Estimating soil moisture from space

A comparison between soil moisture estimates based on satellite retrievals and the E-HYPE model under drought scenarios in France year 2003 and 2011

April 26, 2017

2017-04-14
TVVR 17/5008
Erik Sönegård
Supervisor: Magnus Persson
Examiner: Rolf Larsson

Foreword

I would like to direct my sincerest thanks to my supervisors Professor Magnus Persson, working at Lund University at the Department of Building and Environmental Technology, and my co-supervisor Doctor David Gustafsson, research fellow at the KTH Royal Instituste of Technology and at the Swedish Meteorological and Hydrological Institute (SMHI), for the help, guidance and patience they have given me during the work with this thesis. Furthermore I would like to thank my examiner Associate Professor Rolf Larsson for taking on this task. Among many of the institutions from where data have been gathered to fulfill this thesis I would like to thank the SMHI, the European Space Agency and the European Comission for providing the data. Finally, many thanks to my student colleague Josefin Tollgren for her invaluable comments which made this thesis readable.

The work with this thesis has been performed from the spring of 2016 to the spring of 2017 at the Faculty of Engineering at Lund University.

Preface

The idea for this thesis was sprung out of curiosity, a wish to learn more about satellite remote sensing and an understanding of its limitation and possibilities. To put it into a context, soil moisture was the variable selected for analysis, and drought was the scenario under which this variable would be studied. I hope the reader will find this paper interesting and that it will work for them as an introduction to the subjects of agricultural drought, soil moisture estimation and satellite remote sensing.

Kind regards, Erik Sönegård

Abstract

Economic impacts of drought events can amass to billions of Euros within the EU. Ways to estimate the amount of water in the soil is thus very valuable to gather information which could be used for decision making. This thesis looks at different ways to estimate the amount of water in the soil, the so called soil moisture. The E-HYPE model from the Swedish Meteorological and Hydrological Institute (SMHI) is compared to different estimates based on satellite remote sensing technology. The thesis investigates different ways to make comparisons between these types of models and products during drought events in France year 2003 and 2011. The thesis demonstrates that both types of tools can be used to analyze a drought event and that there is correlation between the different types of data sets.

Abstract

Den ekonomiska effekten av torka kan uppgå till miljarder Euro inom EU. Sätt att uppskatta mängden vatten i marken är därför väldigt värdefull för att samla information som kan användas vid beslutstagande. Denna uppsats undersöker olika sätt man kan uppskatta mängden vatten i marken, den så kallade markfuktigheten. E-HYPE modellen från Sveriges Meterologiska och Hydrologiska Institut (SMHI) jämförs med olika uppskattningar som görs baserade på information från satelliter. Uppsatsen undersöker olika sätt man kan utföra en jämförelse mellan dessa typer av modeller och produkter under ett scenario av torka i Frankrike år 2003 och 2011. Uppsatsen visar att båda typerna av verktyg kan användas för att analysera en händelse av torka och att det finns korrelation mellan dessa olika typer av verktyg.

Abbreviations

SMHI	Swedish Meteorological and Hydrological Institute
ESA	The European Space Agency
ESA CCI	European Space Agency's Climate Change Initiative
ESA CCI SM	European Space Agency´s Climate Change Initiative´s Combined Soil Moisture product version 02.2
Copernicus SWI	Copernicus Soil Water Index version 3
ASCAT SSM	ASCAT Surface Soil Moisture Data
ECV	Earth Climate Variable
UNFCCC	United Nations Framework Convention on Climate Change

Contents

List of Figures

1 Introduction

According to the European Commission (2007) there is a need for enhancing the knowledge of droughts in Europe. This thesis will focus on agricultural land, selection of the time period and the location to be studied will be based upon the impact on crops due to a deficiency of soil moisture, a drought concept known as "agricultural drought". Many European states have moved from crisis management to drought risk management [Commission, 2007], it is therefore probable that tools for assessing agricultural drought events will be of increasing importance in the future. An increased amount of the area and population in the EU have been affected by drought [Commission, 2006]. Estimates of economic impacts of drought events during a 30 year period were as large as 85 billion Euros at EU level, not including social and environmental costs [Commission, 2006]. Estimated average costs of drought impacts on French farms were 110 Million Euros per year since 1989 and the estimated cost due to the drought event of year 2003 amounted to 670 Million Euros [Commission, 2006].

Several methods exist for estimating the amount of water in soil. The Swedish Meteorological and Hydrological Institute (SMHI) has developed a dynamic hydrological model for estimating the runoff in rivers, named HYPE (Hydrological Predictions of the Environment), this model includes variables keeping track on the amount of water in the soil column. Also, there have been a large interest in developing satellite remote sensing tools for estimating the soil moisture, this thesis will look at the soil moisture variables from the European Space Agency Climate Change Initiative Combined Soil Moisture Product and the Copernicus Soil Water Index.

The disposition of this thesis is as follows; A theory section is given to familiarize the reader with the concepts of agricultural drought, soil moisture and satellite remote sensing, followed by the purpose and scientific questions of this thesis. The methodology is then given, defining the time frame for and location of the project, then presenting the individual products used in this thesis followed by a description of how this data has been handled and analyzed. A result section then follows presenting the results from the performed analysis followed by a discussion of the different results. A concluding remark is lastly given to the thesis.

1.1 Background

Here the main theoretical components are presented which are of special importance for the thesis, the concept of agricultural drought, soil moisture and how one can estimate this from space.

1.1.1 Agricultural drought

Agricultural drought is a phenomena occurring when there is crop failure due to a period of declining soil moisture, as stated by the American Meteorological Society "Agricultural drought links meteorological drought characteristics to agricultural impacts (...)" [AMS, 2017]. One has to make a distinction between the terms "drought" and "agricultural drought" since drought events can differ in terms of intensity and duration. Agricultural drought is thus a complex concept since it not only depends on the lack of precipitation but also variables such as soil characteristics, time of drought occurrence and plant types.

1.1.2 Soil Moisture

Soil moisture has an important role in a range of different fields with different objectives, some examples being for making climatic predictions, managing water resources and performing hazard analysis [Cho et al., 2015]. Soil moisture is an important variable for many geophysical processes and has many hydrological implications. Soil moisture relates to the amount of water in the unsaturated zone of the soil column [Petropoulos et al., 2014] and is an Essential Climate Variable (ECV) [GCOS, 2010], being "(. . .) a physical, chemical, or biological variable or a group of linked variables that critically contributes to the characterization of Earth's climate" [Bojinski et al., 2014]. Soils in natural condition always contain some water because of a number of forces affecting the water with the ability to work against the force of gravity [Robinson and Ward, 2000]. Three main forces can be recognized (after Robinson and Ward 2000):

- Capillary forces, due to the surface tension at the interface between air and water in the soil
- Adsorption, due to electrostatic forces between the solid particles and the water
- Osmosis, due to solutes in the water creating an osmotic pressure

Some terms for describing the hydrological status of soils are wilting point and field capacity. The wilting point is defined as "the minimum water content of the soil at which the plants can extract water" [Robinson and Ward, 2000]. The field capacity is the amount of water in the soil after the movement due to gravity ceased [Robinson and Ward, 2000].

One of the difficulties when keeping track of the amount of water in the soil is its spatial and temporal variability [Behari, 2005] meaning that a point sample may be an inefficient way of describing the soil moisture quantities of a larger area. Alternative ways to estimate soil moisture in an area is with the use of hydrological models or by satellite measuring techniques, which will be the tools used in this thesis.

1.1.3 How can soil moisture be estimated from space?

Soil moisture can be estimated by satellites in a number of different ways [Petropoulos et al., 2014], however the main focus in this thesis will be satellites with instruments in the microwave spectrum. Two principal types of remote sensing exists (from [Chung et al., 2015]):

- scatterometers and radars, denoted as *active* instruments, using their own source of electromagnetic energy for measurement
- radiometers, denoted as *passive* instruments, measuring energy that is reflected or emitted from the earth surface

Both active and passive remote sensing in microwave (MW) frequencies make use of the large contrast between the dielectric properties of water and dry soil [Petropoulos et al., 2014] [Behari, 2005], which makes it possible to accurately perform measurement in the order of 5 cm thickness of a surface layer of soil [Behari, 2005]. Water has a large dielectric constant due to the molecules ability to align its dipole moments along an applied field [Behari, 2005] and a dielectric can store electrical energy when the dielectric is placed in an electrical field [Gardner et al., 2000]. Some techniques, making use of the electromagnetic properties of water for measuring soil moisture, take advantage of the water's effect on the bulk dielectric properties of the soil [Petropoulos et al., 2014], being that at certain frequencies the permittivity of pure water is around 80 at 25 C^o while it for most solids, which constitutes the soil, it is often less then 6 [Gardner et al., 2000].

The passive instruments make use of that all objects, with a temperature higher than 0 K, emit radiation in the MW frequencies [Petropoulos et al., 2014]. This radiation is usually expressed as "brightness temperature" which is converted to "physical temperature" by including the surface emissivity, for smooth surfaces approximated from the reflectivity of the surface which can be predicted through the Fresnel reflection equations using the incidence angle and the complex dielectric constant of soil [Petropoulos et al., 2014]. The active instruments measure the difference in power between the transmitted and received electromagnetic radiation [Petropoulos et al., 2014]. Behari (2005) mention four advantages of satellite remote sensing in the MW frequencies

1. "atmosphere is effectively transparent, can be penetrated through clouds, providing all weather clearance in the decimeter range of wavelengths"
2. "vegetation can be treated as semi-transparent, allowing the observation of underlying surfaces"
3. "microwave measurements are independent of the dielectric properties of the target, which for soil are a function of the amount of water present"
4. "measurement is also independent of solar illumination, permitting continuing day or night observations"

1.2 Purpose

The focus on this study will be on investigating soil moisture conditions during an agricultural drought event, making comparisons between the E-HYPE model and other soil moisture products. The selected agricultural event takes place in the French regions "Centre" and "Poitou-Charente" during the year 2011. The soil moisture products studied are soil moisture estimates from the E-HYPE hydrological model version 3 (denoted as the E-HYPE model), the European Space Agency´s Climate Change Initiative´s Combined Soil Moisture product version 02.2 (denoted as the ESA CCI SM product) and the Copernicus Soil Water Index product version 3 (denoted as the Copernicus SWI product).

