-MetCoOp and in the SMOS retrieval model, clay and sand fields are taken from the FAO database. HWSD is a newer and more detailed data set. However, it has some unwanted features, note the contrast along the border between Sweden and Norway indicated in Fig. 8.3. The FAO data set is chosen in this study, to be as consistent as possible with the forcing data.

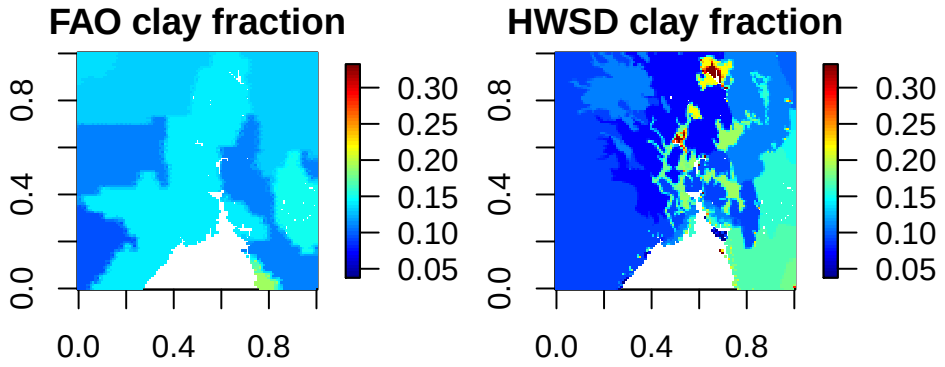


FIGURE 8.3: Clay fractions from FAO data set (left) and HWSD (right) indicate different patterns in detail between fields.

8.4 Initialization

In this study, SURFEX is allowed to spin up for one month ahead of the experiment start to allow the model states to adapt to the forcing. The model is initialized with default SURFEX initialization values for temperatures and soil moisture. Snow cover and amount are set by default to zero. Ideally the model should be run from late summer the previous year to accumulate snow during the winter, and then melt it in the spring. To save simulation time, the spin up starts 1st of May where most of the snow is gone. Fig.8.4 show the domain average and standard deviation of soil temperature and moisture for the experiment period including the spin up in front. The experiment begins around 750 hours (one month). The figure indicates that the top 5 to 6 layers have lost most of their memory of the initial conditions with regard to the soil moisture. However, since the deep layers (8-10) have increasing soil moisture content throughout the period, water continually flows downwards, and the model has not really found its steady state. The spin up is assumed to be sufficient for this study.

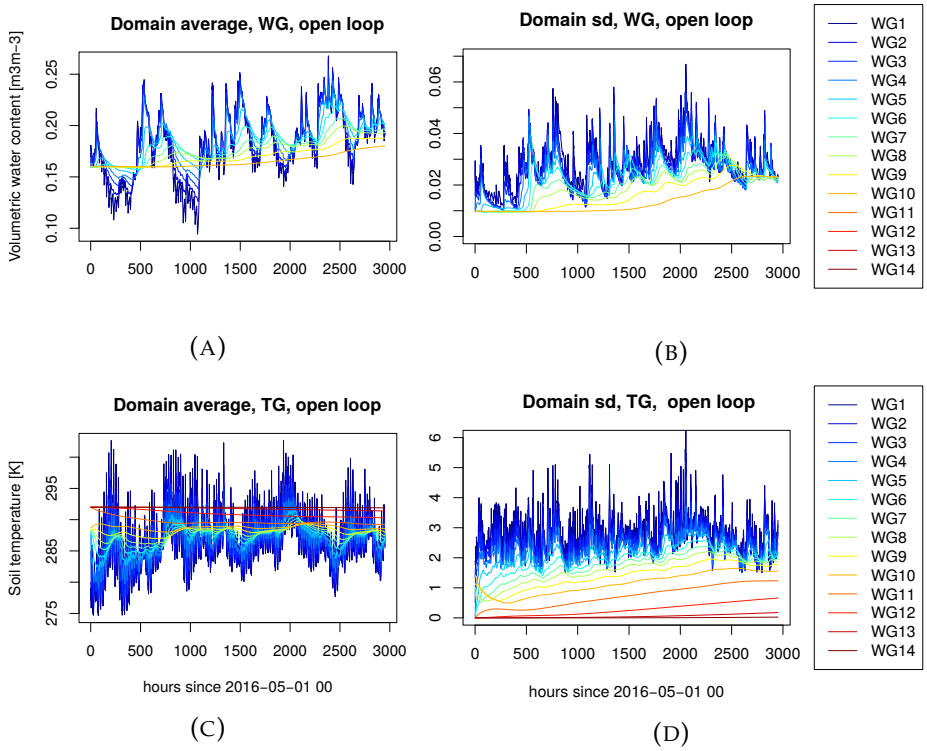


FIGURE 8.4: Time series of domain mean (left) and standard deviation (right) soil moisture (top) and soil temperature (bottom) in all model layers from 1st of May to 1st of September 2016. WG1 (blue) corresponds to the superficial soil layer, and WG14 (red) to the bottom soil layer at 12 meter. Note that soil moisture is only calculated down to the tenth layer, which signals the bottom of the root zone. The experiment start after one month, about 750 hours.

8.5 Forcing

The ground is affected by the atmosphere in multiple ways, such as precipitation, downward radiation, both long and short wave, and other energy fluxes described in Ch. 2. When SURFEX is decoupled from an atmospheric model, it needs input from the atmosphere, forcing. In this work, atmospheric input is obtained from AROME MetCoOp. The forcing consist of variables at lowest model level, which are needed by SURFEX as boundary conditions. Input variables required by SURFEX are: atmospheric temperature, pressure and specific humidity, rain and snow fall, wind speed and direction, long-wave, direct and diffuse shortwave radiation and flux of CO2.

8.6 Experiments

The offline simulation used to spin up the model state throughout May, continued to September 1st. No DA is performed and the simulation is used as a control in this study. Two simulations with DA are performed, one with observations from SMOS and one with observations from the SMAP satellite. All three runs use the same forcing. The assimilation window of the two DA experiments is 6 hours.

name	DA	obs.	obs. error
open loop	-	-	-
sekf_smos	SEKF	SMOS	0.042
sekf_smap	SEKF	SMAP	0.032

TABLE 8.1: List of experiments

Part III

Results and Discussion

Chapter 9

Model Performance

9.1 Validation of Open Loop

The control simulation, open loop, is driven purely by forcing from operational AROME-MetCoOp data. Comparing variables from these two will reveal differences introduced by the different surface schemes, ISBA-DIF and force restore, and different initialization.

Fig. 9.1 shows an over all good correlation for both soil temperature and superficial soil moisture except in some mountain areas. Total root mean square difference (RMSD) is $2.7\,\mathrm{K}$ for soil temperature, and $0.04\,\mathrm{m}^3\,\mathrm{m}^{-3}$ for soil moisture. The most probable cause for the larger RMSD of temperature in the western mountain regions, is the poor initialization of snow in the open loop. Nevertheless, we observe that the same area has relatively low RMSD in soil moisture. The mountain areas in west, where the operational model is likely to have snow cover in parts of the experiment period, a warm bias exist for the open loop. Snow reflects more of the incoming solar radiation, and increases the heating of ground and air. A wet bias is observed over the whole domain reaching as high as 10-$20\,\%$ of the dynamic range.

Similar characteristics are found in Fig. 9.2 for diagnostic variables. Open loop 2 meter temperature (T2M) and specific humidity (Q2M) are in very good agreement with the operational model. Improper initialization of snow is also shown in the screen level variables with warm temperature and moist humidity bias.

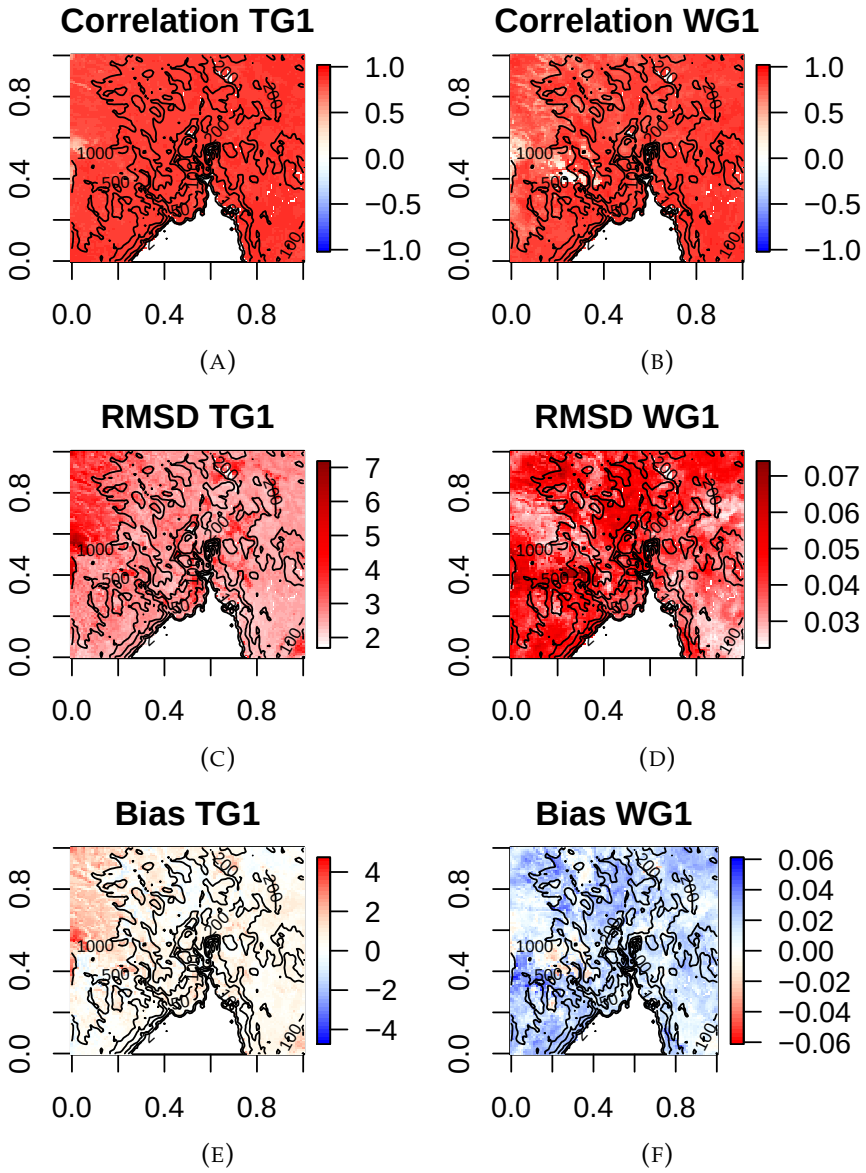


FIGURE 9.1: Comparison of prognostic variables. Open loop simulation with ISBA-DIF and operational AROME-MetCoOp with force restore. superficial soil temperature TG1 (left) and superficial soil moisture WG1 (right). Upper panels show correlation coefficient, middle panels show RMSD, and lower panels show mean open loop minus mean AROME. map axes as Fig. 8.1.

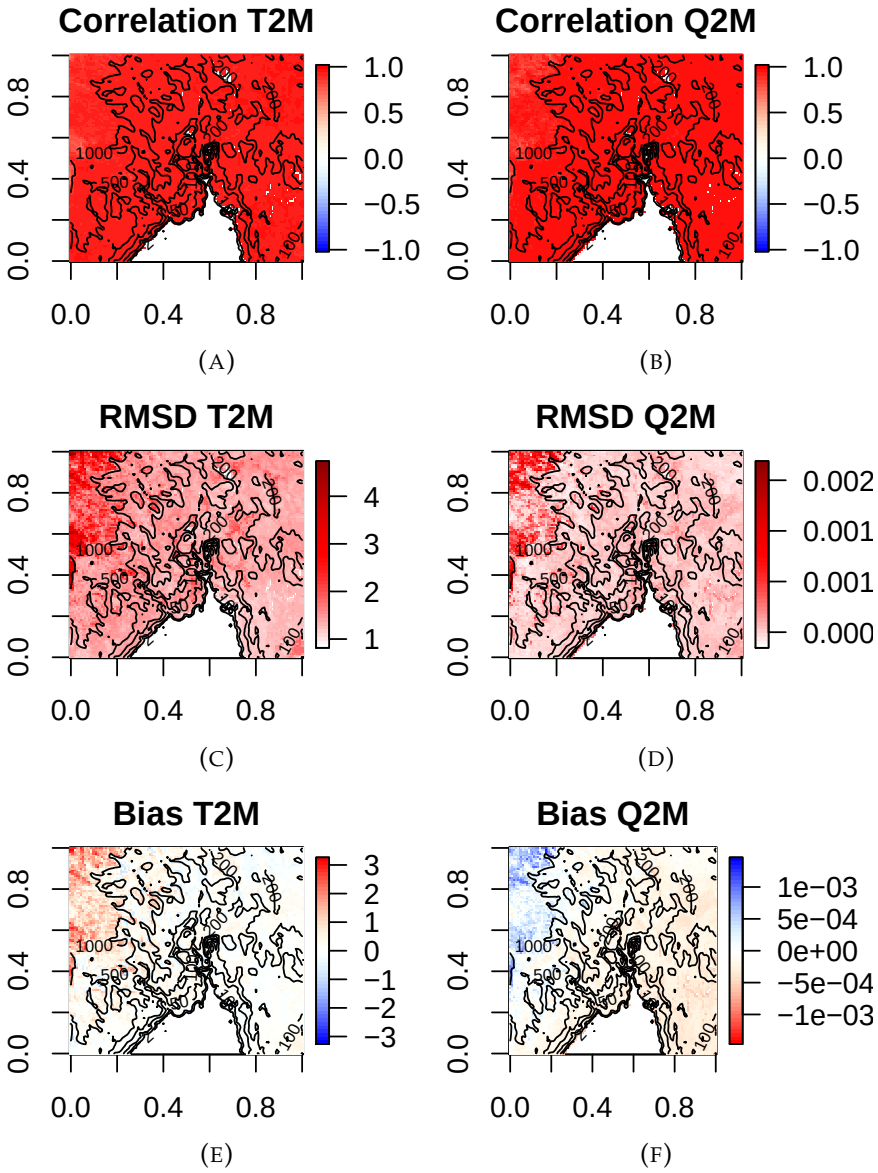
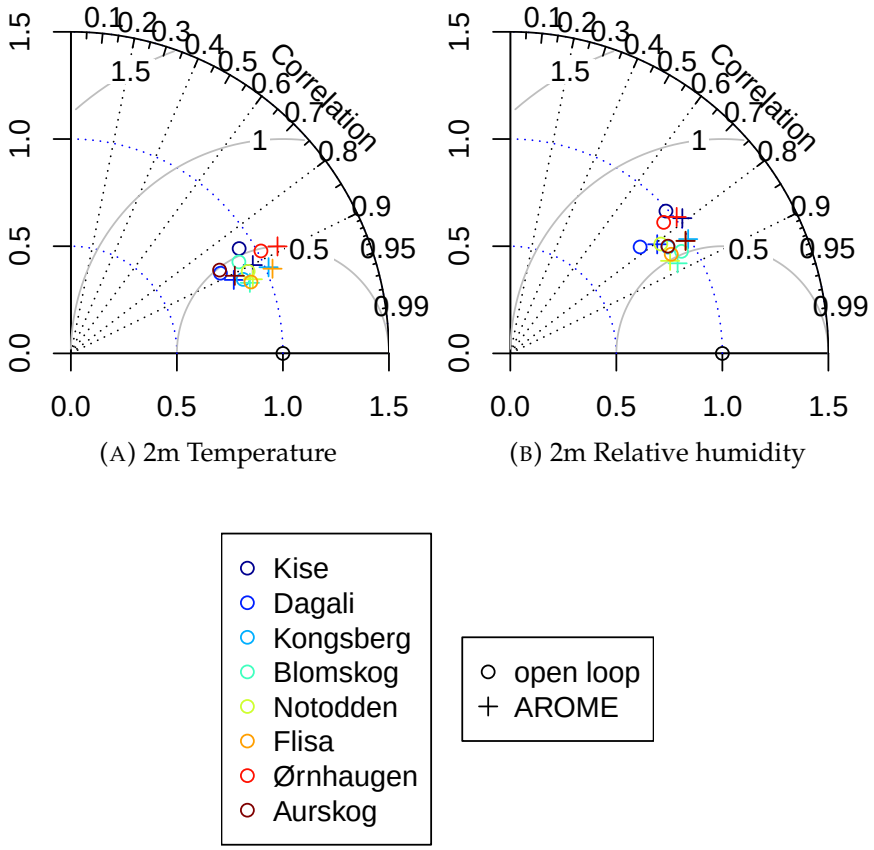


FIGURE 9.2: Comparison of diagnostic variables. Open loop simulation with ISBA-DIF and operational AROME-MetCoOp with. 2 meter temperature T2M (left) and specific humidity (right)

Taylor diagrams presented in Fig. 9.3, show that the operational forecast and the open loop perform very close compared to in situ stations. The operational model has more variability, for both air temperature and relative humidity, and slightly better root mean square error (RMSE). Over all, AROME-MetCoOp has better performance than the open loop. Since the open loop is forced with atmospheric input from the operational model, the full potential of the more advanced surface scheme will be suppressed and it is prevented from evolving naturally. The reduced variability of the open loop could be evidence of this. In episodes of possible feedback processes between surface and atmosphere, it is clear that a coupled model system will have an advantage.