1.2.1 Research questions

- What T factor gives the highest correlation, using Pearson correlation and RMSE, when comparing soil moisture variables from the E-HYPE model to the Copernicus SWI. How high is the correlation?
- How well do the soil moisture variable in the E-HYPE model correlate, using Pearson correlation, to the soil moisture variable from the ESA CCI product?
- What is the best possible correlation, using Pearson correlation, possible to achieve for the year 2011 between the soil moisture variable in the E-HYPE model and the ESA CCI SM product when using aggregated data?
- How do the the soil moisture estimates deviate from long term mean? How do the different products differ in terms of time of deviation and their relative degree.

2 Method

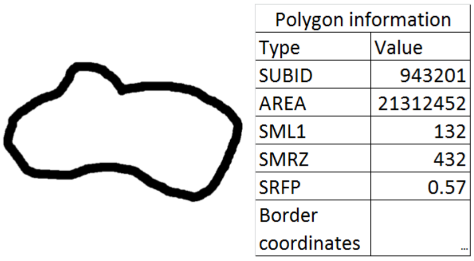
2.1 GIS for beginners

This section aims at giving readers new to handling geographical information system (GIS) information a simplyfied crashcourse in the basic methodology used when performing analysis in this thesis, if the reader have earlier experiences with GIS please read this section anyway since some terminology used in the thesis is introduced.

The E-HYPE data is related to a shape file which consists of a number of so called ”polygons”. As in figure 1, a polygon is a line delimiting an area which can have different types of information associated to it, such as coordinates, size, amount of soil moisture etc., the polygons in the E-HYPE model is referred to as ”segments”.

A raster object is a sheet of squares (usually one use the terminology ”cells” or ”pixels”) spread out in space, where each and every square contains a value (an example is given in figure 2). The Copernicus SWI and ESA CCI SM data are delivered as raster objects (covering the whole world).

The main task in this thesis is thus to attribute cells from the raster object to the polygon object, and there are many ways one can do this. In this thesis the methodology is to put the raster object ”over” the polygon object, and attributing all cells whose center resides within the borders of the polygon object as being a part of the polygon object (as in figure 3). It is important to understand that this will be an estimation, since some cells whose parts reside within the border wont be attributed to the polygon, and some cells being attributed to the polygon object will have parts of the cell lying outside of the polygon. What is then done is to compare these two values, the information of the polygon object with the averaged value of all cells attributed to the polygon from the raster object, and then try to draw conclusions based on this information.



Polygon information	
Type	Value
SUBID	943201
AREA	21312452
SML1	132
SMRZ	432
SRFP	0.57
Border coordinates	...

Figure 1: An imaginary example of a polygon object, a delimited area containing a set of parameters

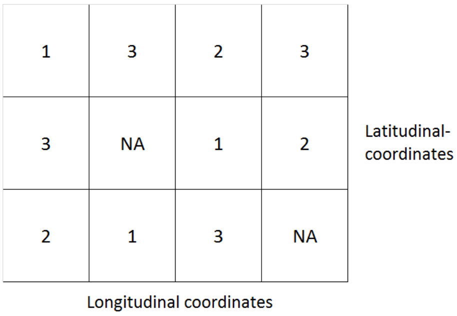


Figure 2: An imaginary example of a raster object where we have pixels distributed in space containing different values

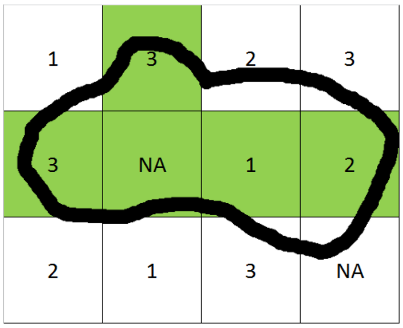


Figure 3: An example of a polygon object overlaying a raster object, the numbers give the value and mark the center of the raster cells, colored cells are attributed to the polygon

2.2 Definition of time frame

To avoid comparing the different products under frozen ground conditions, the time frame is defined using the Surface State Flag (SSF) provided together with the Copernicus SWI product. Data analysis will be made in between the longest period of the year where there are no cells within the selected E-HYPE segment which gives the ground the classification ”Frozen”. The results and chosen time period can be viewed in Appendix A. The data shows that for the segment located in the Centre region, analysis can be made using data for the whole year, however some caution must be taken during analysis of the months March and April since the data indicates ”Temporary water on the surface” which may affect the Copernicus SWI estimations. For the segment located in the Poitou-Charente region there were some short periods where the ground were in a ”Frozen” state, the data analysis will therefore be performed during day 6 up till and including day 347 (a normal year being composed of 365 days and 366 days on leap years).

2.3 Location of study sites

The studied area will be located in France, being a country occurring as one of the sites affected by drought year 2011 [Sepulcre-Canto et al., 2012]. The E-HYPE model over France is divided into 2531 segments, see figure 4. To find individual segments associated with an agricultural drought event, information from two sources (combined with information from academic papers and journals) will be used:

- Monitoring Agricultural ResourceS (MARS) bulletins provided by the European Commission´s science and knowledge service, the Joint Research Centre (JRC)
- Statistics regarding agricultural yields from the statistical office of the European Union, Eurostat

Figure 4: The 2531 E-HYPE segments representing France

To study an agricultural drought event, it is important to specify what type of agricultural crops that are being monitored when looking at drought scenarios. Based on statistics from Eurostat, France is the main producer of wheat and spelt in Europe, as is illustrated in figure 5. To further evaluate the impact of a certain drought event, regional statistics will be studied, instead of looking at statistics on a national level. Regional statistics from Eurostat are not provided for the production of wheat and spelt, but for a collection of crops under the label ”Cereals (excluding rice) for the production of grain (including seed)”, however, wheat and spelt make out the biggest portion of this category which is represented in figure 6, in where each category is presented as percentage of total (where all categories representing a production below 1% are presented under ”Other”). Thus, this statistics will be used as an indicator of the yearly yield of wheat and spelt.

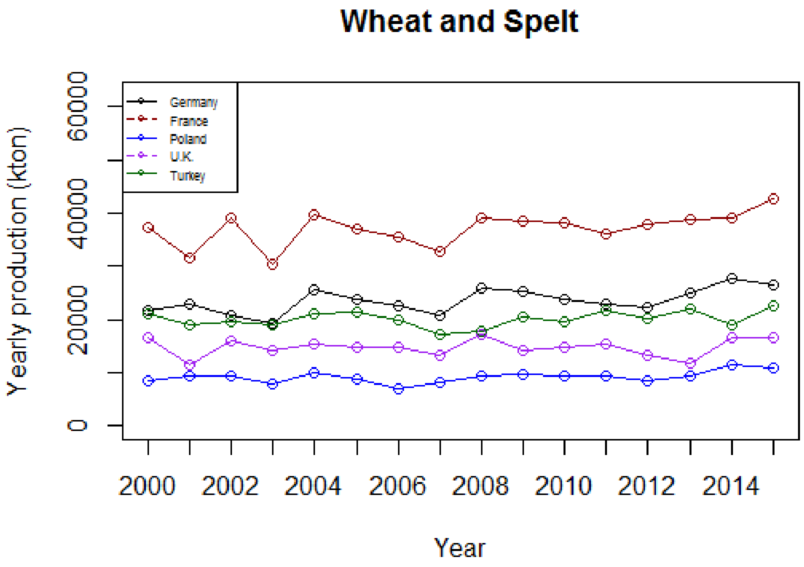


Figure 5: Countries in Europe that, according to statistics from Eurostat, have had a production of wheat and spelt exceeding 10 000 ktons on at least one occasion in between the years 2000-2015. Source: Eurostat

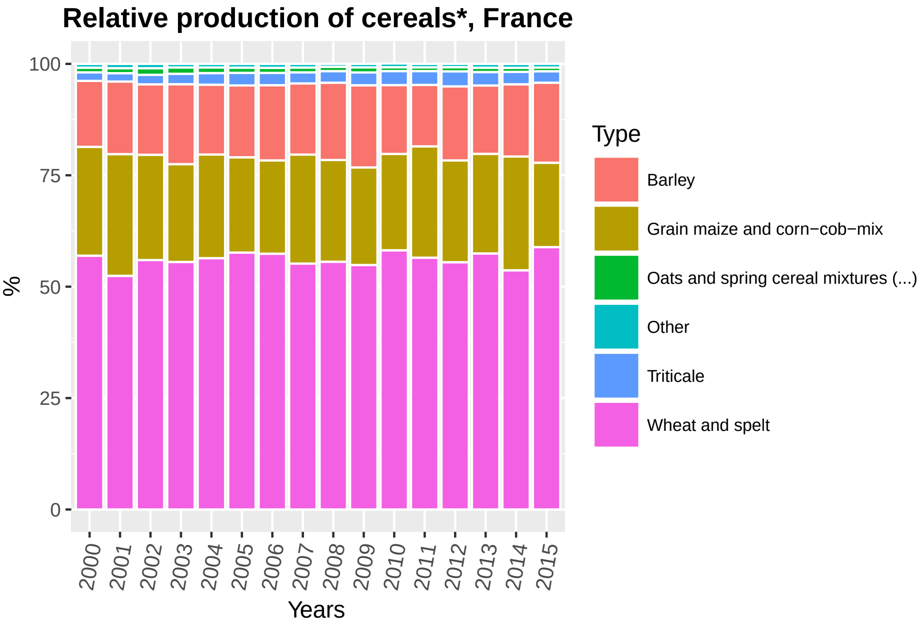


Figure 6: Relative production of cereals* in France. Source: Eurostat
**Cereals (excluding rice) for the production of grain (including seed)*

In the period March-May year 2011 precipitation was very scarce in Western Europe, the accumulated precipitation deficit for most of France from March to the end of May was in the range of 50-80% [MARS, 2011c]. During May-June regional percipitation deficits ranged from -57% in Poitou-Charente to -30% in Centre [MARS, 2011a]. The dry spell ended in the beginning of June, however precipitations in the regions Aquitaine, Midi Pyrenees and Centre Est remained below monthly averages [MARS, 2011c]. In the period June-July there were a lot of precipitation, with high averages of cumulated precipitation almost in the whole of France, with values in the regions Centre and Poitou-Charente between +40% and +50% [MARS, 2011b]. A synthesis of the 2010/2011 season in the MARS bulletin [MARS, 2011c] states that soft wheat and total wheat production in France was below average, affected by the dry spring. Yield statistics from Eurostat of the two regions Centre and Poitou-Charente, see figure 7 and Appendix B, shows that 2011 was the year with the lowest reported yield since year 2003 (the data for year 2012 have been omitted due to inconsistencies between the national and regional data, see Appendix B).