(A) 2m Temperature (B) 2m Relative humidity

FIGURE 9.3: Taylor diagram: Reference observation (2 meter temperature (left) and 2 meter relative humidity (right)) is located on the x-axis (1,0). Distance from origin indicates the normalized standard deviation of the signal, azimuthal angle is related to the correlation coefficient, and distance from the observation represents the RMSE of the model.

9.2 Spatial and Temporal Structures of Soil Moisture

Soil moisture and soil parameters in general, have a large spatial variability (Crow et al., 2012). This heterogeneity has many scales and vary with location. The heterogeneity is dependent on precipitation patterns, vegetation, soil texture, and topography. The spatial resolution of soil parameters are limited in numerical models. Data sets used in this work have spatial resolution of 1 kilometer, and an average over multiple cells are used in the model with 2.5 kilometer resolution. The averaged values could provide sufficient detail to calculate a energy balance for a grid cell of the same size. Nevertheless, the true variation within a grid cell could be large. Satellite products, with about an order of magnitude larger pixels, will not capture every feature in the model.

Following Korres, Reichenau, and Schneider (2013), scaling analysis is performed to reveal how the spatial soil moisture pattern in the model changes through up-scaling. The spatial auto-correlation is calculated first

for the original grid, then a new grid is created by averaging 2 by 2 cells into a new cell reducing the resolution. This process is continued several times to investigate who information is gradually lost. The spatial autocorrelation coefficient used is the Moran's I, formulated in Eq. (9.1), where only neighboring cells are used. The index provides information about how the values are distributed in relation to other values. A value of 1 indicates that all high values are grouped together and low values are grouped together, zero indicates random horizontal distribution and -1 indicates that the values are evenly distributed, or dispersed. This method should determine to what degree satellite products with coarse resolution can capture patterns we also find on finer NWP scales.

$$I = \frac{N}{W} \frac{\sum_i \sum_j w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{\sum_i (x_i - \overline{x})^2} \tag{9.1}$$

N is the number of points counted by i and j, W is the sum of w_{ij}, where w is a weight matrix, which is 0 for all values except neighboring cells with the weight 1 in this study. $\overline{x}$ is the average of all values x.

Passive satellite soil moisture L-band retrieval products have, as described in Ch. 3, a horizontal resolution of about $30\,\mathrm{km}$. NWP models have typical spatial resolutions of 10-$20\,\mathrm{km}$ for global models, and below $3\,\mathrm{km}$ for local area models. The fine resolution of local models is crucial to resolve small scale processes with potentially large impact on society, like convective precipitation. Chen and Avissar (1994) found that horizontal heterogeneity in soil moisture caused circulations that lead to convective cloud formation. In Fig. 9.4, the soil moisture anomaly at August 23rd during the experiment is shown together with the same field upscaled multiple times. The lower right panel has a grid resolution of approximately $36\,\mathrm{km}$, which corresponds to the resolution of the SMAP satellite product. The wet anomaly south-west and dry anomalies south-east and north-west are traceable, but significantly weakened. The soil moisture heterogeneity in Chen and Avissar (1994) was $75\,\mathrm{km}$ wide, at this scale the anomalies shown in the middle row in Fig. 9.4 are still pronounced, but some areas have deviating values, e.g. at (0.4, 0.3), close to Larvik a dry anomaly is no longer traceable. The corresponding satellite anomalies are shown in Fig. 9.5, where some of the features are recognizable.

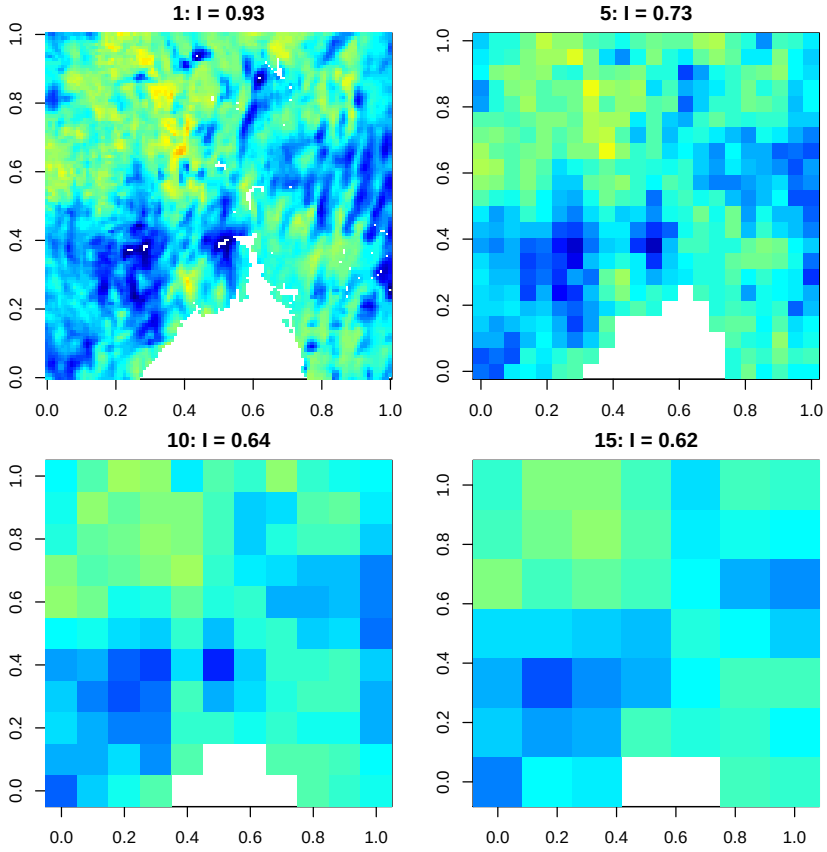


FIGURE 9.4: Aggregated grid, soil moisture anomaly at August 23rd 2016 (randomly chosen). Colors have same zero as Fig. 9.5. The Moran's I, spatial auto correlation coefficient with respect to. neighboring cells, are given on top of each panel indicating the degree of clustering.

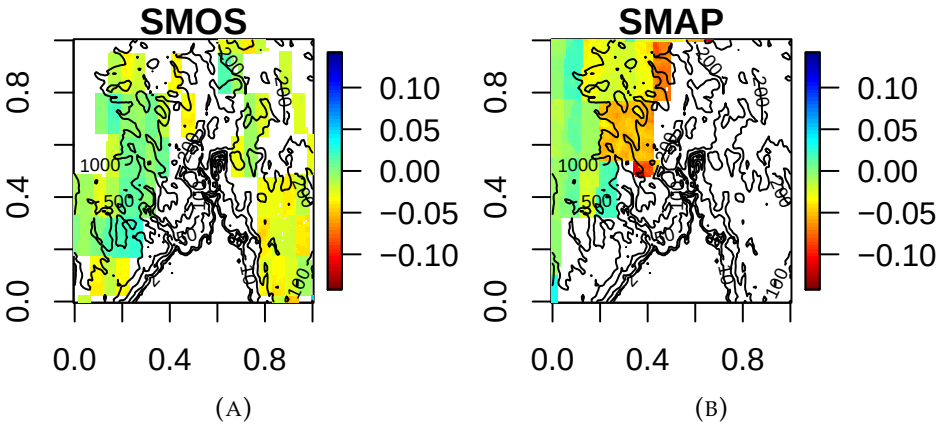


FIGURE 9.5: Observation anomalies for the same time as Fig. 9.4, SMOS (left) and SMAP (right).

9.3 Computational Cost

SURFEX is a relatively cheap model to run, specifically in offline mode. Testing of model setup early in this study, point simulations were performed locally on a laptop. However a post processing cluster at the Norwegian meteorological institute was used for simulations over a domain. Both EKF and EnKF require multiple offline simulations to perform an analysis. The EKF needs one run for each control variable plus a control and the EnKF needs an ensemble. Such a computing problem is called *embarrassingly parallel* [1], which mean that no extra development is needed to run the simulation in parallel. Table 9.1 shows computation time in three different computation setups, for the left and mid column, the ensemble/perturbations are run in a sequence. The right column run the ensemble parallel, making the simulation several times faster.

single processor	16 proc sequential ens.	4 proc per sim. parallelized ens.
15min	3min	1min

TABLE 9.1: Example of time usage for 6 hour forecast with analysis SEKF with seven control variables (8 offline runs) for 12321 data points

[1] https://en.wikipedia.org/wiki/Embarrassingly_parallel

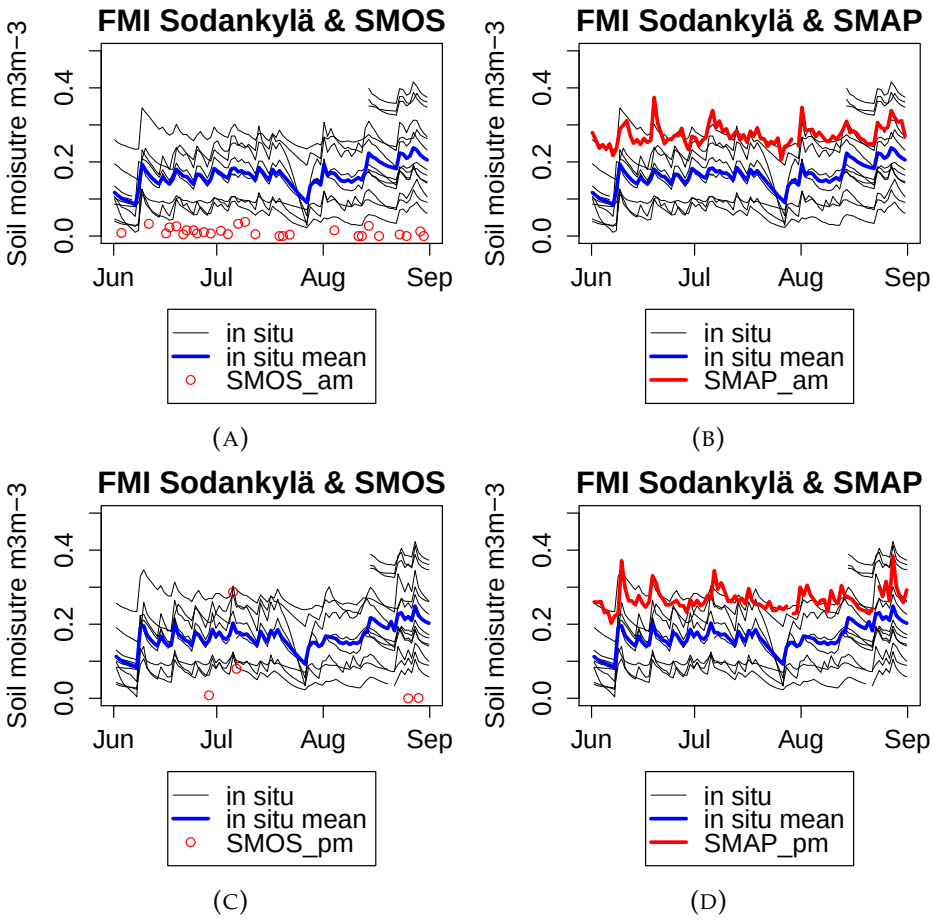
Chapter 10

Observations

10.1 Quality of Satellite Data

10.1.1 Comparison Against an In Situ Network

In validation of soil moisture satellite retrievals, comparison with direct measurements of the quantity under study is a natural first approach. The main problem with this method is that a remote observation cover a large area, whereas an in situ measurement is only representative for a small area around it, see Sect. 3.3. If multiple observation stations are spread out in an area, local variations can be canceled out by taking a mean value over these stations. The station mean should be closer to what the satellite observes. No multistation soil moisture network exist in Norway. The closest multistation network with data for the experiment period is located in Sodankylä, in northern Finland. Sixteen stations with multiple sensors measure soil moisture at different depths. Measurements from these stations at $5\,\mathrm{cm}$ depth are plotted together with soil moisture retrievals from SMOS and SMAP in Fig. 10.1. Morning retrieval (am) and evening (pm) are plotted separately to reveal differences. SMOS data are almost absent and the correlation coefficients, shown in Fig. 10.2 are not statistical significant. SMAP data has, with only one retrieval missing, values every day and agrees well with the in situ measurements. On average over all stations, am and pm SMAP retrievals have a correlation coefficient of 0.42, and 0.47 respectively, and above 0.8 for certain stations, shown in Fig. 10.2. Note that the satellite data used in this comparison are not screened in any way, all provided data are used. The satellite data are not normalized to fit the in situ observations, and a bias of about $0.1\,\mathrm{m}^3\,\mathrm{m}^{-3}$ is found.



(A) (B)

(C) (D)

FIGURE 10.1: FMI Sodankylä soil moisture network JJA 2016: Time series showing bias between stations and satellite soil moisture values. 6 am values are used from in situ stations in the upper panels, and values at 6 pm are used in the lower panels.

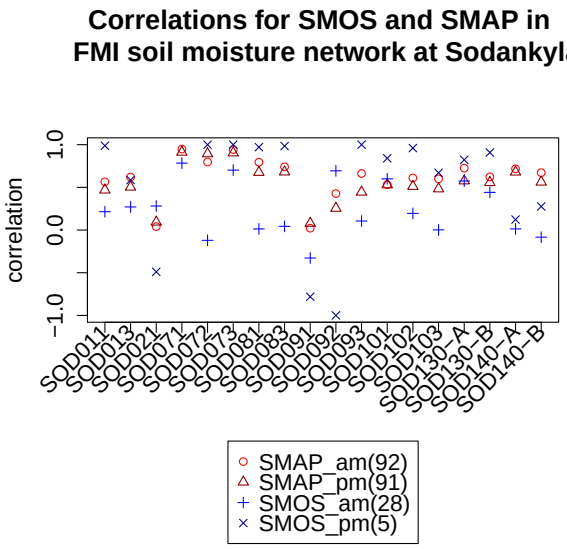


FIGURE 10.2: FMI Sodankylä soil moisture network JJA 2016: Correlation coefficients between satellite retrieval from SMOS, and SMAP and In situ soil moisture at $5\,\mathrm{cm}$, in situ measurements at 6 am and 6 pm are used for corresponding satellite retrievals. 6 am retrievals (light colors) and 6 pm (darker colors).

10.1.2 Quality Flags and Screening Procedures

Level 3 soil moisture products from SMOS and SMAP come with information about the retrieval. In the SMOS data output, a retrieval data quality index (DQX) is included, this is the product of the standard deviations of the retrieval uncertainty, and a measure of the amount of radio frequency interference (RFI). SMAP has a similar value, which is *Estimated '1-sigma' error of the soil moisture output parameter*. Both of these values have the unit $\mathrm{m}^3\,\mathrm{m}^{-3}$. Unfortunately, the SMAP soil moisture error variable was missing from the data used in this study.

In addition, retrieval flags are included that provide information about surface conditions, weather, and landscape, during the retrieval time that could reduce the quality of the retrieval. If a pixel has frozen soil, is covered with snow, contain high fraction of open water or strong topography, the retrieval is usually skipped (Kerr et al., 2012). However, the data sets could still contain data of varying quality, and a screening procedure to remove unreliable values has to be performed.

When evaluating the satellite data, the distribution of soil moisture can provide a good overview of the data quality. The soil moisture retrievals are provided in $\mathrm{m}^3\,\mathrm{m}^{-3}$, hence the values should lie within the dynamic range described in Sect. 2.2.1. The dynamic range used as reference in the following, is taken from SURFEX and may vary from the model used in the retrieval.