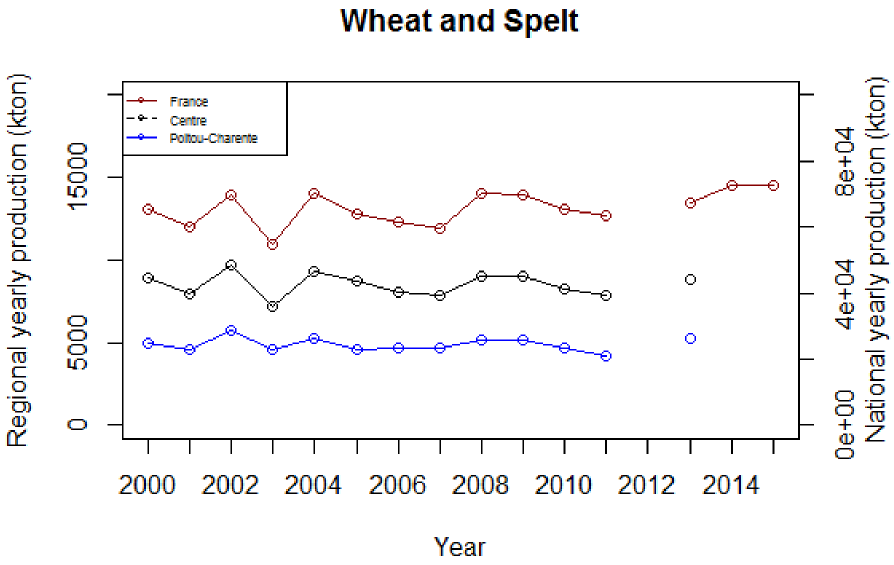


Figure 7: Total production of cereals* in France, Centre and Poitou-Charente. Source: Eurostat
**Cereals (excluding rice) for the production of grain (including seed).*

The regions Centre and Poitou-Charente are thus suitable for studying soil moisture conditions under the agricultural drought event of year 2011, the location of the selected regions can be seen in figure 8. The thesis will also look at soil moisture estimates for year 2002 and 2003, being the years with the highest and lowest reported yields respectively of ”Wheat and Spelt”, in between the years 2000-2013, in the Centre region. Year 2003 is also a year of great droughts over Europe which affected primary productivity [Ciais et al., 2005].

Figure 8: Location of the regions Centre and Poitou-Charente in France, source: Google Earth

2.4 Models and products used in this thesis

2.4.1 The ESA CCI Soil Moisture product

The objective of the ESA CCI is to make use of long term data gathered by ESA together with its Member States and contribute to the Earth Climate Variable (ECV) databases required by the United Nations Framework Convention on Climate Change (UNFCCC) [Downy, 2016]. The Soil Moisture CCI project is one of these programs. The ESA CCI Soil Moisture dataset (product version 02.2) consists of three surface soil moisture dataset: An "Active", "Passive" and "Combined" dataset [Chung et al., 2015]. The focus in this thesis will be the Combined dataset. The ESA CCI SM data files comprise a global soil moisture data set with a daily resolution [Chung et al., 2015]. The ESA CCI SM product is produced by merging data from the Passive and Active data sets and have a temporal resolution of 1 day, the data is provided in volumetric units [m^3 m^{-3}] [Chung et al., 2015]. Soil porosity data from the GLDAS-Noah data set is used for the conversion to volumetric soil moisture measurements. The product is delivered on a grid based on the World Geodetic System 1984 (WGS 84) reference system with a 0.25^ox0.25^o longitude-latitude global array of points. The platforms for producing the ESA CCI SM product include; Nimbus 7, DMSP, TRMM, ERS-1, ERS-2, METOP-A, AQUA, Coriolis, GCOM-W1 and sensors used to create this product [Chung et al., 2015] are presented in table 1

Table 1: Sensors used in the ESA CCI Soil Moisture product, their Passive/Active type, operating frequency and frequency band [GHz]

Sensor	Type	Operating frequency [GHz]
SMMR	Passive	6.6 & 10.69
SSM/I	Passive	19.3
TMI	Passive	10.7
AMI-WS	Active	5.255
ASCAT	Active	5.3
AMSR-E	Passive	6.9 & 10.7
WindSat	Passive	6.8 & 10.7
AMSR2	Passive	6.9 & 7.3 & 10.65

Comparisons of the ESA CCI SM product with modeled soil moisture at 0-10 cm from the Noah LSM version 3.3 and the Variable Infiltration Capacity (VIC) model, using z-scores, show an overall Spearman correlation of 0.76 and 0.85 respectively when comparing the wet season (March-September) during the time period 1992-2013 over East African cropping region [McNally et al., 2016].

Usage and acknowledgment of the ESA CCI Soil Moisture product

The ESA CCI Combined Soil Moisture Product (version 02.2) is developed as a part of the ESA Climate Change Initiative program and is freely available to users after simple registration form. References to the data sets are [Liu et al., 2012] [Liu et al., 2011] [Wagner et al., 2012]

2.4.2 The Copernicus Soil Water Index

The Copernicus Soil Water Index product version 3, denoted as Copernicus SWI, delivers a scaled soil moisture index (indexing from 0-200 representing relative soil moisture between 0-100%) [Paulik, 2016]. In this thesis the SWI was rescaled by dividing the SWI values by a factor of 200, thus giving an index ranging between 0-1. The product models soil moisture in the deeper soil layers using ASCAT Surface Soil Moisture Data (ASCAT SSM) [Paulik, 2016], ASCAT being a scatterometer instrument onboard the METOP satellites operating at 5.255 GHz [Bartalis et al., 2008]. The ASCAT SSM observations represents a thin top soil layer of 1-5 cm [Kidd et al., 2014], the product uses ASCAT SSM data at 25 km resolution and is provided on a regular grid over the globe with lines and columns spaced at 0.1^o on a WGS84 ellipsoid system [Paulik, 2016]. The ASCAT product, representing the soil moisture in the topsoil, is propagated down the soil column by using different T factors, through equation 1 (from [Paulik, 2016]). Available factors being 1, 5, 10, 15, 20, 40, 60 and 100, each factor representing how the soil moisture estimates in the past influence the current SWI [Paulik, 2016].

$$SWI(t_n) = \frac{\sum_i^n SSM(t_i) e^{-\frac{t_n - t_i}{T}}}{\sum_i^n e^{-\frac{t_n - t_i}{T}}} \; for \, t_i \leq t_n \tag{1}$$

Where $SWI(t_n)$ is Soil Water Index at time t_n, $SSM(t_i)$ is retrieved from scatterometer observations from the ASCAT instruments on board the MetOP satellites, t_n is the observation time of the current measurement, t_i are the obsetvation times of the previous measurements (all times are given in Julian days) and T is a factor determining how strongly SSM observations taken in the past influence the current SWI.

Cho et al. [2015] showed that the average correlation for ground based measurements at 20 cm and 30 cm, considering 8 sites in the Korean peninsula, was highest using a characteristic time length (referring to the factor T in equation 1) of 5 (0.75 and 0.64 respectively). For the deepest ground based measurement (i.e. 50 cm) a highest correlation of 0.53 was given using a characteristic time lengths of 5, 10 and 20 [Cho et al., 2015]. The reader should be aware that one of the main weaknesses of the Copernicus SWI, for its implementation in this thesis, is that evapotranspiration is disregarded when calculating the SWI [Paulik, 2016].

Usage and acknowledgment of the Copernicus Soil Water Index Product

In accordance with the data policy stipulated in Paulik 2016, the following citation is provided: *"The product was generated by the land service of Copernicus, the Earth Observation program of the European Commission. The research leading to the current version of the product has received funding from various European Commission Research and Technical Development programmes. The product is based on Metop/ASCAT surface soil moisture distributed by Eumetsat."*

2.4.3 The E-HYPE model

HYPE (Hydrological Predictions for the Environment) is a hydrological model developed by the Swedish Meteorological and Hydrological Institute

(SMHI) for small and large scale assessments of water resources and water quality [Lindstrom et al., 2010]. The E-HYPE model is a HYPE model developed for European basins. The model is used for simulating water flow and transport as well as nitrogen and phosphorus turnover [Lindstrom et al., 2010]. The model gives output variables for the amount of water in different soil layers which can be used for the estimation of the soil moisture content. The input parameters can be different in each model "segment", where each segment represents a delimited area where calculations are processed. The soil routine in the HYPE model divides the soil in each segment in up to three soil layers, each soil layer is divided into a solid part and a pore part, which contains water retention parameters for the wilting point, field capacity and effective porosity as can be seen in figure 9 and 10, from the SMHI website, see [SMHI]. These figures are available under CC Attribution-Share Alike 3.0 Unported, hyperlink available at [Creative Commons].