The histogram of retrieved values during the experiment period, JJA 2016, shown in Fig. 10.3, reveals a high frequency of a few soil moisture values. For SMAP, 90% of these observations were flagged with *The retrieval for this algorithm was attempted but failed*. These soil moisture values are also

above the maximum porosity of the SURFEX domain, and are considered unreliable.

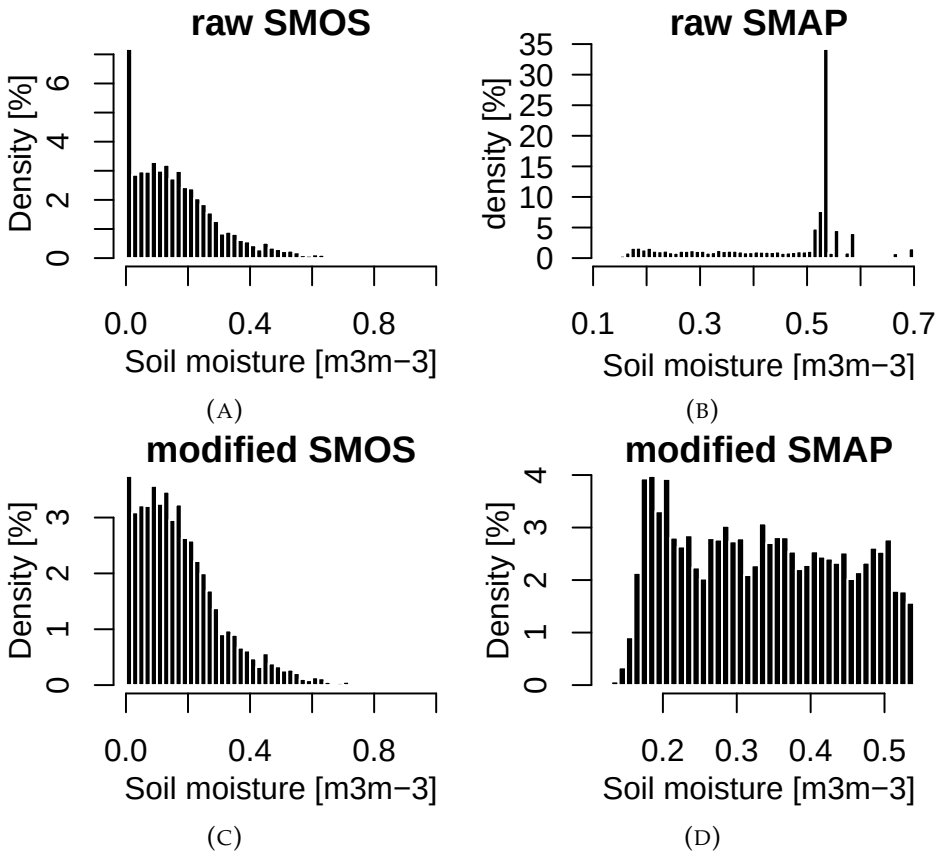


FIGURE 10.3: Distribution soil moisture values before (top) and after (bottom) modification (clean-up) SMOS (left) SMAP (right)

For the SMOS data, a clustering toward zero is found, and a large amount of values equal to zero. Figure 10.4 shows the provided soil moisture DQX (low values are better) plotted versus soil moisture. We notice a dozen values with low quality close to zero, and a general decrease in quality with increasing soil moisture. Good quality values are also found close to zero, but since the spread of the quality is so large compared to values within the dynamic range (blue lines), and the fact that these values are almost zero, below hygroscopic water content thus unphysical, these are considered unreliable. Regarding the values above the higher limit, water content at saturation, despite low reliability, they are few in number and will not influence the distribution when normalizing to model climate.

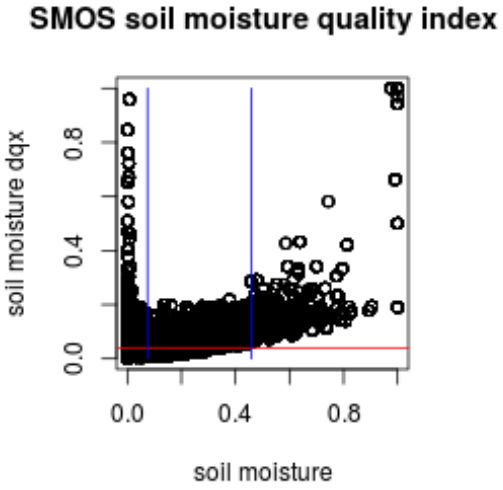


FIGURE 10.4: SMOS DQX vs soil moisture, horizontal (red) line indicates the mission goal of $0.04\,\mathrm{m}^3\,\mathrm{m}^{-3}$, vertical (blue) lines indicates the mean dynamic range of SURFEX (wsat-wwilt) in the domain

In De Lannoy and Reichle (2016) a number of requirements are set to the quality of SMOS data to be used. For the domain of the experiments in this thesis, in the certain time period, setting these requirements will leave no values left for the assimilation in most parts of the domain. Figure 10.5 shows a histogram of the observations after each step of the screening, and a heat map of observations that are left after the screening is shown in Fig. 10.6a.

Wrona et al. (2017) conclude that the SMAP soil moisture product should be used with extreme caution over high latitudes, and that the Equal Area Scalable Earth-2 (EASE-2) grid interpolation of the product is responsible for faulty values. The SMAP and SMOS soil moisture products are interpolated on the global EASE-2 grid, which has a cylindrical projection. A cylindrical projection is not very suitable for high latitudes as it distort shapes significantly.

Jeu et al. (2012) found that SMOS provide low soil moisture values over northern regions compared to retrievals from the active sensors Advanced Microwave Scanning Radiometer for EOS (AMSR-E) and Advanced SCATterometer (ASCAT). No explanation was concluded, and they encourage further investigation. However, the results presented in Fig. 10.1a and Fig. 10.3a, coincides with their findings of a dry bias.

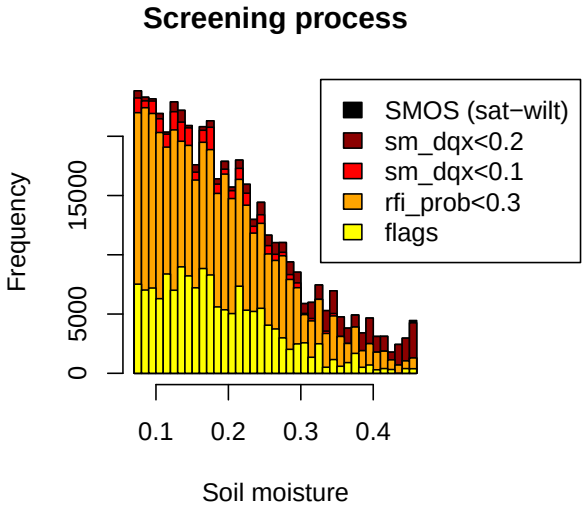


FIGURE 10.5: Histogram of SMOS data after each step in the screening process proposed by De Lannoy and Reichle (2016).

The SMAP data output has a quality parameter, that contains a flag for recommended quality. Figure 10.6b shows number of retrievals with recommended quality during JJA 2016. Only 2.3 % of the SMAP retrievals have the flag for recommended quality. All recommended values are located within a few pixels in high mountain areas.

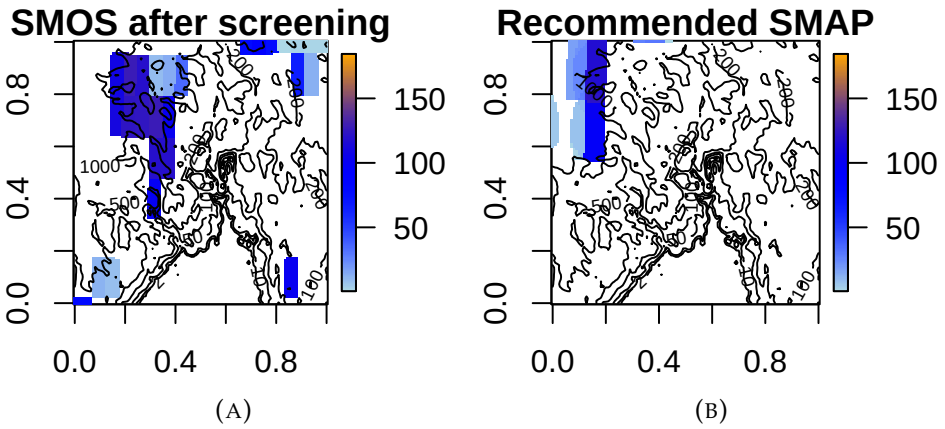


FIGURE 10.6: Number of observations during JJA 2016: heatmap of SMOS soil moisture retrievals after screening following De Lannoy and Reichle (2016)(A), heatmap of SMAP pixels with recommended quality(B).

In this study, for both satellites, values that appear a relatively large amount of times, 450 and 1000 times for SMOS and SMAP respectively, are removed from the data set prior to the normalization process described in Sect. 6.2. This approach ensures that the distribution of observations after the normalization is of similar shape to the SURFEX control. The number of observations after the screening procedure is shown in Fig. 10.7. Large parts of the domain, have no observation throughout the experiment period.

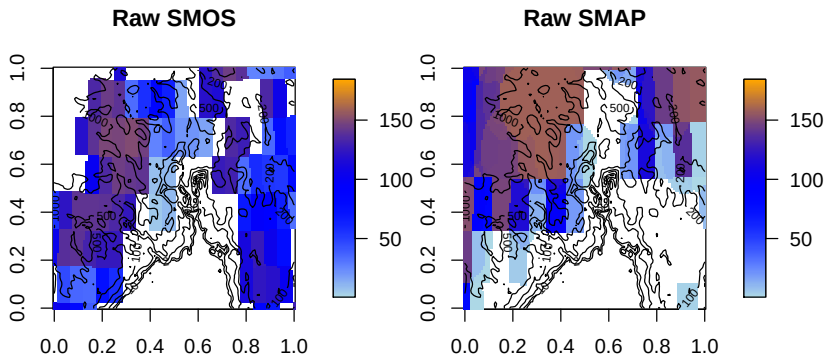


FIGURE 10.7: Number of observations from each satellite, SMOS(left) and SMAP (right) during JJA 2016 used in experiments.

The second row of Fig. 10.3, shows the distribution of satellite soil moisture values after the quality control. The soil moisture values outside the dynamical range of available water are intentionally not removed, because of the sparse data coverage.

Fig. 10.8 shows the consequence of removing the data with respect to the mean retrieval, where the SMAP data is modified significantly. SMOS is over all less affected. Fig. 10.8 also indicate a large bias between the two satellites.

Taking the average of the DQX values of all soil moisture retrievals from SMOS, after the screening, later used in the experiment, results in a value of $0.055\,\mathrm{m}^3\,\mathrm{m}^{-3}$, higher than the mission goal of $0.04\,\mathrm{m}^3\,\mathrm{m}^{-3}$.

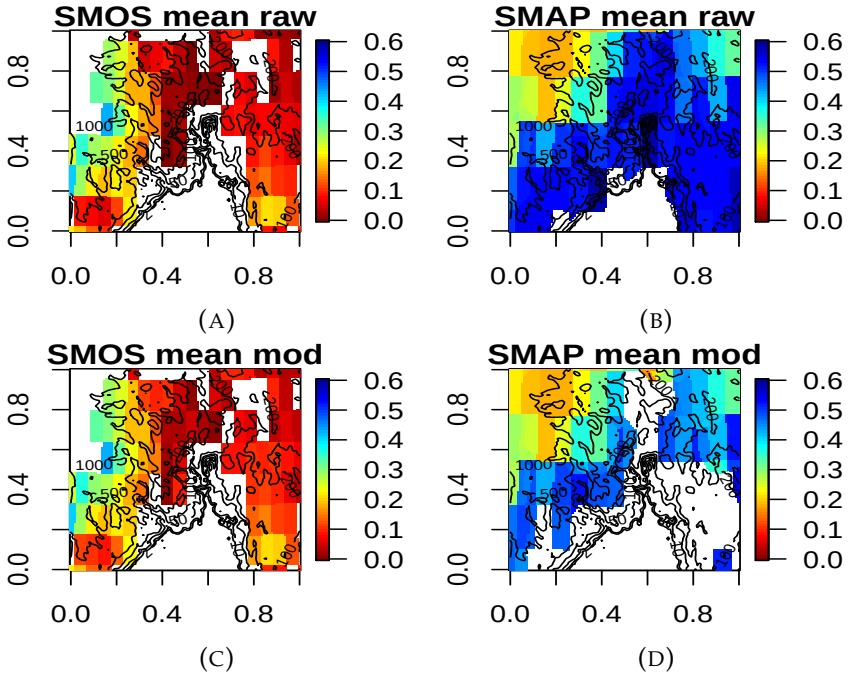


FIGURE 10.8: Mean raw SMOS and SMAP soil moisture (top) and mean modified SMOS and SMAP soil moisture (bottom).

10.2 Comparison with Open Loop

A similar evaluation of data done in Ch. 9, is done for satellite products after the modifications described in Sect. 10.1.2. Satellite soil moisture data are compared with the open loop simulation, and results are shown in Fig. 10.9. The upper panel, showing correlation coefficients, indicates that SMOS data are over all weakly negative correlated. SMAP data are more spread, but have some areas with positive correlation. Pixels with high negative or positive correlations, have a low number of observations and have no significant result, indicated in Fig. 10.7. SMAP has a relatively large wet bias for the whole domain. SMOS has both positive and negative bias depending on the location. Since areas with no data cover also are white, use Fig. 10.7 as a reference to identify these.

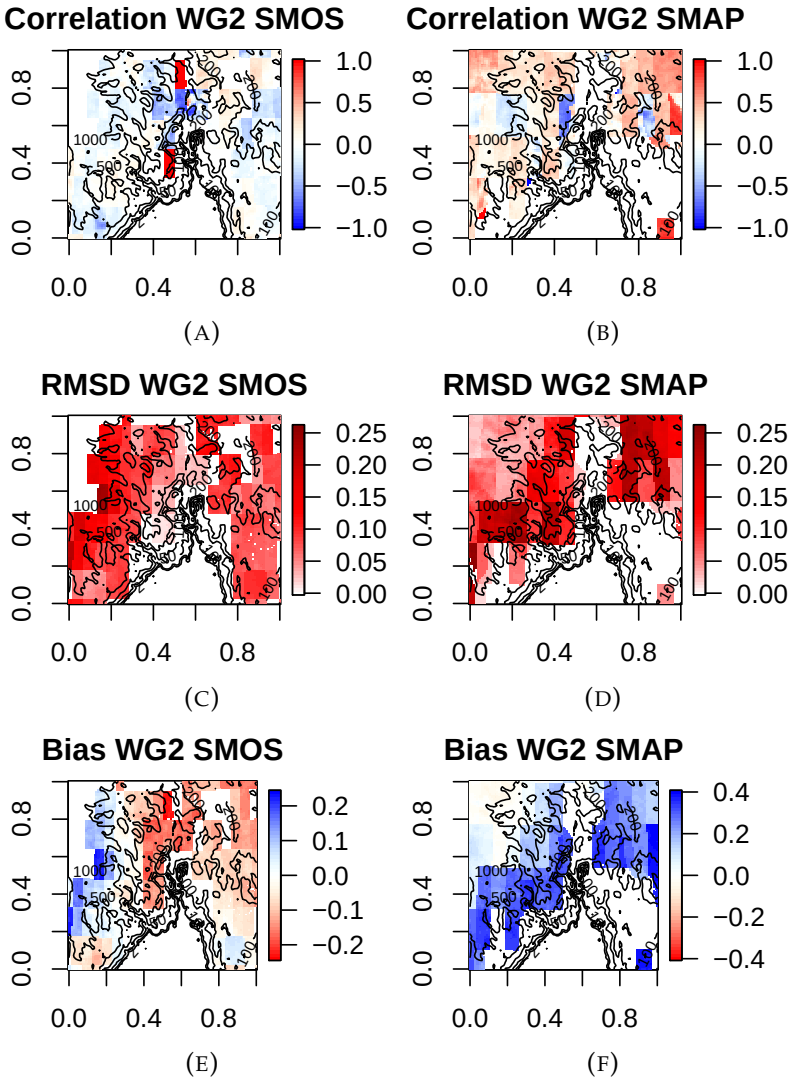
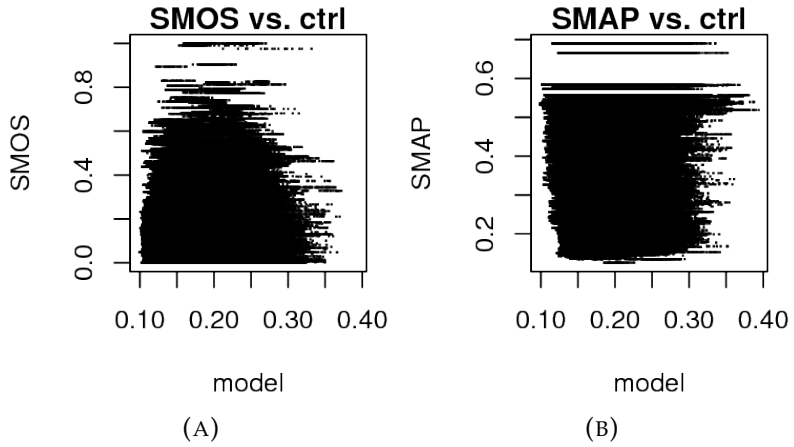


FIGURE 10.9: Comparison of satellite soil moisture products, SMOS (left) and SMAP (right), with open loop. Upper panels show correlation coefficient, middle panels show RMSD, and lower panels show bias. The modified data sets are used in this figure

Figure 10.10 shows satellite data plotted against the model control for all grid cells during JJA. Satellite retrievals and the open loop model simulation show no sign of correlation.



(A) (B)

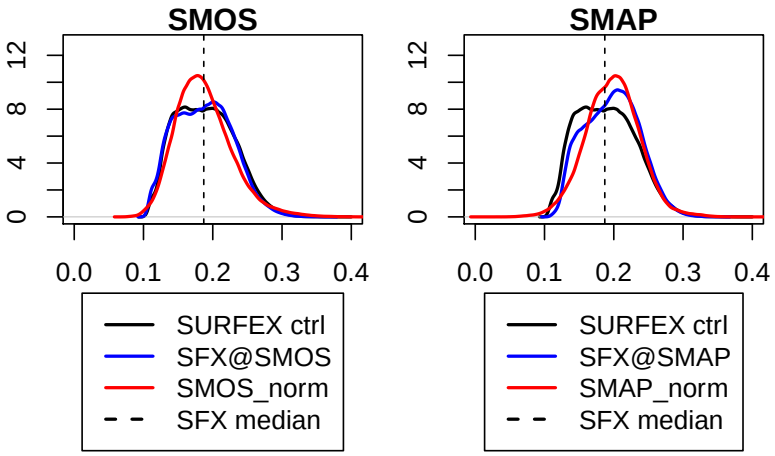
FIGURE 10.10: Satellite vs. model soil moisture, SMOS left and SMAP right.

10.3 Scaling Observations

After quality control is performed, the normalization method described in Sect. 6.2 is applied to the data. The resulting soil moisture distributions are shown in Fig. 10.11 where the model climate is included together with the results. We notice that the width of the distribution is reduced from $0.6\,\mathrm{m}^3\,\mathrm{m}^{-3}$ and $0.4\,\mathrm{m}^3\,\mathrm{m}^{-3}$ to about $0.2\,\mathrm{m}^3\,\mathrm{m}^{-3}$ for SMOS and SMAP respectively. Rather than removing unphysical values from the data, they are assumed to be representative for their respective extreme and normalization will redistribute them within the dynamical range. Normalizing the data in this manner, will reduce the variance and thus the amplitude of random errors.

The blue line in Fig. 10.11a indicates that the model values used in the normalization of SMOS are representative for the total model climate. The normalized satellite data are skewed slightly to the dry side of the model, this is caused by the skewed distribution of the raw satellite retrievals shown in Fig. 10.3c. The model soil moisture values used to scale SMAP retrievals (blue line in Fig. 10.11b) differ from the total model distribution, shifting to wetter values. Even though the satellite data are evenly distributed (Fig. 10.3d), the slightly wetter model values cause the normalized satellite data to also be shifted.

The number of possible observations is the number of grid cells times the number of satellite overpasses, $111 \times 111 \times 184 = 2.3 \times 10^6$. The number of observations after cleaning and normalization is 6.3×10^5 and 6.9×10^5 for SMOS and SMAP respectively, which is about $30\,\%$ of the total possible amount.



(A) (B)

FIGURE 10.11: Distributions of soil moisture from SMOS (left) and SMAP (right) after normalization to model climate. Black line shows the SURFEX (SFX) distribution for the entire domain for every time step, blue lines indicate model values with a corresponding observation from the satellite used in the normalization, and red line is the scaled observations later used in the assimilation.

The wet shift in model values of SMAP retrievals, suggests that the retrieval algorithm fail to provide soil moisture values in dry conditions. The number of observations for each overpass is plotted in Fig. 10.12, which show that SMAP is relatively consistent compared to SMOS, which varies to a larger degree. During the beginning of August, SMOS has a minimum of retrievals per overpass. Figure 8.4 shows that this was a relatively wet period. If SMOS fail to retrieve soil moisture in wet soil conditions, this could explain the dry shift in Fig. 10.3c.

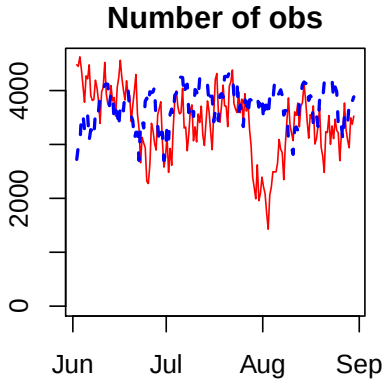


FIGURE 10.12: Number of observations from SMOS (solid red line) and SMAP (dashed blue line) within the domain at each overpass, 7 day moving average. Every grid cell in the model domain with an observation is counted, one satellite pixel can thus contribute to multiple observations in the domain.

Chapter 11

SURFEX offline data assimilation

11.1 Residuals and Innovations

A simple method to verify that the DA system is working, is to compare the residuals to the innovations (Barbu et al., 2011; Albergel et al., 2017). The residual is defined as the difference between observation and analysis (O-A). Innovation is the observation minus the forecast (O-F). The reason for doing DA is to reduce the distance between model and observation, thus should the residuals be closer to zero than the innovations.

Histograms of residuals, and innovations are shown in Fig. 11.1. The residuals have a sharper distribution than the innovations, the analyses are thus closer to the observations than the forecasts. This indicates a working DA system. For both SMOS and SMAP the distributions are close to symmetric around zero, which is expected due to the normalized satellite data. However, the efficiency of the analysis is related to the difference between the innovation and the residual, which is slightly larger on the positive side, that is when the satellite observation is wetter than the model.

In Kalman theory the errors are assumed to be normally distributed, the distribution of innovations should thus also be normal. Following Fairbairn et al. (2015), we compare the the distribution of innovations with a Gaussian curve with same standard deviation, shown with blue line in Fig. 11.1. The innovations are in fairly good agreement with the normal distribution, except that they are slightly sharper, especially SMOS.

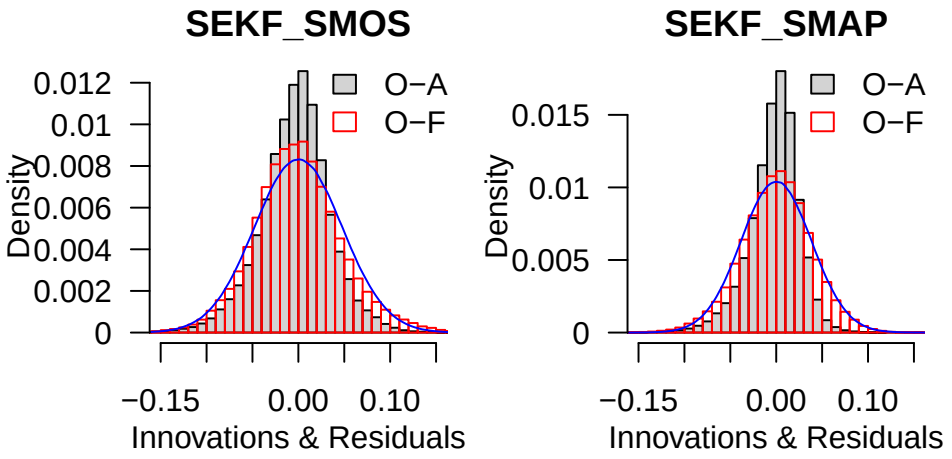


FIGURE 11.1: Distribution of residuals (O-A) and innovations (O-F) from DA with observations from SMOS(left) and SMAP respectively, blue line indicates the normal distribution with standard deviation equal to innovations.

11.2 Observation Errors

The χ^2-test presented in Sect. 7.3, resulted in consistent DA systems with observation errors of $0.042\,\mathrm{m^3\,m^{-3}}$ and $0.032\,\mathrm{m^3\,m^{-3}}$ for SMOS and SMAP respectively. Both satellite missions aim at an uncertainty below $0.04\,\mathrm{m^3\,m^{-3}}$, the obtained observation errors are thus in the right area. In Ch. 10, the satellite data are found to be of questionable quality. However, through normalization, compressing the range including the noise, these values could represent larger errors of the original data.

11.3 The Linearized Observation Operator

The diagnostics described above, only cover the observation and the model equivalent, soil moisture at 1-4 cm depth, and not the analysis of the other control variables. In a SEKF the Jacobian $\mathbf{H}$, is what provides weight to the control variables. This needs to be investigated in detail. Water transport in the soil is a slow process compared to the fast changes in weather and atmospheric conditions. During a 6 hour time window, the atmospheric forcing will only reach to a certain depth of the ground. Further, the 1-4 cm depth, will have limited communication with deeper layers.

In the light of DA it is thus a maximum depth, of which it is appropriate to make an analysis based on an observation in the top layer. Figure 11.2 shows the Jacobians for each control variable, soil moisture from 0-80 cm.

Since clay and sand fractions are important parameters for the transport of soil moisture as described in Sect. 2.2.2, pixels of high clay fraction (left) and high sand fraction (right) are thus included in the figure to clarify the differences. The values on the y-axis are the change of soil moisture content in layer 2 with respect to a perturbation in the ith control variable. A perturbation in layer 6 or 7, 40-60 cm and 60-80 cm, does not introduce a significant change in the observed layer, 1-4 cm, during a 6 hour assimilation

window. The sensitivity due to a perturbation in layer 5, 20-40 cm, is dependent on the soil texture, but in both cases it is relatively small compared to layer 2-4.

As layer 3, 4-10 cm, is quite similar in both soil texture regimes, layer 4, 10-20 cm, and the observed layer have the largest differences. The top surface layer, 0-1 cm, has low sensibility in both cases, this is because the layer is more affected by the atmospheric forcing than a small perturbation during a 6 hour assimilation window. Using a 24 hour assimilation window, Albergel et al. (2017) did not include the surface layer in the DA.

Being familiar with the physical processes of water transport in the soil, and expecting that there is a maximum depth, of which a perturbation will affect the observed layer during the assimilation window, Fig. 11.2 indicates that the Jacobian of the observation operator is calculated correctly, and that the implementation of additional control variables is done properly. Furthermore, the figure suggest that layer 6, 7, and most likely 1, can be removed from the control vector. This means that the number of perturbed simulations can be halved, and computational cost reduced.

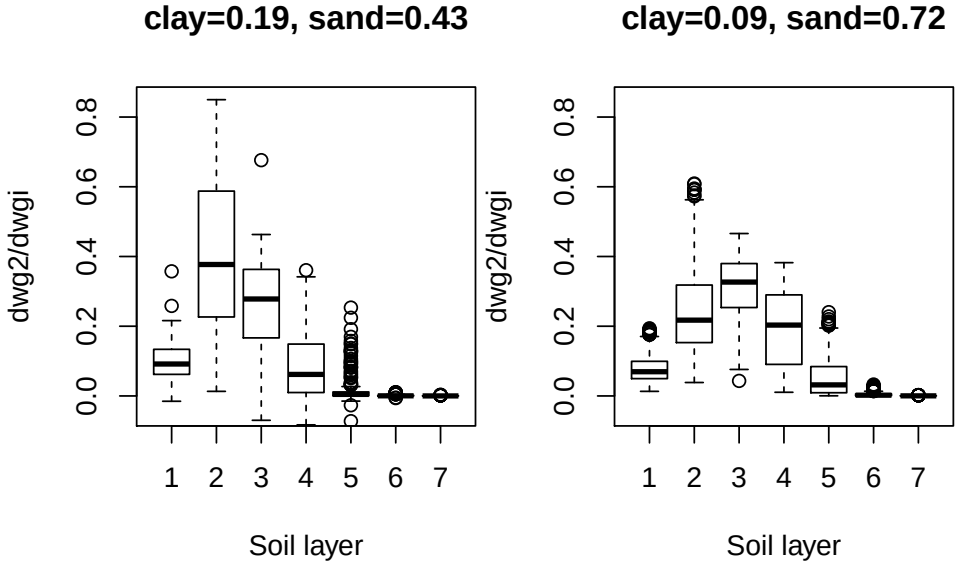


FIGURE 11.2: Jacobian **H** for a grid cell rich in clay (left) and sand (sand) respectively. The distribution of Jacobians are similar for both simulations regardless of satellite product.

The SEKF assimilation scheme adjust the weights according to the physical conditions in the specific case when the analysis is made, the spread of the Jacobians, for each control variable, in Fig. 11.2 is an evidence of this. Figure 11.3 shows the Jacobians for wet and dry events. In Fig. 11.4 the sensitivity, Jacobian, is plotted against soil moisture. The figures reveal that each soil layer have different dependencies to the observed layer with respect to soil moisture. The observed layer is most sensitive to perturbations in the same layer when the soil is dry. Perturbations in deeper layers influence the observed layer most in wet conditions.
In wet conditions, that is when the model is wet, satellite observations are more likely to be drier than the model. Claiming this is justified with the

uncorrelated, random behavior of satellite data shown in Fig. 10.10. Increments in layers below $10\,\mathrm{cm}$ are thus likely to be negative, and water is removed. For dry soil conditions observations are likely to be wetter and since the upper layers now have large Jacobians, water is added to these layers.

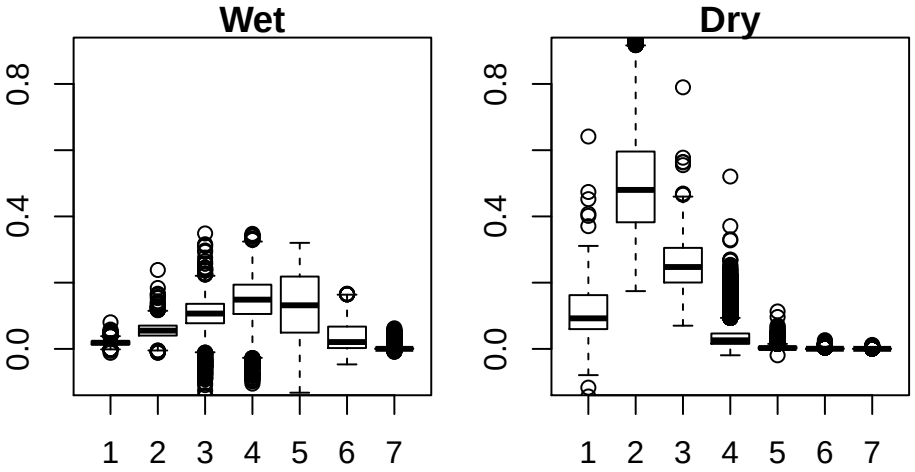


FIGURE 11.3: Jacobian **H** for wet (mean=0.278) (left) and dry (mean=0.129) (right) conditions (in the observed layer)

In Fig. 11.4 the components (control variables WG2, WG3, WG4) of the linearized observation operator are plotted against soil moisture content in the observed layer. Each control variable have a different soil moisture value, in which gives a maximum change in the observed layer. The maximum is shifted toward wetter values with increasing depth of the control variable. The differences due to soil texture are better illustrated in Fig. 11.2.