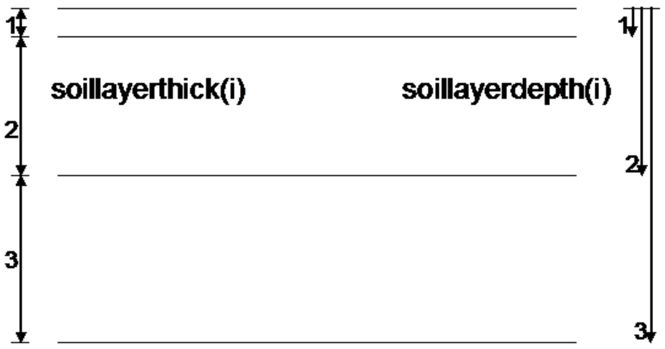


Figure 9: HYPE soil layers, soil layer 1, 2 and 3 is represented by the variable soillayerthick(i) on the left hand side

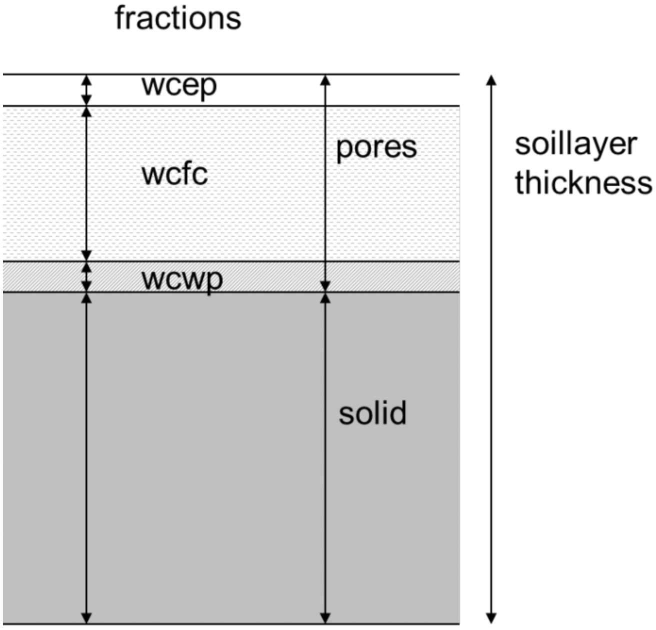


Figure 10: HYPE soil water retention parameters, wilting point (wcwp), field capacity (wcfc) and effective porosity (wcep)

The first two soil layers can be considered as the root depth, being the layers from which evapotranspiration is made possible [Lindstrom et al., 2010]. This thesis will focus on three soil moisture variables, SML1, representing the amount of water in layer 1 in millimeters, SMRZ, representing the amount of water in soil layer 1 and 2 combined in millimeters and SRFP, representing the amount of water in soil layer 1 and 2 combined as fraction of pore volume, these are further described in section 2.4.4.

2.4.4 Summary of studied product outputs

The outputs studied in this thesis, together with the product and model name and description, is given in table 2 and 3

Table 2: Model outputs studied in this thesis with descriptions

Model name	Model outputs	Description [units]
E-HYPE model version 3	*SML1*	Soil moisture, upper soil layer (not including standing water) [mm]
	SMRZ	Soil moisture root zone (upper two soil layers) (not including standing water) [mm]
	SRFP	Soil moisture root zone (upper two soil layers) (not including standing water) as fraction of pore volume

Table 3: Product outputs studied in this thesis with descriptions

Product name	Product outputs	Description [units]
ESA CCI Combined Soil Moisture Product (version 02.2)	*The ESA CCI SM product output*	$[m^3 m^{-3}]$
Copernicus Soil Water Index (SWI) product version 3	SWI_{T1}	The Copernicus SWI when using a factor T=1 $[m^3 m^{-3}]$
	SWI_{T5}	The Copernicus SWI when using a factor T=5 $[m^3 m^{-3}]$
	SWI_{T10}	The Copernicus SWI when using a factor T=10 $[m^3 m^{-3}]$
	SWI_{T15}	The Copernicus SWI when using a factor T=15 $[m^3 m^{-3}]$
	SWI_{T20}	The Copernicus SWI when using a factor T=20 $[m^3 m^{-3}]$
	SWI_{T40}	The Copernicus SWI when using a factor T=40 $[m^3 m^{-3}]$
	SWI_{T60}	The Copernicus SWI when using a factor T=60 $[m^3 m^{-3}]$
	SWI_{T100}	The Copernicus SWI when using a factor T=100 $[m^3 m^{-3}]$

2.5 Converting HDF5 data to Geotiff

The Copernicus SWI data is delivered in a HDF5 format. To analyze this data, the data is converted to Geotiff using the *gdal_translate* command, according to the instructions from the Copernicus Global Land Service website, see reference [Copernicus], in the OsGeow4 program, which is a project under the Open Source Geospatial Foundation (OSGeo), freely available at the website, see reference [OSGeo]

2.6 GlobeCover 2009

To determine weather the studied segment in the E-HYPE model represents an area which is made up of agricultural land, the GlobeCover 2009 product is used. The GlobeCover 2009 product delivers a global land cover map using the Plate-Carée projection and the WGS 84 ellipsoid, with a pixel resolution of $1/360^o$ [Arino et al., 2011]. Twenty-two land cover classes are defined with the United Nations Land Cover Classification System [Arino et al., 2011]. The land cover classifications are defined using the Meris instrument onboard the ENVISAT satellite, combined with other external data sets when handling gaps, flooded forests and water bodies [Arino et al., 2011]. When performing validation of the product using 2190 selected, globally distributed, points the overall accuracy reached 67.5 % [Arino et al., 2011].

Usage and acknowledgment of the Globcover 2009 product

Credit for the Globcover 2009 product goes to ESA and Université catholique de Louvain, Copyright notice: ©ESA 2010 and UCLouvain. The product is available at ESA DUE GlobCover website, see reference [Globecover].

2.7 Attributing spatial information to a polygon object

The segments in the E-HYPE model is in the form of polygons differing in spatial extent and having irregular shapes. A E-HYPE segment is regarded to be associated to a geographical region if its centroid lies within the region of interest, using Google Earth as a reference to define the regions extent. A segment within the region of interest is selected for further investigation with the condition that ≥90% of the associated land cover information attained by the GlobeCover 2009 product indicates that the segment is composed of "Rainfed croplands".

The data contained in the raster cells of the Copernicus SWI and ESA CCI SM products are attributed to the E-HYPE polygons using the *extract* method delivered in the *raster* package [Hijmans, 2015] in the R-software, the extract method returns the values of the cells in a raster object whose center is covered by the studied polygon. Of all the segments which fulfill the criteria, one is randomly selected for further analysis.

Using the above stated criterias, two segments in the E-HYPE model is selected for further analysis, presented in table 4 and 5. The location of the two segments in France can be viewed in figure 11

Table 4: Information regarding the selected E-HYPE segment in the Centre region of France

Centre Region	
E-HYPE SUBID	9564959
Number of Globcover cells	2466
Degree of Rainfed cropland [%]	91.9

Table 5: Information regarding the selected E-HYPE segment in the Poitou-Charente region of France

Poitou-Charente Region	
E-HYPE SUBID	9525076
Number of Globcover cells	2682
Degree of Rainfed cropland [%]	91.8

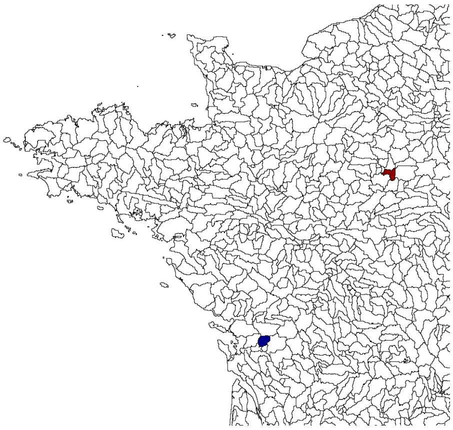


Figure 11: Location of the two selected segments of the E-HYPE model over France, segment 9564959 (red color, located in the Centre region) and 9525076 (blue color, located in the Poitou-Charente region)

2.8 Statistical tools used for analysis

The data in the raster cells from the ESA CCI SM and Copernicus SWI products are associated and averaged over the selected region using the arithmetic mean, see equation 2

$$\frac{1}{n}\sum_{i=1}^{n} x_i \tag{2}$$

Where x refer to the soil moisture estimate for pixel i in a sample containing n values. The averaged data from the ESA CCI SM and Copernicus SWI products are compared to the soil moisture variables in the E-HYPE model.

Correlation analysis is performed between the different data sets using the Pearson correlation coefficient, presented in equation 3 and, when applicable, the Root Mean Square Error (RMSE) method, presented in equation 4.

$$r = \frac{\sum_{k=1}^{i}(x_k - \bar{x})(y_k - \bar{y})}{\sqrt{\sum_{k=1}^{i}(x_k - \bar{x})^2}\sqrt{\sum_{k=1}^{i}(y_k - \bar{y})^2}} \tag{3}$$

Where x_i is a sample from the first data set $x_1, ..., x_n$ containing n values, y is a sample from the second data set $y_1, ..., y_n$ containing n values, $\bar{x}$ and $\bar{y}$ is the averages for the first and second data set calculated analogous to equation 2.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(x_{1,t} - x_{2,t})^2}{n}} \tag{4}$$

Where t is the time for the observation, n is the total number of observations, $x_{1,t}$ is the output from product "1" at time t and $x_{2,t}$ is the output from product "2" at time t.

To further try to improve the correlation between the ESA CCI SM product and the SML1 variable, the ESA CCI SM estimates will be averaged over a certain number of daily events before comparison to a single SML1 estimate, this since the SML1 variable represents a soil layer of deeper depths, thus trying to compensate for the large impact a smaller rainfall event would have on a thinner soil layer compared to the small impact it would have on a deeper soil layer. A new variable, q, is introduced, representing the number of days which are averaged. The variables x and y in equation 3 will be a new set of variables where the x data set will be composed of SML1 estimates for days given by a q day interval over the year, and the y variable data set will be composed of summed up ESA CCI SM variables according to equation 2. So if earlier the x (representing E-HYPE) and y (representing ESA CCI SM) data sets were composed of daily soil moisture estimates during year 2011 according to $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ where n is the total number of estimates, then the new x and y variables would be:

$$x = (x_{1*q)}, x_{2*q}, ..., x_n) \tag{5}$$

and

$$y = (\frac{1}{q}\sum_{t=(1)}^{q*1} y_t, \frac{1}{q}\sum_{t=(q*1+1)}^{q*2} y_t, ..., \frac{1}{q}\sum_{t=(q*(n-1)+1)}^{q*n} y_t) \tag{6}$$

Another way to compare soil moisture estimates from different soil layer depths are to use the z-score, as done by [McNally et al., 2016]. The way to compute the z-score is presented in equation 7 (from [Sepulcre-Canto et al., 2012]). The daily E-HYPE and ESA CCI SM soil moisture estimates are assumed to be normally distributed over time.

$$z = \frac{X_t - \overline{X}}{\sigma} \tag{7}$$

Where X_t is the daily soil moisture estimate for a certain day and year, $\overline{X}$ is the long term mean for that day over the studied number of years and σ the standard deviation for the date used to calculate the mean. When calculating long term means and standard deviations for each of the 365 days of the year, data from year 2000 to year 2011 will be used. Leap years will be treated as non leap years, thus the 366th day will be neglected when calculations are performed.