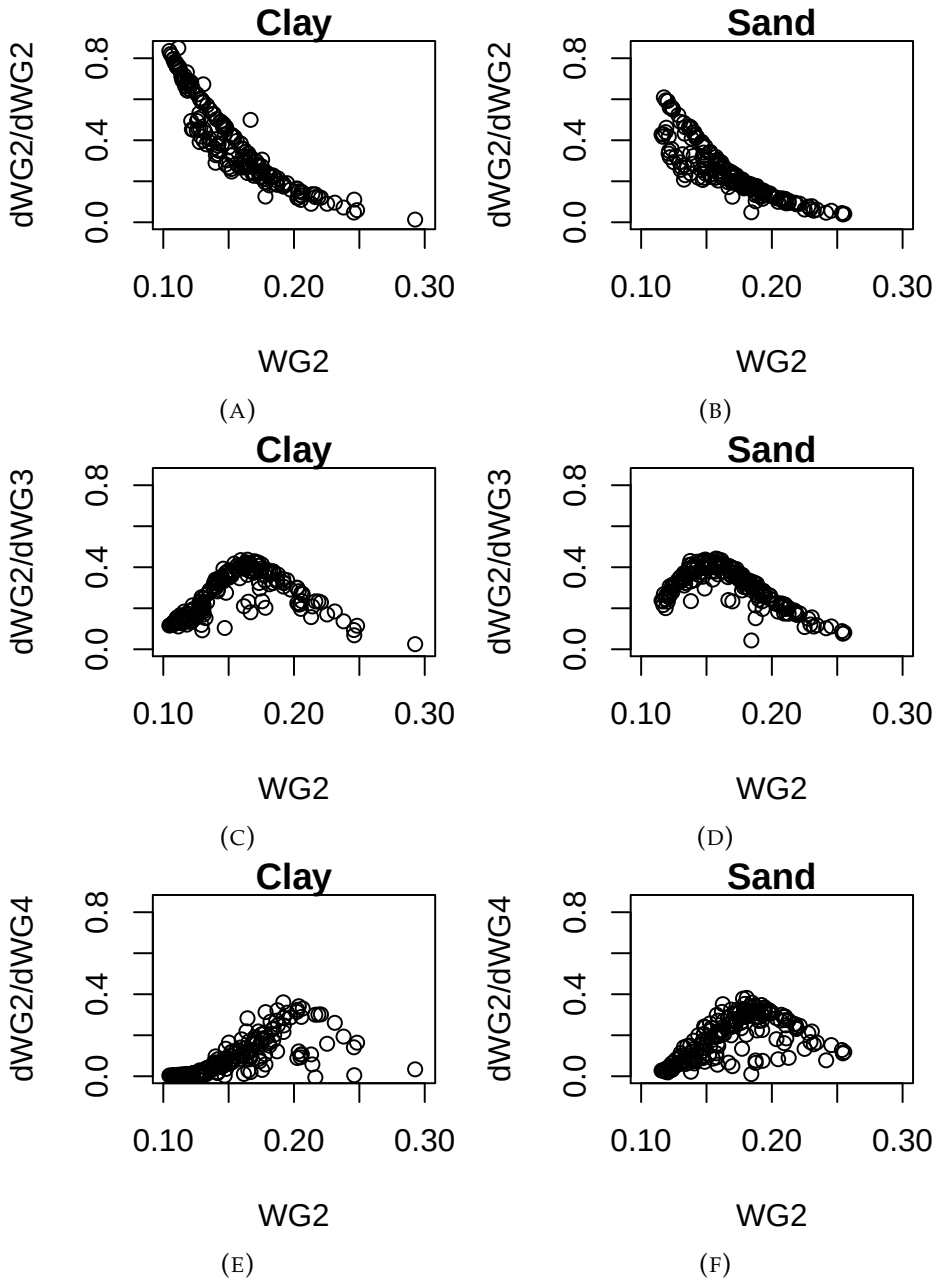


FIGURE 11.4: Jacobian vs WG2 (soil moisture in observed layer) for grid cells with high clay (left) and sand (right) fraction, for soil layers 2 (top), 3 (middle), and 4 (bottom)

11.4 Mean Increments

Mean increments relative to the mean absolute value of the increments for each control variable are shown in Fig. 11.5 and 11.6. The figures confirms the claim from the previous section. The three upper layers (0-10 cm) have a positive (wet) mean increment, and the lower 4 layers (10-80 cm) have a negative (dry) mean increment. Mean absolute increments are plotted in Fig. 11.7 on a logarithmic y-scale. Strong colors in Fig. 11.5 and 11.6 indicate an over represented sign of the increments. Most pronounced are increments in layer 1, 2, and 5.

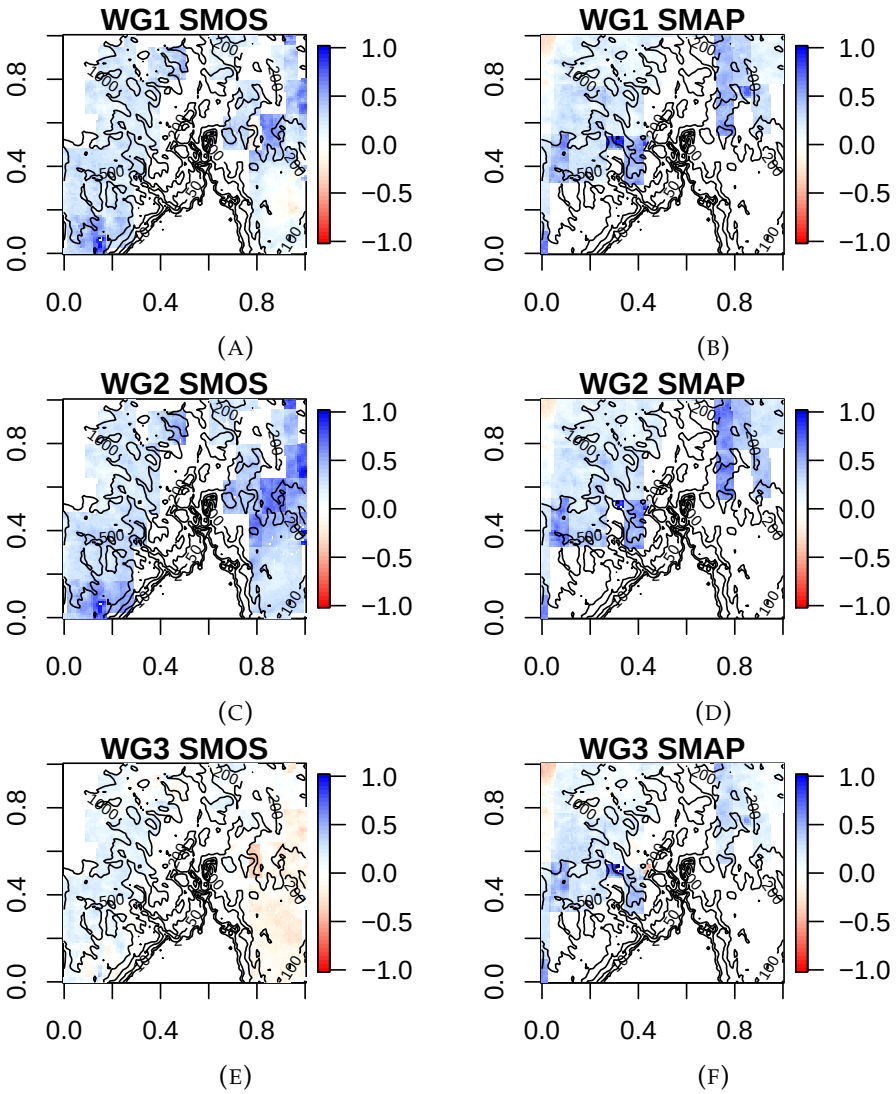


FIGURE 11.5: Mean increments relative to the mean absolute value of increments for sekf_smos (left) and sekf_smap (right), layer 1, 2, and 3

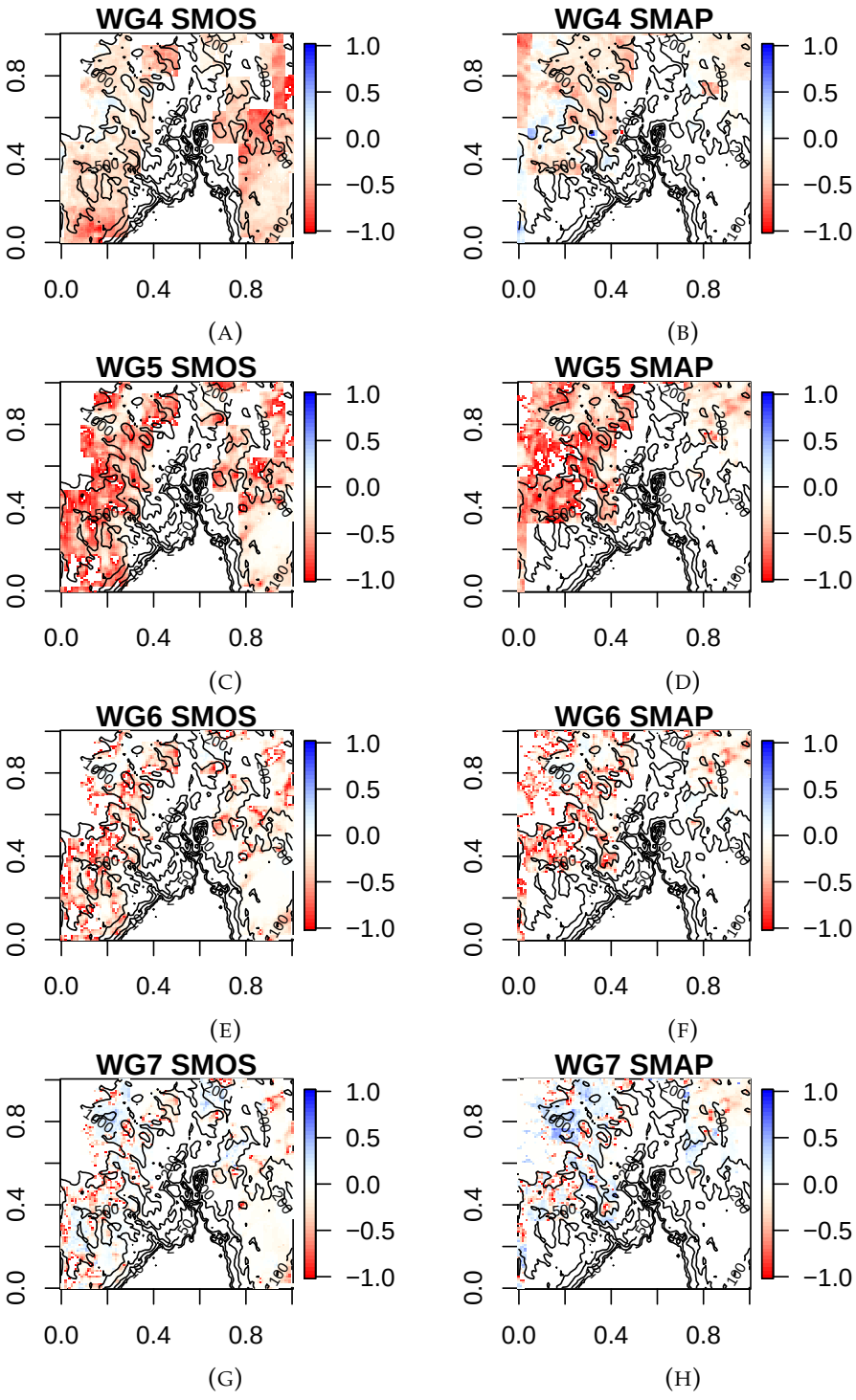


FIGURE 11.6: Same as Fig. 11.6, for layer 4, 5, 6, and 7

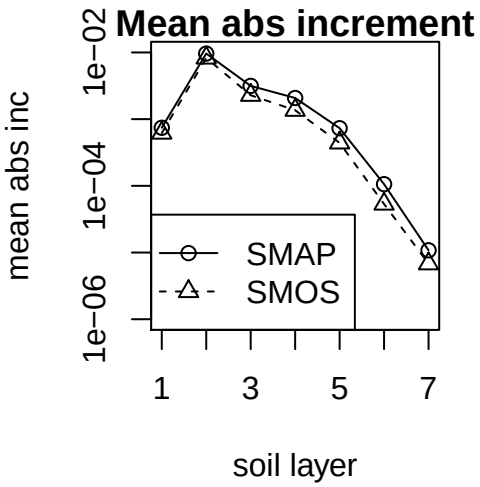


FIGURE 11.7: Mean absolute value of increments for each control variable, log y-axis

In total, for all points, throughout the three months, the means of all increments are $0.36 \times 10^{-3}\,\mathrm{m}^3\,\mathrm{m}^{-3}$ for sekf_smos and $0.34 \times 10^{-3}\,\mathrm{m}^3\,\mathrm{m}^{-3}$ for sekf_smap. The mean of all innovations has opposite sign in the two experiments, with $-0.46 \times 10^{-3}\,\mathrm{m}^3\,\mathrm{m}^{-3}$ in sekf_smos and $0.20 \times 10^{-3}\,\mathrm{m}^3\,\mathrm{m}^{-3}$ for sekf_smap. In the sekf_smos experiment, the satellite is over all drier than the model, however, in total more water is added.

The results in Fig. 11.2, Fig. 11.5, and 11.6 are similar to what Albergel et al. (2017) presented in their Table 2, and Figure 5 respectively, regardless of different observations, domain, and time period. Model set up is relatively similar in this and their study, and the tendency to add water in upper layers and remove water in deeper layers, is caused by the same model physics. If the satellite data had good correlation with the model, and the amplitude was larger, so that innovations changed signs in extreme conditions, the result would be different. Otherwise, satellite observations seem to pull the model toward the mean value, reducing extremes.

Chapter 12

Forecast Improvements

The reason for doing DA is to obtain the best estimate of the state under study. In NWP this means bringing the model closer to the true state of the system, which again should improve the forecast skill. As discussed in Ch. 11, the assimilation works in the sense that it brings the model closer to the normalized satellite observations. Is this closer to the true state, and does the increments of this magnitude improve the forecast? These are difficult questions to answer, since no detailed in situ observation network of soil moisture exist for the area. To fully evaluate the possible improvements of the forecast, simulations coupled with an atmospheric model should be studied for a number of different scenarios. In the following chapter, the available tools are used to investigate to what extent the DA of satellite soil moisture, can improve the forecast with respect to temperature and humidity.

12.1 Comparison Against Screen Level Observations

No atmospheric model is used in this study, mainly because of limited time. However, SURFEX calculates screen level variables, 2 meter temperature and humidity, as diagnostic variables. These variables are standard quantities measured at in situ stations. Note that these values are not updated in SODA, and changes in them are caused by soil moisture increments. Model results are compared with observations from a selection of stations shown in Fig. 8.1. By comparing model simulations with and without DA of satellite data, we can investigate if the screen level variables are closer or not to the observations, and thus get an indication of a possible improvement. Since SURFEX is run offline, an increased water vapor flux from the ground to the atmosphere will not change the atmospheric conditions in the next forecast. Differences in screen level variables, are thus solely local effects due to surface conditions.

In Fig. 12.1 the relative RMSE, $RMSE_{rel}$, of 2 meter temperature, see Eq. (12.1), is plotted versus the RMSD between the SEKF-run and the open loop. The relative RMSE is the relative improvement done with DA with respect to in situ measurements of temperature. For all stations and with both satellites, the model simulations with DA of soil moisture have a small improvement in 2 meter temperature, on the order of 0.1-1 percent relative to the open loop. The number of observation stations used is very small compared to the large domain, and the result might not be statistically significant.

$$RMSE_{rel} = \frac{RMSE_{da} - RMSE_{ol}}{RMSE_{ol}} \tag{12.1}$$

$RMSE_{da}$ is the RMSE of the SEKF-run and $RMSE_{ol}$ is the RMSE of the open loop to observations.

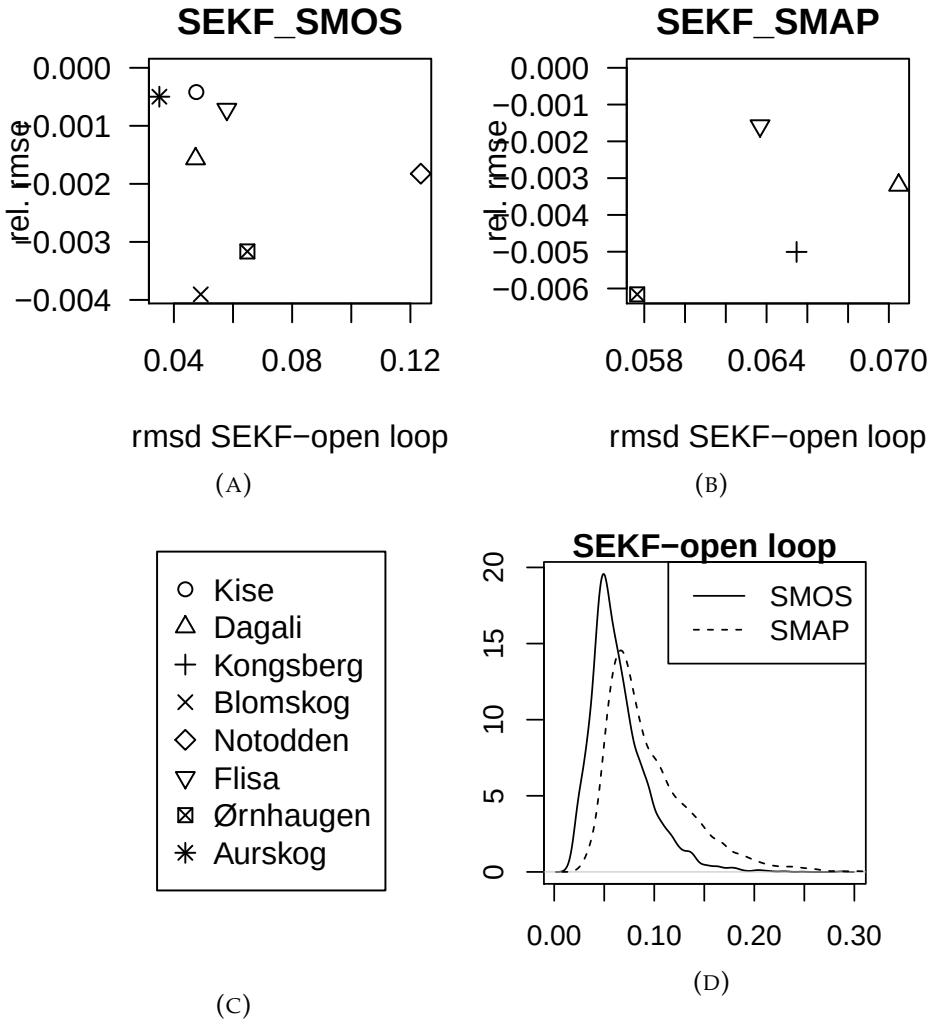
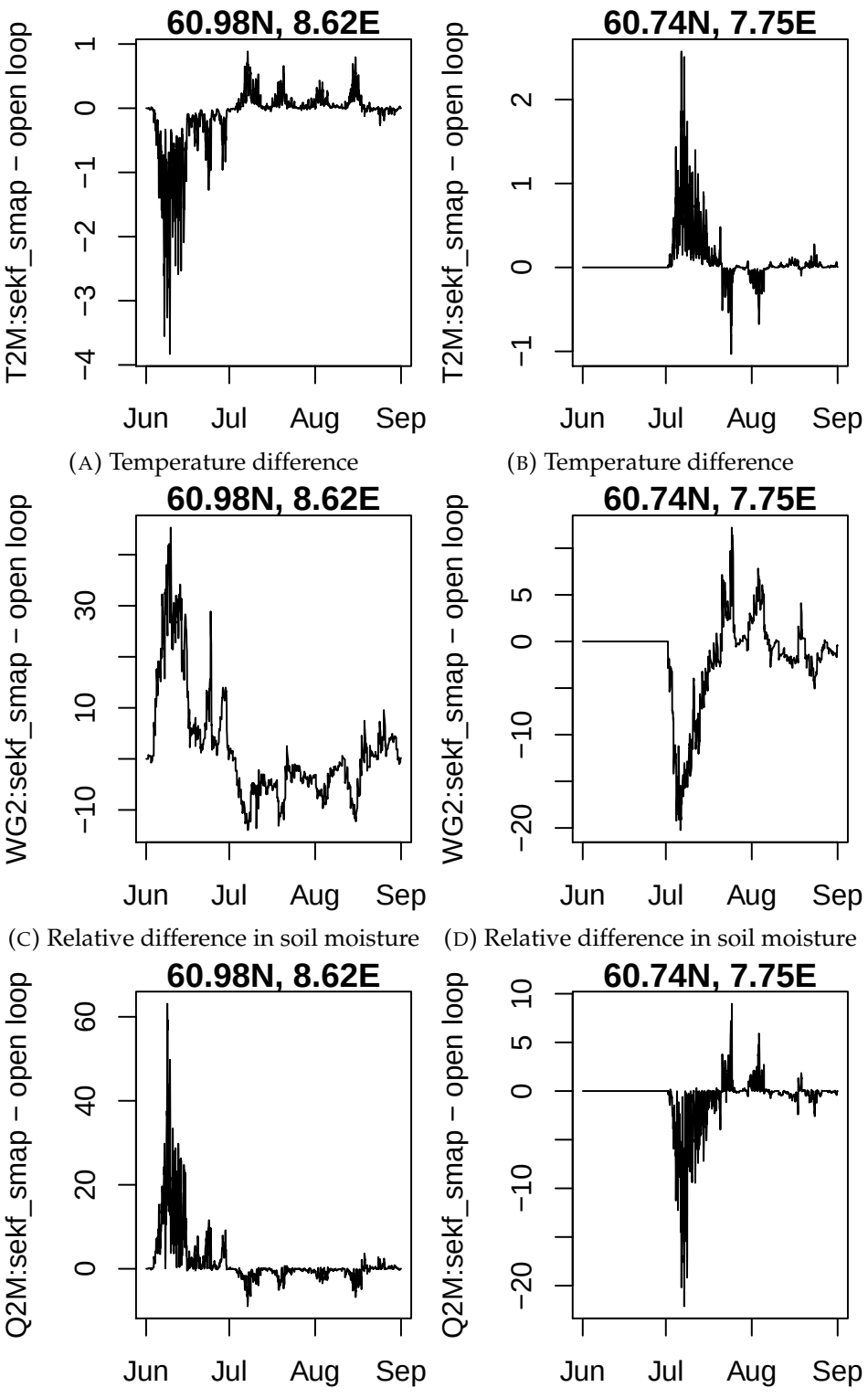


FIGURE 12.1: Upper panels show relative RMSE (between simulation and observations) vs. RMSD between SEKF_SMOS (left), SEKF_SMAP (right) and open loop. Distribution of RMSD between DA-runs and open loop is shown in lower right panel.

In the operational forecast produced at MetCoOp, screen level observations of temperature and humidity are used in the land surface DA. A horizontal interpolation is done to the station data with the Code d'Analyze Nécessaire à Action de Recherche Petite Echelle Grande Echelle (ARPEGE) pour ses Rejets et son Initiali-zation (CANARI) tool (Taillefer, 2002). These observations are then used to update soil moisture and soil temperature (Müller et al., 2017). OI is the DA method. Compared to EKF, OI does not calculate the linearized observation operator. Instead a constant relation between the observed variable and the control variable is used.

Even though the RMSE of 2 meter temperature is barely improved over the whole experiment period, the impact of soil moisture increments can be more pronounced in specific cases.

To identify cases where soil moisture plays an important role for screen level anomalies, time series are plotted in Fig. 12.2 for the two grid cells with the maximum, positive and negative, temperature difference between SEKF-runs and open loop. We observe that both plots have a maximum in the beginning of the simulation. In the warm case, Fig. 12.2b, no observation is available before July, hence SEKF_SMAP and open loop are equal to this point. Both points are located in the mountains and snow could be an explanation for the large differences. Remember that there is no snow present in the model, because of the short spin up and inadequate initialization of snow. Nevertheless, the maximum in the warm case, shown in Fig. 12.2b, coincides with a warm anomaly for the other location, about $54\,\mathrm{km}$ away, shown in Fig. 12.2a. The relative difference in soil moisture for the same locations (Fig. 12.2c and 12.2d) indicates a negative correlation between soil moisture and temperature, which is expected since a larger portion of the available energy is used to evaporate water and less to heat surface and air. The lower panel shows the relative difference in specific humidity with as much as 60% in the cold case (left).



(A) Temperature difference

(B) Temperature difference

(C) Relative difference in soil moisture

(D) Relative difference in soil moisture

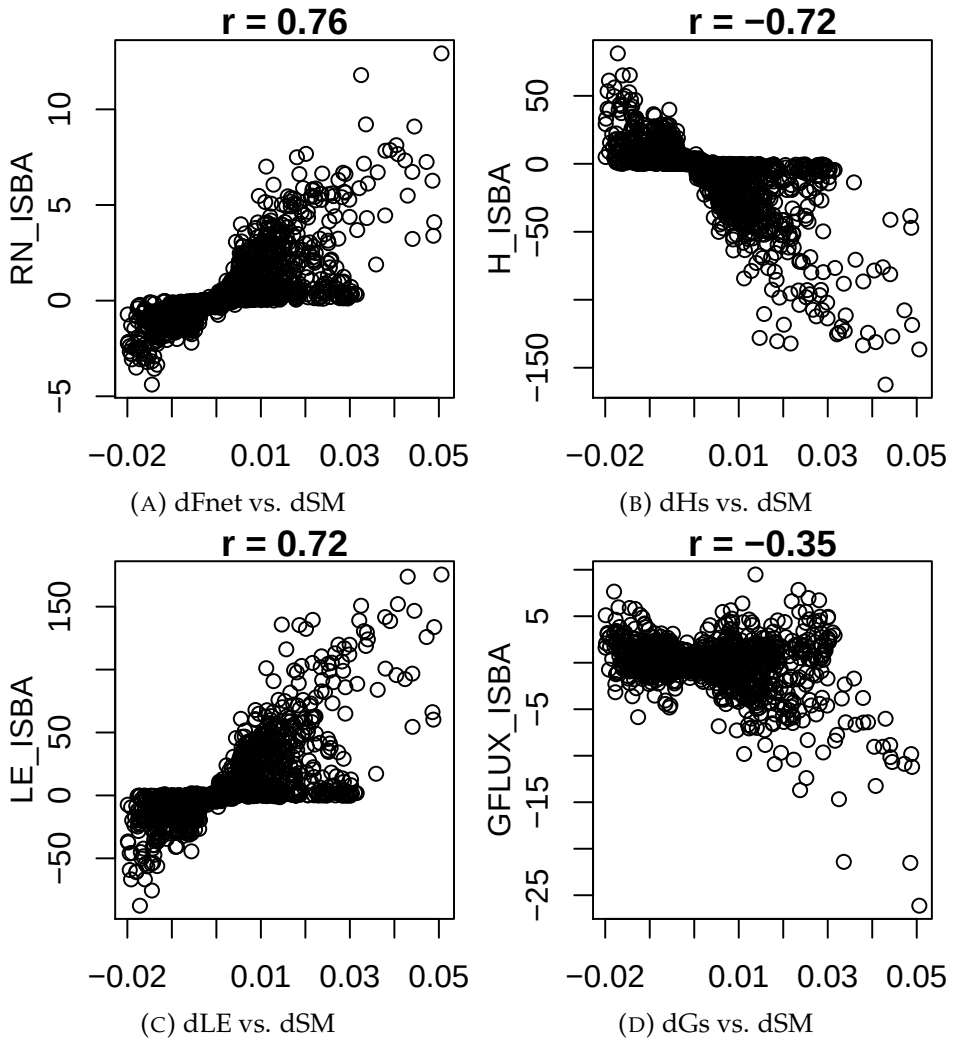
(E) Relative difference in specific humidity

(F) Relative difference in specific humidity

FIGURE 12.2: Difference (sekf_smos - open loop) between SEKF_SMAP and open loop at the grid cells with largest negative (left), and largest positive (right) difference in the domain. 2 meter temperature [K] (top) and soil moisture [%] at 1-4 cm (middle), and relative difference [%] in specific humidity (bottom)

12.2 Surface Fluxes

Surface fluxes calculated by the surface scheme, act as boundary conditions for the atmosphere. Changes in these, induced by soil moisture increments, are thus important to identify. Figure 12.3 shows changes induced in components of the energy balance, due to differences in soil moisture between sekf_smos and open loop simulations.



(A) dFnet vs. dSM (B) dHs vs. dSM

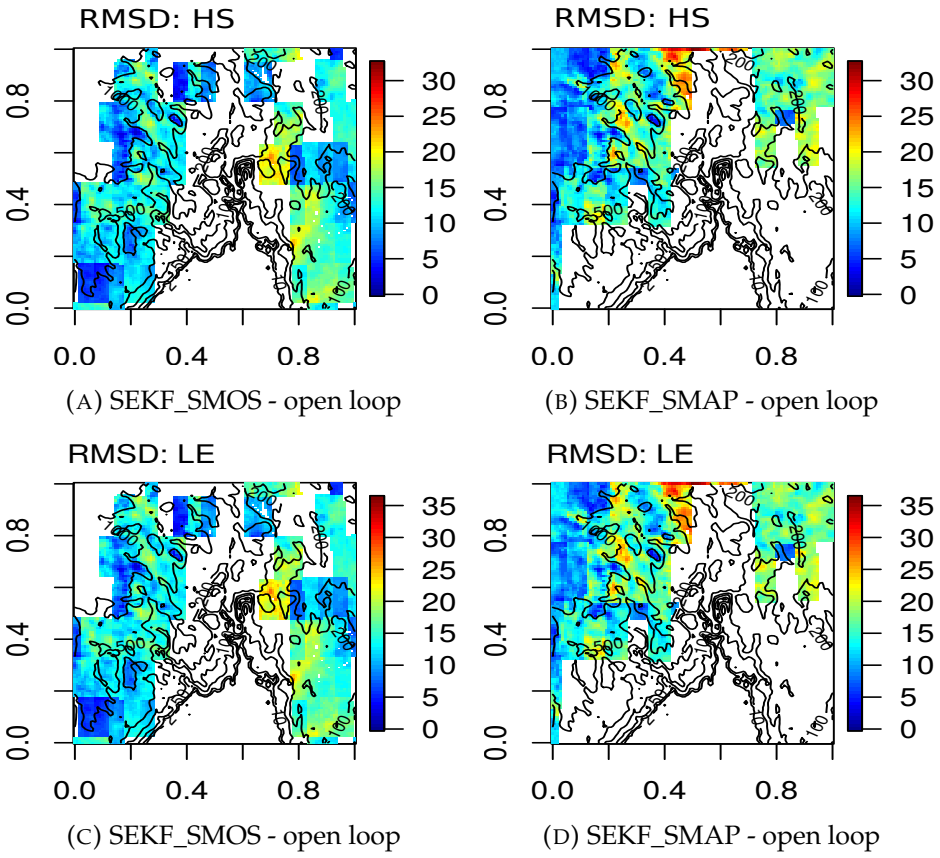
(C) dLE vs. dSM (D) dGs vs. dSM

FIGURE 12.3: Change in surface energy balance due to soil moisture changes: Fnet is net radiation, Hs sensible heat flux, LE latent heat flux, and Gs is ground heat flux.

Added water leads to less heating of the surface. This leads to increased net radiation as the outgoing component is reduced, and a decrease of sensible heat flux. Latent heat flux is then increased, as more water and energy is available for evaporation. Note that latent and sensible heat fluxes are negatively correlated, and an order of magnitude larger compared to the other components. The ground flux has no clear dependency, but shows a reduction for large positive soil moisture differences. As higher water contents reduce the heating, thus decrease ground flux, they also increase the thermal conductivity resulting in an increased ground flux. The two

opposing effects are indicated in the figure with both positive and negative correlations. The wet differences with a reduced ground flux coincides with dry soil moisture content (not shown), this implies that the effect of reduced heating of the surface dominate in dry conditions. Anyhow, most impact of soil moisture increments is found in latent and sensible heat fluxes.

RMSD for latent and sensible heat flux between DA and open loop-run are shown in Fig. 12.4. The RMSD range up to $30\,\mathrm{W\,m^{-2}}$, which is approximately 30% of the mean latent heat flux. As shown above, soil moisture increments have highest impact on latent and sensible heat flux. Furthermore, these vary opposite to each other and thus affecting the Bowen ratio. Decreased Bowen ratio, causes a colder, moister, and shallower planetary boundary layer, increasing cloud formation and reducing incoming short wave radiation (Jaeger and Seneviratne, 2011). The adjustment of the energy and mass fluxes will not affect the state of the atmosphere in offline simulations, but could change the characteristics of the planetary boundary layer and down stream effects in an coupled experiment.



(A) SEKF_SMOS - open loop

(B) SEKF_SMAP - open loop

(C) SEKF_SMOS - open loop

(D) SEKF_SMAP - open loop

FIGURE 12.4: Root mean square difference of sensible (top) and latent (bottom) heat flux [$\mathrm{W\,m^{-2}}$] between SEKF-run and open loop

Chapter 13

Summary and Conclusions

Soil moisture is a crucial component of the hydrological cycle, and the land surface energy balance. As it is important for the partitioning of available energy into sensible and latent heat flux, it influences temperature and humidity, locally, and potentially downstream. Accurate initialization and evolution of soil moisture are thus important in NWP on both short and long time scales (De Rosnay et al., 2014; Balsamo et al., 2014; Hurk et al., 2012; Weisheimer et al., 2011; Jaeger and Seneviratne, 2011; Koster et al., 2011; Hurk, Ettema, and Viterbo, 2008; Drusch and Viterbo, 2007; Koster et al., 2004; Koster, 2004; Douville et al., 2000).