2.9 Softwares used in the thesis

2.9.1 The R software

For performing calculations and analysis of geographical information the R software [R Core Team, 2015] is used. R is a language and environment for statistical computing and graphics and is available as Free Software under the terms of the Free Software Foundation's GNU General Public License in source code form.

To extract the Copernicus SWI data using the R software, the rhd5 package is used [Fischer and Pau].

R-studio As compilator for the R software R-studio is used which is an integrated development environment (IDE) for R working under the AGPL v3 license available in an open source edition.

3 Results

The soil moisture estimates from the E-HYPE model, the ESA CCI SM and the Copernicus SWI product can be viewed in Appendix C

3.1 Finding the factor T giving the highest correlation between studied Copernicus SWI and E-HYPE outputs

The complete result from the comparison between the E-HYPE variables and the Copernicus SWI using different T factors can be viewed in Appendix D, however the T factors delivering the highest correlation and lowest root mean square error when compared to the E-HYPE variables SML1, SMRZ and SRFP are presented in table 6 and 7 which are the results from comparisons made in figure 12 and 13.

Table 6: Correlation and RMSE between E-HYPE SML1, SMRZ, SRFP and SWI_{T1}, SWI_{T5}, SWI_{T10}, SWI_{T20}, SWI_{T40}, SWI_{T60}, SWI_{T100}, for segment 9564959 (located in the Centre region), the best results are marked in bold

E-HYPE variable	T factor	Pearson correlation	RMSE
SML1	5	**0.856**	NA
SMRZ	40	**0.903**	NA
SRFP	40	**0.899**	0.102
SRFP	100	0.799	**0.075**

Table 7: Correlation and RMSE between E-HYPE SML1, SMRZ, SRFP and SWI_{T1}, SWI_{T5}, SWI_{T10}, SWI_{T20}, SWI_{T40}, SWI_{T60}, SWI_{T100}, for segment 9525076 (located in the Poitou-Charente region), the best results are marked in bold

E-HYPE variable	T factor	Pearson correlation	RMSE
SML1	1	**0.840**	NA
SMRZ	20	**0.911**	NA
SRFP	20	**0.911**	0.136
SRFP	100	0.738	**0.082**

To extract additional information regarding the dynamics of the data from the time series containing the E-HYPE SRFP and the Copernicus SWI data, these data sets are sorted from minimum to maximum value and then again compared, the results are viewed in figure 14

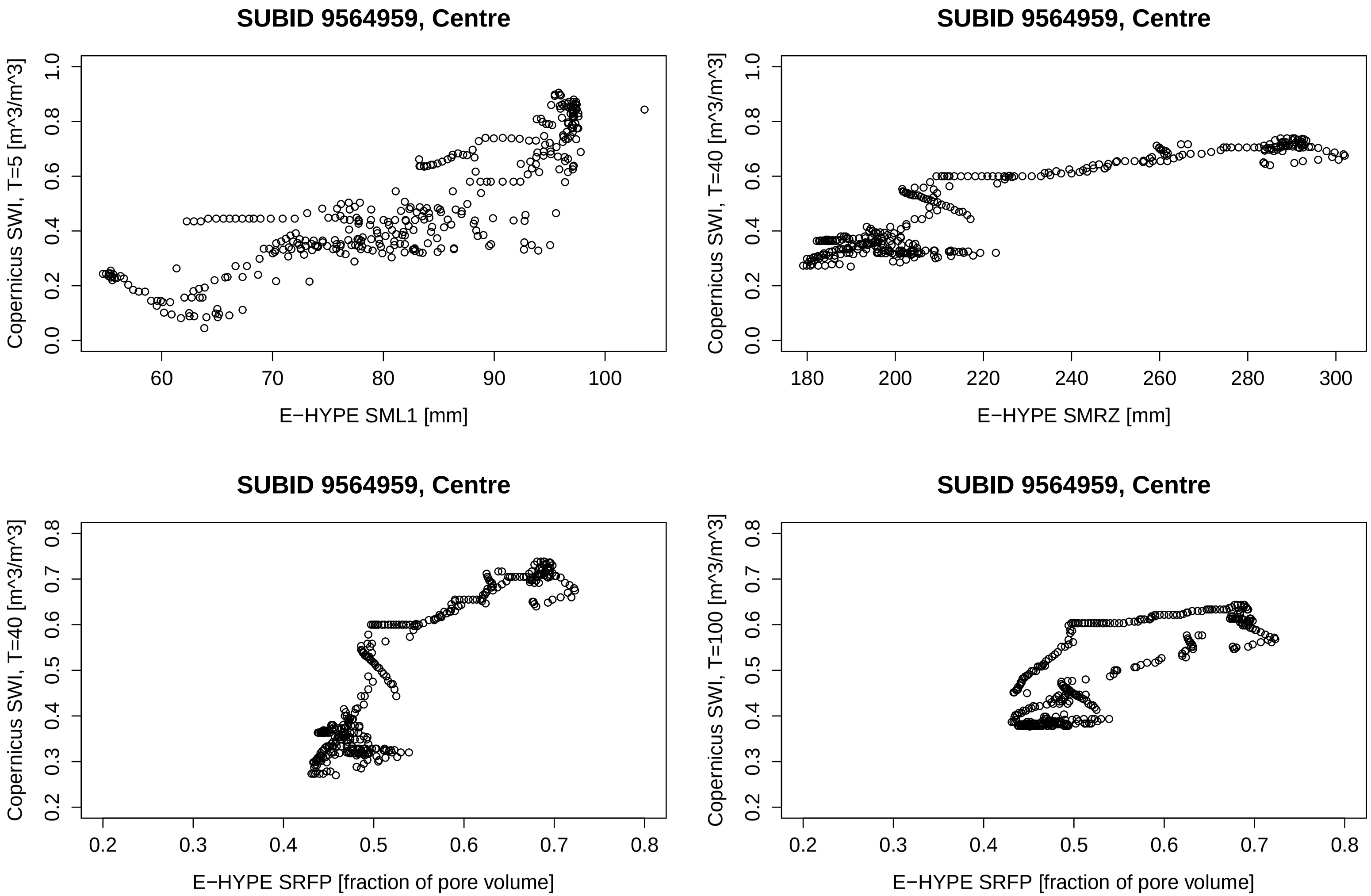


Figure 12: The E-HYPE SML1, SMRZ, SRFP variables plotted against the Copernicus SWI with T factors equal to 5, 40 and 100, for the 365 days during the year 2011, in segment 9564959 (located in the Centre region)

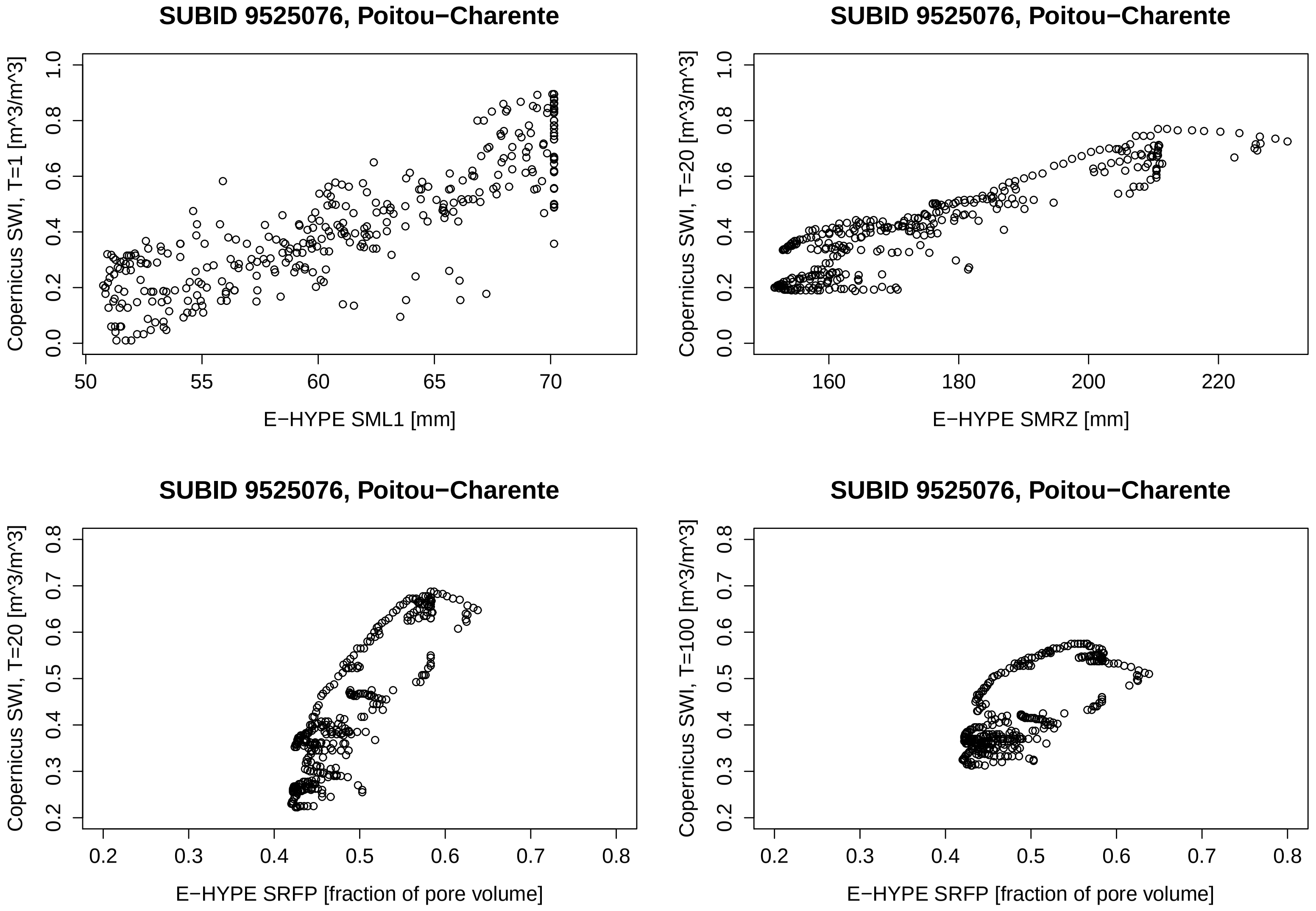


Figure 13: The E-HYPE SML1, SMRZ, SRFP variables plotted against the Copernicus SWI with T factors equal to 1, 20 and 100, for the 342 studied days during the year 2011, in segment 9525076 (located in the Poitou-Charente region)

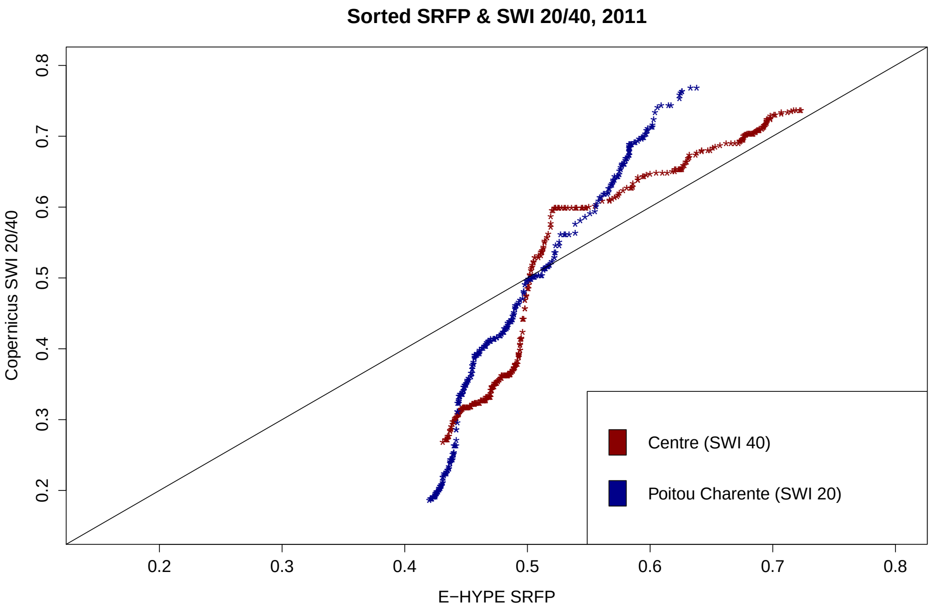


Figure 14: Sorted E-HYPE SRFP and Copernicus 20 and 40 data for the two segments located in the Centre and Poitou Charente region

3.2 Comparison between the E-HYPE SML1 variable and the ESA CCI SM product

The results when performing Pearson correlation between the E-HYPE SML1 variable and the ESA CCI SM data for the two studied segments are given in table 8. The E-HYPE SML1 variable plotted against the ESA CCI SM data for the two segments can be seen in figure 15.

Table 8: Correlation between E-HYPE SML1 and ESA CCI, for segment 9564959 (located in the Centre region) and 9525076 (located in the Poitou-Charente region)

Segment/Region	E-HYPE variable	Pearson correlation
Centre	SML1	0.795
Poitou-Charente	SML1	0.760

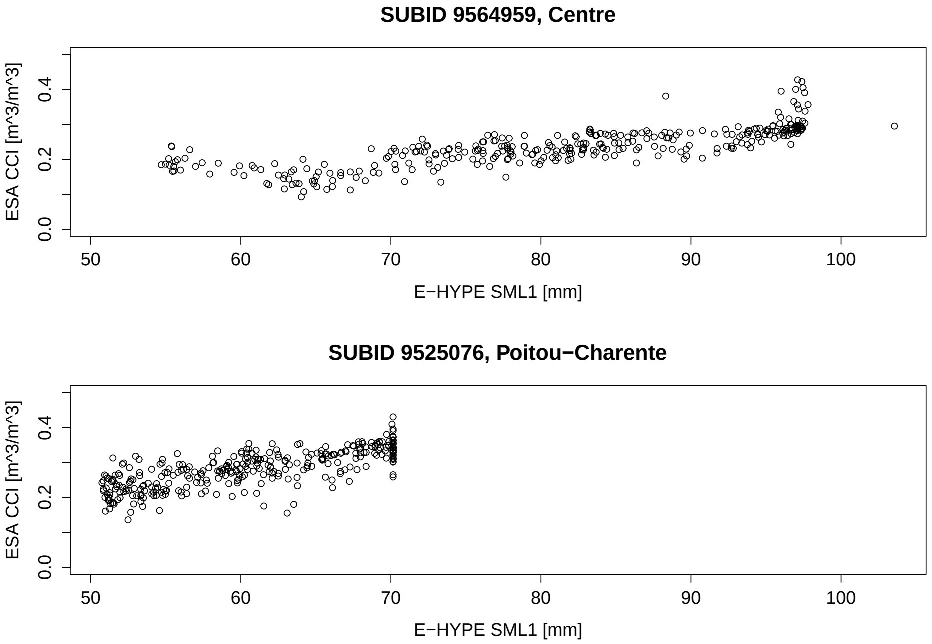


Figure 15: The E-HYPE SML1, SMRZ, SRFP variables plotted against the ESA CCI SM data, for the 365 and 342 studied days during the year 2011, for segment 9525076 (located in the Centre region) and segment 9525076 (located in the Poitou-Charente region)

3.3 Comparison between the E-HYPE SML1 variable and aggregated ESA CCI SM data

The correlation with aggregated ESA CCI SM estimates with single E-HYPE SML1 values are given in table 9.

Table 9: Correlation between E-HYPE SML1 and averaged ESA CCI SM estimates for segment 9564959 (located in the Centre region) and 9525076 (located in the Poitou-Charente region), the best results are marked in bold

q-values	Pearson correlation, Centre (no. valid obs.)	Pearson correlation, Poitou-Charente (no. valid obs.)
1	0.795 (342)	0.760 (328)
2	0.832 (164)	0.836 (157)
3	0.835 (103)	0.856 (100)
4	0.865 (74)	0.836 (74)
5	0.839 (55)	0.826 (56)
6	0.845 (44)	**0.884 (45)**
7	0.814 (35)	0.827 (37)
8	0.823 (28)	0.824 (33)
9	0.807 (22)	0.840 (27)
10	**0.873 (20)**	0.791 (23)

3.4 Comparison of Z-scores between the E-HYPE SML1 variable and the ESA CCI SM product

The E-HYPE SML1 and ESA CCI SM z-scores for the Centre and Poitou-Charente region is presented in figure 16 and 17, the correlation between them can be seen in table 10. A comparison between the E-HYPE SMRZ and ESA CCI SM z-scores for year 2002, 2003 and 2011 for the Centre region can be seen in figures 18 to 21, to make a review of this data easier the same graphs are presented in enlarged versions in Appendix E.

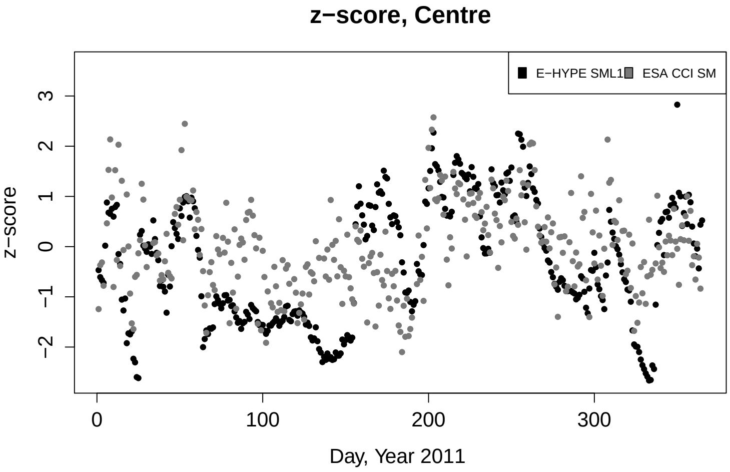


Figure 16: z-score of the E-HYPE SML1 variable and the ESA CCI SM product for segment 9525076 (located in the Centre region) for the 365 studied days year 2011

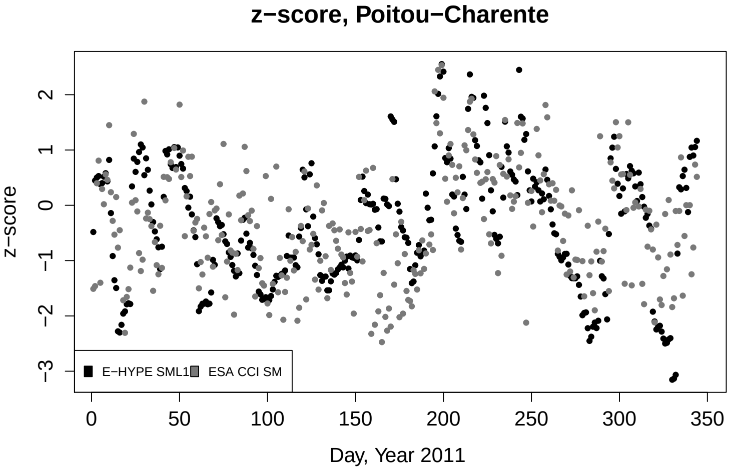


Figure 17: z-score of the E-HYPE SML1 variable and the ESA CCI SM product for segment 9525076 (located in the Poitou-Charente region) for the 342 studied days year 2011

Table 10: Pearson correlation and RMSE between the z-scores of the E-HYPE SML1 variable and the ESA CCI SM product in the Centre and Poitou-Charente region

Region	Pearson correlation	RMSE
Centre	0.537	1.08
Poitou-Charente	0.542	1.01

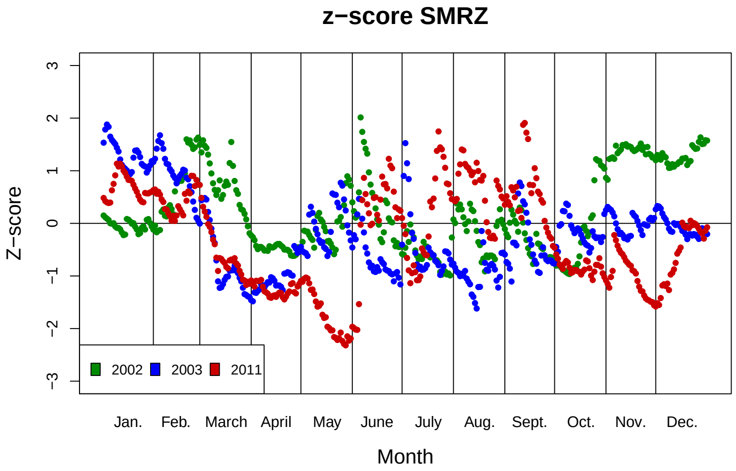


Figure 18: z-score of the E-HYPE SMRZ variable in segment 9525076 (located in the Centre region) for the 365 studied days years 2002, 2003 and 2011. Vertical lines separates the individual months and the horizontal line show the average soil moisture estimate.