The SURFEX land surface model is set up in an offline environment with atmospheric input from the operational NWP model AROME-MetCoOp. The simulations cover south-eastern Norway, during the summer 2016. The advanced multilayer diffusive scheme ISBA-DIF is used to provide detailed soil moisture and temperature profiles.

Daily level 3 soil moisture products from the SMOS and SMAP satellites are evaluated and interpolated onto the model grid. The observations are quality controlled, then normalized to model climate using a linear re scaling method.

To investigate the impact of soil moisture analyses, experiments include one control simulation without DA, open loop, and two simulations with DA using a SEKF and observations from the two satellites SMOS and SMAP. Tuning is performed to the observation errors, obtaining a consistent DA system.

In this study, SODA is optimized for DA of satellite soil moisture retrievals in conjunction with the ISBA-DIF scheme. An option to use soil moisture in the $1\text{-}4\,\mathrm{cm}$ layer as observed variable is added to represent satellite products valid at few centimeters. The control vector is extended to make analysis for soil moisture in the top seven layers, ranging to $80\,\mathrm{cm}$ below surface, to cover most of the root zone. The DA system is verified by investigating the linearized observation operator and analysis increments.

The impact of soil moisture increments is validated through comparison with the open loop, the operational model, and in situ measurements of screen level variables.

Good agreement is found between the open loop and the operational forecast in prognostic variables, superficial soil temperature and moisture, and diagnostic variables at screen level, temperature and humidity. Compared to in situ observations for eight stations, the operational model has slightly better performance with respect to air temperature and relative humidity.

The satellite retrievals, do not describe soil moisture content over the domain well. The data have poor spatial coverage, noisy signal, and unrealistically high and low values. Dry bias in the SMOS data product is a known problem at high latitudes, (Jeu et al., 2012), the same problem is found in this study. Screening procedures suggested in De Lannoy and Reichle (2016), flag most of the available retrievals and leave few left for the DA. In this work, retrievals of poor or questionable quality are, however, used. The only criterion is that the distribution of the normalized observations, is similar to that of the model.

Verification of the modifications in SODA is done by investigating the linearized observation operator and its values for the extended control vector. Decaying dependency between the observed layer and layers of increasing depth is found, similar to results in Albergel et al. (2017), with similar modifications to SODA. Further investigation of the linearized observation operator shows that observations are weighted more in the upper layers in dry conditions, and deeper layers in wet conditions. Since the satellite observations are noisy and uncorrelated with the model, observations are more likely to be drier when the model is wet, and wetter when model is dry. This results in over all added water to the top $10\,\mathrm{cm}$, and removed water in the layers below, during the experiment period.

Mean increments are compared to results in Draper, Mahfouf, and Walker (2011) and Albergel et al. (2017), and similar results are found as in the latter. In the former study, the force restore model with two layers is used instead of the diffusion version, and the sum of increments in the root zone was dependent on the observations used. Unless observations are correlated and have larger amplitude than the model, the mean increments will be distributed as in the results of this thesis. This is due to the linearized observation operator, directly mimicking the model physics of the ISBA-DIF.

Comparison of screen level variables for a selection of stations, shows small but consistent improvement in 2 meter temperature for the whole period. The amount of retrievals, surface parameters, and weather conditions presents large variations in analysis increments and impact on screen level variables. It is thus possible that more significant improvements were made in periods, and locations with no available in situ measurement. Since the simulations are decoupled from an atmospheric model, down stream effects of soil moisture increments are not described. However, soil moisture increments are able to induce large changes in surface fluxes of latent and sensible heat, primarily affecting the Bowen ratio.

Chapter 14

Outlook

To avoid influence from the surface model used in the satellite soil moisture retrieval, one can assimilate brightness temperature directly, instead of soil moisture (De Lannoy and Reichle, 2016). By using a RT-model in combination with a land model, a more consistent system should be obtained. More observations should be possible to retrieve, when the state of the surface is better known with respect to ice, snow and land cover types.

Using an EKF with evolving $\mathbf{B}$-matrix, will be very likely to produce a different distribution of increments through the soil profile, compared to the SEKF. Draper, Mahfouf, and Walker (2009) find that the EKF produces larger maximum and minimum values of increments, compared to the simplified analysis. Testing the EKF with an evolving background error covariance matrix, in conjunction with the ISBA-DIF, is a natural way of continuing the work in this thesis.

Compared to the EKF, ensemble based Kalman filters can account for errors in the forcing as well the model soil state. Choosing perturbations is a challenge and requires attention. However, ensemble forecasting systems, such as MetCoOp Ensemble Prediction System (MEPS), already produces ten members operationally, which can be used as forcing for offline experiments. The EnKF is in thus a candidate for future development of operational land DA.

Running SURFEX offline, constrain the full potential of soil moisture analyses, with respect to improving weather forecasts, as energy fluxes do not change the atmospheric state. Effects down stream of a soil moisture increment, must thus be investigated through a coupled surface atmosphere model system, which require much more computational power than what is used in this study. Investigation of soil moisture perturbations in an atmospheric model, can reveal the impact of increments of different magnitudes. Coupled experiments are encouraged in future studies.

Static surface and soil parameters, like clay and sand fractions, have significant effect on the soil moisture transport and dynamic range. The static fields used for soil texture in this study and in the operational surface scheme, lack the desired level of detail, and have room for improvement. Gathering better data could significantly improve the soil state in the model. A Short time Augmented EKF (STAEKF), used to add corrections to surface model parameters along with state increments has shown promising results (Carrassi et al., 2012). This method can in principle produce more realistic fields of static parameters, if the parameterizations are good, and further remove bias of e.g. soil moisture in the model.

List of Abbreviations

ALADIN Aire Limitée Adaptation Dynamique Développement International

AMSR-E Advanced Microwave Scanning Radiometer for EOS

AROME Applications of Research to Operations at MEsoscale

ARPEGE Action de Recherche Petite Echelle Grande Echelle

ASCAT Advanced SCATterometer

ATBD algorithm theoretical basis document

BLUE Best Linear Unbiased Estimate

CANARI Code d'Analyze Nécessaire à ARPEGE pour ses Rejets et son Initiali-zation

CATDS Centre Aval de Traitement des Données SMOS

CDF cumulative density function

CMEM Community Microwave Emission Modeling Platform

CNRM Centre National de Recherches Météorologiques

DA data assimilation

DIF DIFfusion

DQX data quality index

EASE-2 Equal Area Scalable Earth-2

ECMWF European Centre for Medium-Range Weather Forecasts

EnKF Ensemble Kalman Filter

EKF Extended Kalman Filter

FAO Food and Agricultural Organization (United nations)

FR force restore

GMAO Global Modeling and Assimilation Office

HARMONIE HIRLAM ALADIN Research on Mesoscale Operational NWP in Europe

HIRLAM HIgh Resolution Limited Area Model

HWSD Harmonized World Soil Database

ISBA Interaction Sol-Biosphère-Atmosphère

ISMN International Soil Moisture Network

JJA June, July, August

KF Kalman filter

LPRM Land Parameter Retrieval Model

MetCoOp Meteorological Co-operation on Operational NWP

MEPS MetCoOp Ensemble Prediction System

NSIDC National Snow and Ice Data Center

NWP numerical weather prediction

OI Optimal interpolation

PGD physiography data

RFI radio frequency interference

RMSD root mean square difference

RMSE root mean square error

RT radiative transfer

SEKF simplified extended Kalman filter

SMAP Soil Moisture Active Passive

SMOS Soil Moisture and Ocean Salinity

SODA SURFEX Offline Data Assimilation

STAEKF Short time Augmented EKF

SURFEX SURFace EXternalisée

List of Figures

List of Tables

Bibliography

Albergel, C. and S. Munier (2017). *Tehcnical news for the assimilation part SODA*. URL: http://www.umr-cnrm.fr/surfex/spip.php?article429 (visited on 11/22/2017).

Albergel, C. et al. (May 30, 2017). "Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX_v8.0: LDAS-Monde assessment over the Euro-Mediterranean area". In: *Geosci. Model Dev. Discuss.* 2017, pp. 1–53. ISSN: 1991-962X. DOI: 10.5194/gmd-2017-121. URL: http://www.geosci-model-dev-discuss.net/gmd-2017-121/ (visited on 06/01/2017).

Arya, S. Pal (2001). *Introduction to Micrometeorology*. Google-Books-ID: tRE0sBZ8u4YC. Academic Press. 448 pp. ISBN: 978-0-12-059354-5.

Balsamo, G. et al. (Dec. 1, 2007). "A Land Data Assimilation System for Soil Moisture and Temperature: An Information Content Study". In: *Journal of Hydrometeorology* 8.6, pp. 1225–1242. ISSN: 1525-755X. DOI: 10.1175/2007JHM819.1. URL: http://journals.ametsoc.org/doi/abs/10.1175/2007JHM819.1 (visited on 02/14/2017).

Balsamo, G. et al. (2014). "Representing the Earth surfaces in the Integrated Forecasting System: Recent advances and future challenges". In: *ECMWF Research Department Technical Memorandum* 729, p. 50. URL: http://www.academia.edu/download/42026627/Representing_the_Earth_surfaces_in_the_I20160204-30232-grcbx4.pdf (visited on 09/30/2016).

Balsamo, Gianpaolo et al. (June 1, 2009). "A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System". In: *Journal of Hydrometeorology* 10.3, pp. 623–643. ISSN: 1525-755X. DOI: 10.1175/2008JHM1068.1. URL: http://journals.ametsoc.org/doi/abs/10.1175/2008JHM1068.1 (visited on 11/16/2017).

Barbu, A. L. et al. (July 22, 2011). "Assimilation of Soil Wetness Index and Leaf Area Index into the ISBA-A-gs land surface model: grassland case study". In: *Biogeosciences* 8.7, pp. 1971–1986. ISSN: 1726-4189. DOI: 10.5194/bg-8-1971-2011. URL: http://www.biogeosciences.net/8/1971/2011/ (visited on 09/19/2017).

Barbu, A. L. et al. (Jan. 14, 2014). "Integrating ASCAT surface soil moisture and GEOV1 leaf area index into the SURFEX modelling platform: a land data assimilation application over France". In: *Hydrol. Earth Syst. Sci.* 18.1, pp. 173–192. ISSN: 1607-7938. DOI: 10.5194/hess-18-173-2014. URL: https://www.hydrol-earth-syst-sci.net/18/173/2014/ (visited on 09/20/2017).

Batjes, N. H. (Mar. 1, 1997). "A world dataset of derived soil properties by FAO–UNESCO soil unit for global modelling". In: *Soil Use and Management* 13.1, pp. 9–16. ISSN: 1475-2743. DOI: 10.1111/j.1475-2743.1997.tb00550.x. URL: http://onlinelibrary.wiley.com/

doi/10.1111/j.1475-2743.1997.tb00550.x/abstract (visited on 11/29/2017).

Bauer, Peter, Alan Thorpe, and Gilbert Brunet (Sept. 2, 2015). "The quiet revolution of numerical weather prediction". In: *Nature* 525.7567, nature14956. ISSN: 1476-4687. DOI: 10.1038/nature14956. URL: https://www.nature.com/articles/nature14956 (visited on 11/16/2017).

BBC (Apr. 30, 2015). "The birth of the weather forecast". In: *BBC News*. URL: http://www.bbc.com/news/magazine-32483678 (visited on 09/22/2017).

Behari, Jitendra (Mar. 11, 2006). *Microwave Dielectric Behaviour of Wet Soils*. Google-Books-ID: dsM4wfFaa7AC. Springer Science & Business Media. 177 pp. ISBN: 978-1-4020-3288-2.

Beljaars, Anton C. M. et al. (Mar. 1, 1996). "The Anomalous Rainfall over the United States during July 1993: Sensitivity to Land Surface Parameterization and Soil Moisture Anomalies". In: *Monthly Weather Review* 124.3, pp. 362–383. ISSN: 0027-0644. DOI: 10.1175/1520-0493(1996)124<0362:TAROTU>2.0.CO;2. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(1996)124%3C0362%3ATAROTU%3E2.0.CO%3B2 (visited on 09/30/2016).

Bengtsson, Lisa et al. (Feb. 9, 2017). "The HARMONIE–AROME Model Configuration in the ALADIN–HIRLAM NWP System". In: *Monthly Weather Review* 145.5, pp. 1919–1935. ISSN: 0027-0644. DOI: 10.1175/MWR-D-16-0417.1. URL: http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-16-0417.1 (visited on 04/27/2017).

Bjerknes, Vilhelm (1904). "Das Problem der Wettervorhersage, betrachtet vom Standpunkte der Mechanik und der Physik. (The problem of weather prediction, considered from the viewpoints of mechanics and physics)". In: *Meteorol. Z* 21, pp. 1–7.

Carrassi, Alberto et al. (Oct. 1, 2012). "Short time augmented extended Kalman filter for soil analysis: a feasibility study". In: *Atmospheric Science Letters* 13.4, pp. 268–274. ISSN: 1530-261X. DOI: 10.1002/asl.394. URL: http://onlinelibrary.wiley.com/doi/10.1002/asl.394/abstract (visited on 11/23/2017).

Chen, Fei and Roni Avissar (Dec. 1, 1994). "Impact of Land-Surface Moisture Variability on Local Shallow Convective Cumulus and Precipitation in Large-Scale Models". In: *Journal of Applied Meteorology* 33.12, pp. 1382–1401. ISSN: 0894-8763. DOI: 10.1175/1520-0450(1994)033<1382:IOLSMV>2.0.CO;2. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0450(1994)033%3C1382:IOLSMV%3E2.0.CO%3B2 (visited on 11/12/2017).

Cole, Steve and Alan Buis (Sept. 1, 2015). *NASA Soil Moisture Radar Ends Operations, Mission Science Continues*. SMAP. URL: https://smap.jpl.nasa.gov/news/1247/nasa-soil-moisture-radar-ends-operations-mission-science-continues/ (visited on 09/08/2017).

Crow, Wade T. et al. (June 1, 2012). "Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products". In: *Reviews of Geophysics* 50.2, RG2002. ISSN: 1944-9208. DOI: 10.1029/2011RG000372. URL: http://onlinelibrary.wiley.com/doi/10.1029/2011RG000372/abstract (visited on 09/25/2017).

De Lannoy, Gabriëlle J. M. and Rolf H. Reichle (Dec. 15, 2016). "Assimilation of SMOS brightness temperatures or soil moisture retrievals into a land surface model". In: *Hydrology and Earth System Sciences* 20.12, pp. 4895–4911. ISSN: 1607-7938. DOI: 10.5194/hess-20-4895-2016. URL: http://www.hydrol-earth-syst-sci.net/20/4895/2016/ (visited on 10/12/2017).

De Rosnay, Patricia de et al. (Mar. 11, 2009). "AMMA Land Surface Model Intercomparison Experiment coupled to the Community Microwave Emission Model: ALMIP-MEM". In: *Journal of Geophysical Research* 114 (D5). ISSN: 0148-0227. DOI: 10.1029/2008JD010724. URL: http://doi.wiley.com/10.1029/2008JD010724 (visited on 11/28/2017).

De Rosnay, Patricia de et al. (July 1, 2013). "A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF". In: *Quarterly Journal of the Royal Meteorological Society* 139.674, pp. 1199–1213. ISSN: 1477-870X. DOI: 10.1002/qj.2023. URL: http://onlinelibrary.wiley.com/doi/10.1002/qj.2023/abstract (visited on 09/20/2017).

De Rosnay, Patricia de et al. (May 1, 2014). "Initialisation of Land Surface Variables for Numerical Weather Prediction". In: *Surveys in Geophysics* 35.3, pp. 607–621. ISSN: 0169-3298, 1573-0956. DOI: 10.1007/s10712-012-9207-x. URL: http://link.springer.com/article/10.1007/s10712-012-9207-x (visited on 01/24/2017).

Desroziers, G. et al. (Oct. 1, 2005). "Diagnosis of observation, background and analysis-error statistics in observation space". In: *Quarterly Journal of the Royal Meteorological Society* 131.613, pp. 3385–3396. ISSN: 1477-870X. DOI: 10.1256/qj.05.108. URL: http://onlinelibrary.wiley.com/doi/10.1256/qj.05.108/abstract (visited on 03/29/2017).