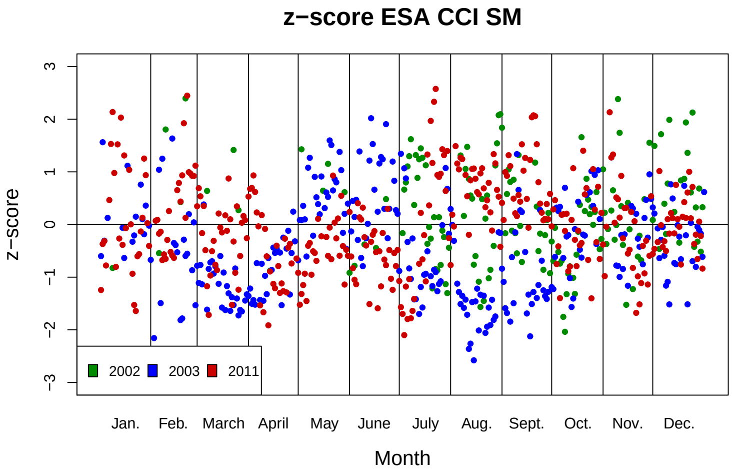


Figure 19: z-score of the ESA CCI SM product in segment 9525076 (located in the Centre region) for the 365 studied days years 2002, 2003 and 2011. Vertical lines separates the individual months and the horizontal line show the average soil moisture estimate.

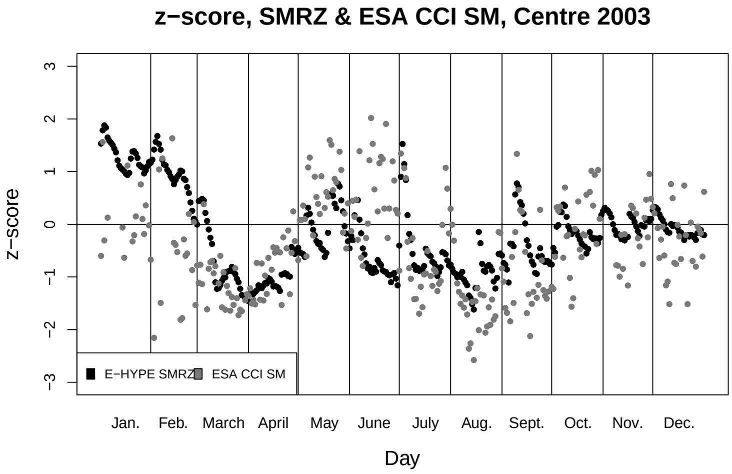


Figure 20: z-score of the E-HYPE SMRZ and the ESA CCI SM product in segment 9525076 (located in the Centre region) for the 365 studied days year 2003. Vertical lines separates the individual months and the horizontal line show the average soil moisture estimate.

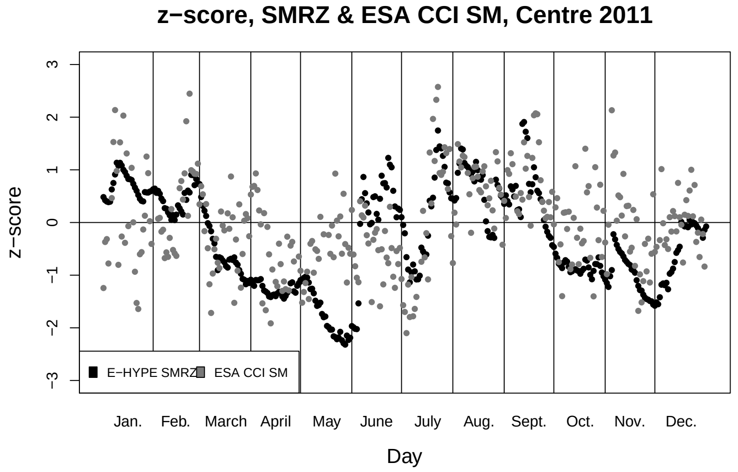


Figure 21: z-score of the E-HYPE SMRZ and the ESA CCI SM product in segment 9525076 (located in the Centre region) for the 365 studied days year 2011. Vertical lines separates the individual months and the horizontal line show the average soil moisture estimate.

4 Discussion

Before the individual results are discussed the reader should have noted that the SML1 variable for the segment 9525076 located in the Poitou-Charente region seems to reach some kind of maximum around 72 mm of soil moisture, probably due to internal model dynamics. Comparing it with the SML1 variable from the Centre region it shows that E-HYPE model model each segment differently in terms of layer structure and dynamics. Since the Copernicus SWI and ESA CCI SM products estimates relative soil moisture (which rarely reached a maximum) this makes comparison with the SML1 in the Poitou-Charente region difficult when performing correlation analysis.

4.1 Finding the factor T giving the highest correlation between studied Copernicus SWI and E-HYPE outputs

The comparison between the E-HYPE variables and the Copenicus SWI estimates in segment 9564959 in the Centre region indicated that the SML1 variable had the highest Pearson correlation to SWI_{T5} (r=0.856), SMRZ to SWI_{T40} (r=0.903) and SRFP with SWI_{T40} (r=0.899). For segment 9525076 in the Poitou-Charente region indicated that the SML1 variable had the highest Pearson correlation to SWI_{T1} (r=0.840), SMRZ to SWI_{T20} (r=0.911) and SRFP with SWI_{T20} (r=0.911).

The analysis investigating the correlation and RMSE when comparing E-HYPE variables with the Copernicus SWI using different T factors shows that the correlations are consistent in power, but different. Segment 9564959 in the Centre region demands a higher T factor to be used in comparison to the same E-HYPE variables in segment 9525076 in the Poitou-Charente region, one can only speculate in causes but it could possibly be due to different ground textures at the two sites. It shows that the soil water index (SWI) indeed gives different results depending on the location of the site.

For the Centre data it seems that the different data sets, studying qualitatively the figures showing the highest Pearson correlation, alter in shape and form between the SML1 layer, the SRMZ and SRFP layer. SMRZ and SRFP shows similar behavior, however, looking at the SRFP layer, it seems that the error is larger for lower degrees of saturation.

The data from Poitou-Charente shows a more coherent structure regarding the SML1 layer compared to the Centre data. The similarity between the SMRZ and SRFP variable in the Poitou-Charente region are not as distinct as compared to those in the Centre region.

One can conclude that the results from the comparison between the E-HYPE variables and the Copernicus SWI data gives different results depending on the segment which is studied, thus if one is to use the Copernicus SWI data together with the E-HYPE variables there could be need of an individual assessment for each segment in the model to find where and when correlation is strongest and under what conditions.

When looking at the range dynamics between the sorted series of E-HYPE SRFP and Copernicus SWI one can note that for the Poitou Charente region the span of the SWI_{20} soil moisture estimates is larger compared to that of the E-HYPE SRFP data, the span is though relatively linear between the minimum and maximum values for this segment. Looking at the range of the data for the Centre region there is a rather sharp deviation between the sorted series, the distribution of "low" soil moisture estimates are different from the distribution of the "high" soil moisture estimates. All together, the data strengthen the notion that there is a relationship between the E-HYPE SRFP and Copernicus SWI data sets but put in question the strength of this relationship.

4.2 Comparison between the E-HYPE SML1 variable and the ESA CCI SM product

The Pearson correlation between the E-HYPE SML1 variable and the ESA CCI SM product was 0.795 and 0.760 for segment 9564959 in the Centre region and segment 9525076 in the Poitou-Charente region respectively.

Comparing the different soil moisture estimates one can, as was expected, see more fluctuations in the data representing smaller soil layers. Interestingly, one can note that the ESA CCI SM data never reaches a saturation of more than 50% while this is possible for the modeled estimates from the E-HYPE and Copernicus models, representing deeper soil layers.

The correlation between the E-HYPE SML1 data and the ESA CCI SM data is weaker than those making comparison to the Copernicus SWI, which is also expected since the ESA CCI data only represents the first centimeters of the soil layer while the SML1 variable goes deeper.

4.3 Comparison between the E-HYPE SML1 variable and aggregated ESA CCI SM data

The correlation between the E-HYPE SML1 variable and the ESA CCI SM data was able to be improved by averaging data over a number of days before making comparison. The best results were yielded when aggregating 10-day periods for segment 9564959 in the Centre region (r=0.873) and when aggregating 6-day periods for segment 9525076 in the Poitou-Charente region (r=0.884).

Unfortunately the results seem to be quite unstable and it is hard to distinguish a certain number of days which are generally to be considered optimal. The results are also unstable in the sense that the correlations are deviating between the two regions, where an improvement in the data for one region could mean a deterioration in the other. Interestingly , just creating a mean over two days improve the correlation directly, indicating that this type of smaller aggregation could be of interest for further studies.

4.4 Comparison of z-scores between the E-HYPE SML1 variable and the ESA CCI SM product

The z-score demonstrates the deviation from the mean soil moisture during the period 2000-2011, the earlier dryer soil moisture conditions during the first months of year 2011, which then increases above normal after the high amount of rainfalls occurring in June-July, is clearly visible looking at the data. For the Centre region the dry period is more clearly visible in comparison to the ESA CCI SM estimates, which is not the case for the Poitou-Charente region. The Pearson correlation and RMSE shows that the correlation between the z-scores are quite weak (r=0.537 and r=0.542 in the two segments). Possibly one can speculate that this deviation is due to the fact that it is possible for the E-HYPE model to react differently to long and short term drought events due to a deeper soil layer compared to the ESA CCI SM product.

To simplify for the reader, it is pointed out that the rest of the analysis in section 4.4 is referring to figures 18 to 21.

4.4.1 E-HYPE SMRZ estimates for year 2002, 2003 and 2011 in the Centre region

Looking at the variation of the E-HYPE SMRZ data for the three years 2002, 2003 and 2011 one note that the soil moisture estimates for year 2003 and 2011 were higher than average for the first two months and then decline well below average in March and April. Interestingly, the soil moisture estimates for year 2011 continue to decline during May while the 2003 estimates increases to an average level. The data for June to August show that the SMRZ data for year 2003 remain well below average with few exceptions for the whole period while the 2011 data is pending. The 2003 data then recuperates to normal levels during September and stay there for the rest of the year, the 2011 data falls and stays a level below average but recuperates in December. The SMRZ data for year 2002 is pending most of the year but is below average in a number of occasions, indicating that having high soil moisture levels might not be the only thing determining weather the yield accounts are to be above normal.