Desroziers, Gerald and Serguei Ivanov (Apr. 1, 2001). "Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation". In: *Quarterly Journal of the Royal Meteorological Society* 127.574, pp. 1433–1452. ISSN: 1477-870X. DOI: 10.1002/qj.49712757417. URL: http://onlinelibrary.wiley.com/doi/10.1002/qj.49712757417/abstract (visited on 03/29/2017).

Dingman, S. Lawrence (2008). *Physical hydrology*. 2. ed., reissued. OCLC: 845365023. Long Grove, Ill: Waveland Press. 646 pp. ISBN: 978-1-57766-561-8.

Dorigo, W. A. et al. (May 30, 2011). "The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements". In: *Hydrology and Earth System Sciences* 15.5, pp. 1675–1698. ISSN: 1607-7938. DOI: 10.5194/hess-15-1675-2011. URL: http://www.hydrol-earth-syst-sci.net/15/1675/2011/ (visited on 11/29/2017).

Douville, Hervé et al. (June 1, 2000). "Evaluation of the Optimum Interpolation and Nudging Techniques for Soil Moisture Analysis Using FIFE Data". In: *Monthly Weather Review* 128.6, pp. 1733–1756. ISSN: 0027-0644. DOI: 10.1175/1520-0493(2000)128<1733:EOTOIA>2.0.CO;2. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(2000)128%3C1733%3AEOTOIA%3E2.0.CO%3B2 (visited on 09/30/2016).

Draper, C. S., J.-F. Mahfouf, and J. P. Walker (Oct. 27, 2009). "An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme". In:

Journal of Geophysical Research: Atmospheres 114 (D20), p. D20104. ISSN: 2156-2202. DOI: 10.1029/2008JD011650. URL: http://onlinelibrary.wiley.com/doi/10.1029/2008JD011650/abstract (visited on 09/20/2017).

Draper, C. S., J.-F. Mahfouf, and J. P. Walker (Jan. 27, 2011). "Root zone soil moisture from the assimilation of screen-level variables and remotely sensed soil moisture". In: *Journal of Geophysical Research: Atmospheres* 116 (D2), p. D02127. ISSN: 2156-2202. DOI: 10.1029/2010JD013829. URL: http://onlinelibrary.wiley.com/doi/10.1029/2010JD013829/abstract (visited on 02/14/2017).

Drusch, Matthias and Pedro Viterbo (Feb. 2007). "Assimilation of Screen-Level Variables in ECMWF's Integrated Forecast System: A Study on the Impact on the Forecast Quality and Analyzed Soil Moisture". In: *Monthly Weather Review* 135.2, pp. 300–314. ISSN: 0027-0644, 1520-0493. DOI: 10.1175/MWR3309.1. URL: http://journals.ametsoc.org/doi/abs/10.1175/MWR3309.1 (visited on 09/30/2016).

Entekhabi, Dara et al. (Oct. 1, 1999). "An Agenda for Land Surface Hydrology Research and a Call for the Second International Hydrological Decade". In: *Bulletin of the American Meteorological Society* 80.10, pp. 2043–2058. ISSN: 0003-0007. DOI: 10.1175/1520-0477(1999)080<2043:AAFLSH>2.0.CO;2. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0477%281999%29080%3C2043%3AAAFLSH%3E2.0.CO%3B2 (visited on 11/16/2017).

Evensen, Geir (Nov. 1, 2003). "The Ensemble Kalman Filter: theoretical formulation and practical implementation". In: *Ocean Dynamics* 53.4, pp. 343–367. ISSN: 1616-7341, 1616-7228. DOI: 10.1007/s10236-003-0036-9. URL: http://link.springer.com/10.1007/s10236-003-0036-9 (visited on 10/04/2017).

Fairbairn, D. et al. (Dec. 16, 2015). "Comparing the ensemble and extended Kalman filters for in situ soil moisture assimilation with contrasting conditions". In: *Hydrol. Earth Syst. Sci.* 19.12, pp. 4811–4830. ISSN: 1607-7938. DOI: 10.5194/hess-19-4811-2015. URL: https://www.hydrol-earth-syst-sci.net/19/4811/2015/ (visited on 09/20/2017).

Gelb, Arthur (1974). *Applied Optimal Estimation*. MIT Press. URL: https://mitpress.mit.edu/books/applied-optimal-estimation (visited on 11/22/2017).

Hartmann, Dennis L. (1994). *Global physical climatology*. International geophysics v. 56. San Diego: Academic Press. 411 pp. ISBN: 978-0-12-328530-0.

Hawkins, Ed and Rowan Sutton (Aug. 1, 2009). "The Potential to Narrow Uncertainty in Regional Climate Predictions". In: *Bulletin of the American Meteorological Society* 90.8, pp. 1095–1107. ISSN: 0003-0007. DOI: 10.1175/2009BAMS2607.1. URL: http://journals.ametsoc.org/doi/abs/10.1175/2009BAMS2607.1 (visited on 11/16/2017).

Hurk, Bart van den, Janneke Ettema, and Pedro Viterbo (Feb. 1, 2008). "Analysis of Soil Moisture Changes in Europe during a Single Growing Season in a New ECMWF Soil Moisture Assimilation System". In: *Journal of Hydrometeorology* 9.1, pp. 116–131. ISSN: 1525-755X. DOI: 10.1175/2007JHM848.1. URL: http://journals.ametsoc.org/doi/abs/10.1175/2007JHM848.1 (visited on 09/30/2016).

Hurk, Bart van den et al. (Jan. 1, 2012). "Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe". In: *Climate Dynamics* 38.1, pp. 349–362. ISSN: 0930-7575, 1432-0894. DOI: 10.1007/s00382-010-0956-2. URL: https://link.springer.com/article/10.1007/s00382-010-0956-2 (visited on 11/09/2017).

Jacquette, Elsa et al. (Oct. 22, 2010). "SMOS CATDS level 3 global products over land". In: Remote Sensing for Agriculture, Ecosystems, and Hydrology XII. Vol. 7824. International Society for Optics and Photonics, 78240K. DOI: 10.1117/12.865093. URL: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/7824/78240K/SMOS-CATDS-level-3-global-products-over-land/10.1117/12.865093.short (visited on 11/29/2017).

Jaeger, E. B. and S. I. Seneviratne (May 1, 2011). "Impact of soil moisture–atmosphere coupling on European climate extremes and trends in a regional climate model". In: *Climate Dynamics* 36.9, pp. 1919–1939. ISSN: 0930-7575, 1432-0894. DOI: 10.1007/s00382-010-0780-8. URL: http://link.springer.com/article/10.1007/s00382-010-0780-8 (visited on 01/15/2017).

Jeu, Richard de et al. (Jan. 1, 2012). *Evaluation of SMOS soil moisture with other existing satellite products*. Vol. 352. 25 pp.

John M. Wallace (2006). *Atmospheric science: an introductory survey*. In collab. with Peter V. Hobbs. 2nd ed. Vol. vol. 92. International geophysics series. Amsterdam: Elsevier. xvi+483. ISBN: 978-0-12-732951-2.

Jones, Hamlyn G. and R. A. Vaughan (2010). *Remote sensing of vegetation: principles, techniques, and applications*. OCLC: ocn500822994. Oxford ; New York: Oxford University Press. 353 pp. ISBN: 978-0-19-920779-4.

Kalman, Rudolph Emil (1960). "A new approach to linear filtering and prediction problems". In: *Journal of basic Engineering* 82.1, pp. 35–45.

Kerr, Y. H. et al. (May 2012). "The Smos Soil Moisture Retrieval Algorithm". In: *IEEE Transactions on Geoscience and Remote Sensing* 50.5, pp. 1384–1403. ISSN: 0196-2892. DOI: 10.1109/TGRS.2012.2184548.

Kerr, Yann H. (Jan. 25, 2007). "Soil moisture from space: Where are we?" In: *Hydrogeology Journal* 15.1, pp. 117–120. ISSN: 1431-2174, 1435-0157. DOI: 10.1007/s10040-006-0095-3. URL: http://link.springer.com/10.1007/s10040-006-0095-3 (visited on 09/26/2017).

Korres, W., T. G. Reichenau, and K. Schneider (Aug. 19, 2013). "Patterns and scaling properties of surface soil moisture in an agricultural landscape: An ecohydrological modeling study". In: *Journal of Hydrology* 498, pp. 89–102. ISSN: 0022-1694. DOI: 10.1016/j.jhydrol.2013.05.050. URL: http://www.sciencedirect.com/science/article/pii/S0022169413004368 (visited on 08/31/2017).

Koster, R. D. (Aug. 20, 2004). "Regions of Strong Coupling Between Soil Moisture and Precipitation". In: *Science* 305.5687, pp. 1138–1140. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1100217. URL: http://www.sciencemag.org/cgi/doi/10.1126/science.1100217 (visited on 11/18/2016).

Koster, R. D. et al. (Oct. 2011). "The Second Phase of the Global Land–Atmosphere Coupling Experiment: Soil Moisture Contributions to Subseasonal Forecast Skill". In: *Journal of Hydrometeorology* 12.5, pp. 805–822. ISSN: 1525-755X, 1525-7541. DOI: 10.1175/2011JHM1365.1. URL: http://

journals.ametsoc.org/doi/abs/10.1175/2011JHM1365.1 (visited on 11/18/2016).

Koster, Randal D. et al. (Dec. 2004). "Realistic Initialization of Land Surface States: Impacts on Subseasonal Forecast Skill". In: *Journal of Hydrometeorology* 5.6, pp. 1049–1063. ISSN: 1525-755X, 1525-7541. DOI: 10.1175/JHM-387.1. URL: http://journals.ametsoc.org/doi/abs/10.1175/JHM-387.1 (visited on 11/18/2016).

Lahoz, William A. and Gabriëlle J. M. De Lannoy (May 2014). "Closing the Gaps in Our Knowledge of the Hydrological Cycle over Land: Conceptual Problems". In: *Surveys in Geophysics* 35.3, pp. 623–660. ISSN: 0169-3298, 1573-0956. DOI: 10.1007/s10712-013-9221-7. URL: http://link.springer.com/10.1007/s10712-013-9221-7 (visited on 10/19/2016).

Lahoz, William Albert, Boris Khattatov, and Richard Ménard, eds. (2010). *Data assimilation: making sense of observations*. OCLC: ocn499067426. Heidleberg ; New York: Springer. 718 pp. ISBN: 978-3-540-74702-4 978-3-540-74703-1.

Le Moigne, P. (Aug. 3, 2012). "SURFEX SCIENTIFIC DOCUMENTATION". In: URL: http://www.cnrm-game-meteo.fr/surfex/IMG/pdf/surfex_scidoc_v2-2.pdf (visited on 09/29/2016).

Lorenz, Edward N. (Mar. 1963). "Deterministic Nonperiodic Flow". In: *Journal of the Atmospheric Sciences* 20.2, pp. 130–141. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0469%281963%29020%3C0130%3ADNF%3E2.0.CO%3B2 (visited on 03/20/2017).

Mahfouf, J.-F. et al. (Apr. 27, 2009). "A comparison of two off-line soil analysis schemes for assimilation of screen level observations". In: *Journal of Geophysical Research: Atmospheres* 114 (D8), p. D08105. ISSN: 2156-2202. DOI: 10.1029/2008JD011077. URL: http://onlinelibrary.wiley.com/doi/10.1029/2008JD011077/abstract (visited on 02/14/2017).

Miyakoda, K., G. D. Hembree, and R. F. Strickler (Apr. 1, 1979). "Cumulative Results of Extended Forecast Experiments II: Model Performance for Summer Cases". In: *Monthly Weather Review* 107.4, pp. 395–420. ISSN: 0027-0644. DOI: 10.1175/1520-0493(1979)107<0395:CROEFE>2.0.CO;2. URL: http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(1979)107%3C0395%3ACROEFE%3E2.0.CO%3B2 (visited on 11/07/2017).

Mueller, Brigitte and Sonia I. Seneviratne (July 31, 2012). "Hot days induced by precipitation deficits at the global scale". In: *Proceedings of the National Academy of Sciences* 109.31, pp. 12398–12403. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1204330109. URL: http://www.pnas.org/content/109/31/12398 (visited on 11/16/2017).

Müller, Malte et al. (Jan. 11, 2017). "AROME-MetCoOp: A Nordic Convective-Scale Operational Weather Prediction Model". In: *Weather and Forecasting* 32.2, pp. 609–627. ISSN: 0882-8156. DOI: 10.1175/WAF-D-16-0099.1. URL: http://journals.ametsoc.org/doi/abs/10.1175/WAF-D-16-0099.1 (visited on 04/27/2017).

ONeill, Peggy (2016). *SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 4*. DOI: 10.5067/OBBHQ5W22HME.

ONeill, Peggy et al. (Sept. 14, 2015). *Algorithm Theoretical Basis Document Level 2 & 3 Soil Moisture (Passive) Data Products*. URL: https://smap.jpl.nasa.gov/documents/.

Reichle, R. H. (2004). "Bias reduction in short records of satellite soil moisture". In: *Geophysical Research Letters* 31.19. ISSN: 0094-8276. DOI: 10.1029/2004GL020938. URL: http://doi.wiley.com/10.1029/2004GL020938 (visited on 10/30/2017).

Rosnay, P. de et al. (June 16, 2002). "Impact of a physically based soil water flow and soil-plant interaction representation for modeling large-scale land surface processes: PHYSICALLY BASED SOIL HYDROLOGY IN GCM". In: *Journal of Geophysical Research: Atmospheres* 107 (D11), ACL 3–1–ACL 3–19. ISSN: 01480227. DOI: 10.1029/2001JD000634. URL: http://doi.wiley.com/10.1029/2001JD000634 (visited on 11/16/2017).

Schalie, R. van der et al. (Mar. 1, 2016). "Global SMOS Soil Moisture Retrievals from The Land Parameter Retrieval Model". In: *International Journal of Applied Earth Observation and Geoinformation*. Advances in the Validation and Application of Remotely Sensed Soil Moisture - Part 1 45 (Part B), pp. 125–134. ISSN: 0303-2434. DOI: 10.1016/j.jag.2015.08.005. URL: http://www.sciencedirect.com/science/article/pii/S0303243415300179 (visited on 11/28/2017).

Schmugge, Thomas (Dec. 1, 1998). "Applications of passive microwave observations of surface soil moisture". In: *Journal of Hydrology* 212 (Supplement C), pp. 188–197. ISSN: 0022-1694. DOI: 10.1016/S0022-1694(98)00209-1. URL: http://www.sciencedirect.com/science/article/pii/S0022169498002091 (visited on 09/26/2017).

Seneviratne, Sonia I. et al. (May 2010). "Investigating soil moisture–climate interactions in a changing climate: A review". In: *Earth-Science Reviews* 99.3, pp. 125–161. ISSN: 0012-8252. DOI: 10.1016/j.earscirev.2010.02.004. URL: http://www.sciencedirect.com/science/article/pii/S0012825210000139 (visited on 03/21/2017).

Seneviratne, Sonia I. et al. (Oct. 16, 2013). "Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment: GLACE-CMIP5 EXPERIMENT". In: *Geophysical Research Letters* 40.19, pp. 5212–5217. ISSN: 00948276. DOI: 10.1002/grl.50956. URL: http://doi.wiley.com/10.1002/grl.50956 (visited on 09/30/2016).

Taillefer, F. (Nov. 30, 2002). *CANARI technical documentation (2002)*. http://www.umr-cnrm.fr/gmapdoc. URL: http://www.umr-cnrm.fr/gmapdoc/spip.php?article3 (visited on 11/24/2017).

Weisheimer, Antje et al. (Mar. 16, 2011). "On the predictability of the extreme summer 2003 over Europe". In: *Geophysical Research Letters* 38.5, p. L05704. ISSN: 1944-8007. DOI: 10.1029/2010GL046455. URL: http://onlinelibrary.wiley.com/doi/10.1029/2010GL046455/abstract (visited on 01/15/2017).

Wrona, Elizabeth et al. (May 16, 2017). "Validation of the Soil Moisture Active Passive (SMAP) satellite soil moisture retrieval in an Arctic tundra environment". In: *Geophysical Research Letters* 44.9, pp. 4152–4158. ISSN: 1944-8007. DOI: 10.1002/2017GL072946. URL: http://onlinelibrary.wiley.com/doi/10.1002/2017GL072946/full (visited on 11/17/2017).