4.4.2 ESA CCI SM estimates for year 2002, 2003 and 2011 in the Centre region

The ESA CCI SM data for years 2002, 2003 and 2011 is harder to interpret since the variation is higher in the data. The ESA CCI SM data for year 2003 seems to show declining soil moisture values from the start of the year which recuperates in the end of April, showing soil moisture estimates above average for the months May and June, mostly below average the months July until September and then pending around 0 for the rest of the year. The data for year 2011 tends to be pending around average for January and May, pending around average in March but with the weight on estimates below average, lower than average values for the months March to the mid of July, being above average until September and then pending around 0 for the rest of the year. The data for year 2002 is somewhat scarce for the first half of the year and hard to interpret, showing mostly estimates pending around average for the months July to the mid of September, then lower than average until the mid of October and then pending around average but with the weight on estimates above average

4.4.3 Comparison between the ESA CCI SM data and the E-HYPE SMRZ variable for year 2003 and 2011 in the Centre region

The ESA CCI SM data has generally a higher degree of variation compared to the E-HYPE SMRZ variable, making it harder to study general trends. Looking more closely at the comparison between the E-HYPE SMRZ and ESA CCI SM product for the years 2003 and 2011, a number of things are noted.

For year 2003 the ESA CCI SM data show a rather similar behavior compared to the SMRZ estimates. Both estimates shows declining soil moisture values in the beginning of the year which recuperates in the end of April, the first two months being a bit harder to analyze for the ESA CCI SM data due to the variation. An interesting difference is that while the SMRZ data continues to be below average in the months May and June the ESA CCI SM data is above average. This could indicate that while the top soil layer has been receiving water and thus increased its soil moisture, the available water for plants etc. is still below average, as indicated by the SMRZ estimates, thus the E-HYPE model could be giving us additional information about the amount of soil moisture in the ground. But it could also be that the ESA CCI SM data catches processes not included into the E-HYPE model, for example added irrigation systems that has not been included in the E-HYPE model, further investigation could thus be needed, correlate the event to rainfall data for example, to investigate which of the cases is most probable. Both data sets stay mostly below average for the months July and August but the ESA CCI SM data indicates a much fiercer soil moisture deficit compared to the SMRZ estimates. Both data sets are pending around 0 for the rest of the year.

Looking at the comparison for year 2011, the ESA CCI SM data is highly varying during the first two months January and February, showing soil moisture estimates both above and below average, while the E-HYPE

SMRZ data shows constantly "positive" soil moisture estimates above average. From the beginning of February to the end of May the E-HYPE SMRZ variable is under almost constant decline, a development which is not seen as clearly in the ESA CCI SM data, even though it shows mostly "negative" soil moisture values below average. The SMRZ soil moisture variable is drastically changed in June presenting a rather high amount of soil moisture content in the soil compared to the earlier months, a process not as visible in the ESA CCI SM data. During July the SMRZ variable firstly declines and then recuperates to levels well above average, a process quite visible also in the ESA CCI SM data but with higher/lower values in both extremes. Both data sets show rather similar behavior during the months August to September being mostly above average, however for the last three months October to December, the SMRZ variable declines and shows largely soil moisture estimates below average, while the ESA CCI SM products shows high degree of variation pending both above and below average.

This analysis indicates that the E-HYPE model and the ESA CCI SM estimates could be useful for evaluating drought dynamics with emphasis on soil moisture data, it is probable that it is not only viable to study degree of deviation during a year of drought, but also the amount of time in which a certain deviation is existing, during which periods etc. Furthermore the comparison between the E-HYPE and ESA CCI SM data show that the two methods could possibly be used to complement each other under certain conditions, but that there could be great advantages in using modeled soil moisture estimates to get a clearer picture of the soil moisture situation in the deeper soil layers during an agricultural drought event.

4.5 Looking towards the future

On a planet with limited resources and with an increased strain on commodities such as water, it is likely that tools for accurately estimating the soil moisture status of the ground will be of increased importance. Different models are already used in drought alert systems, improving the ability for for decision makers to take action when alert is raised. This thesis only scratched on the surface of the numerous different E-HYPE variables available, and how these could be used to assess drought scenarios, however the thesis shows that the E-HYPE model has the ability to be used for drought event assessment.

There is an extensive ongoing research on how to assess soil moisture from space, and an increase in tools available for modelers, researchers and even citizens to use. These estimates are continuously improving and they can pose a fantastic source of information also for modelers looking at the deeper soil layers. This thesis show that there exist similarities between the E-HYPE modeled soil moisture in the deeper soil layers and the estimated amount of soil moisture in the topsoil, using satellite remote sensing technology. In the future, the information from remotely sensed satellite data could possibly be used as a tool for assessment of the soil moisture variable in the E-HYPE model under conditions when the validation shows poor results, surface soil moisture information could even be incorporated into the E-HYPE model to try to improve results, more studies on this subject is thus needed.

4.6 Limitations and Sources of errors

One of the thesis's limitations is that it only looked at two agricultural drought events at only two locations, and another mayor source of error is the limited amount of knowledge of the actual conditions on the studied sites. To be able to perform analysis, conclusions have been made on data covering a large area and then this data have been down scaled and said to be representative for a much smaller area, for example in terms of annual yield output and the actual occurrence of a agricultural drought event.

The claim that a agricultural drought event occurred in the studied regions during the period can also be questioned, this assumption is made based on yield reports covering large regions of France and the conclusion that there was a loss in agricultural yield due to low soil moisture conditions is only mentioned as a probable cause and no deeper analysis was made to conclude that this was the case (time being a limiting factor when working on this thesis).

The z-score is calculated based upon 11 years of data, while longer time series of both the E-HYPE and ESA CCI SM data are available, this was not done due to time shortage and also due to a shorter amount of agricultural yield data. The assumption that the data is normally distributed should also be checked properly.

5 Conclusion

The aim of this thesis was originally to try to study the impact of an agricultural drought event by looking at soil moisture data from different data sources. However, the concepts of agricultural drought was shown to be difficult to quantify in a way which made it possible to make comparison, the relationship between soil moisture and agricultural drought is strong but still very dynamic. Different drought response mechanisms used by plants at different growth states makes it hard to connect the degree of agricultural yield loss to the degree of soil moisture deviation, at least with by my original knowledge of the subject and the time given for a thesis.

One thing the thesis do show is that it is possible for any individual to take part and make use of different types of data regarding soil moisture and also get their hands of the tools to analyze this data, free of charge for the user.

In the end, this thesis showed that correlation exist between the estimates from the E-HYPE model compared to the Copernicus SWI and the ESA CCI SM estimates, and that it is possible to improve the correlation to the ESA CCI SM data by using a type of analysis where one use aggregated data over time. The thesis also showed that it is possible to make analysis of drought events using information from the E-HYPE model´s SMRZ variable and the ESA CCI SM satellite product data, where they showed similar behavior of deviations from their means over the course of two different periods of drought. Possibly there exists times where these two data sets could be used to complement one another, when information from one product contradicts the information of the other.

As a suggestion for future studies, it would be interesting to limit the scope to using only one of the different models or products mentioned in the thesis and then look more closely into multiple individual agricultural drought events (for example by looking more closely into the year 2003 where there were major droughts occurring in Europe). This would make it possible to get a deeper understanding of how one can interpret the soil moisture estimates and its possible effects on agricultural crops. Investigating how one can connect data from multiple sources such as the E-HYPE model, the European Drought Observatory (combining soil moisture estimates with radiative information about plant growth status) and actual local agricultural yield data from selected regions, would also be of great interest to get a deeper understanding of the model's performance in the area of agricultural drought. Also, there would be interesting to see if one can incorporate remotely sensed data into the E-HYPE model for improving performance in segments where the validation have yielded poor results.

6 References

References

AMS. Ams website. Available at: https://www.ametsoc.org/ams/index.cfm/about-ams/ams-statements/statements-of-the-ams-in-force/drought/, 2017. Accessed 2017-04-10, Adopted by the AMS Council on 19 September 2013.

Oliver Arino, Sophie Bontemps, Eric van Bogaert, Pierre Defourny, and Vasileios Kalogirou. Globcover 2009 - products description and validation report. Available at: `http://due.esrin.esa.int/page_globcover.php`, 2011. Accessed: 2016-11-13.

Z. Bartalis, V. Naeimi, S. Hasenauer, and W. Wagner. Ascat soil moisture product handbook. ascat soil moisture report series, 2008. Institute of Photogrammetry and Remote Sensing, Vienna University of Technology, Austria.

Jitendra Behari. *Microwave Dielectric Behaviour of Wet Soils*, volume 8th. Anamya Publishers, New Delhi, India, 2005.

S. Bojinski, T. C. Verstraete, C. Peterson, A. Simmons, and M. Zemp. The concept of essential climate variables in support of climate research. *Bulletin of the American Meteorological Society*, 95:1431–1443, 2014. doi:10.1175/BAMS-D-13-00047.1.

Eusang Cho, Minha Choi, and Wolfgang Wagner. An assessment of remotely sensed surface and root zone soil moisture through active and passice rensors in northeast asia. *Remote Sensing of Environment*, 160:166–179, 2015.

R. Chung, D.and Kidd, W. Wagner, W. Dorigo, and R. De Jeu. Esa cci soil moisture product specification document. Available at: `http://www.esa-soilmoisture-cci.org/node/145`, 07 2015. Accessed: 2016-11-02.

Ph. Ciais, M. Reichstein, N. Viovy, A. Granier, J. Ogee, V. Allard, M. Aubinet, N. Buchmann, Chr. Bernhofer, A. Carrara, F. Chevallier, N. De Noblet, A. D. Friend, P. Friedlingstein, T. Grunwald, B. Heinesch, P. Keronen, A. Knohl, G. Krinner, D. Loustau, G. Manca, G. Matteucci, F. Miglietta, J. M. Ourcival, D. Papale, K. Pilegaard, S. Rambal, G. Seufert, J. F. Soussana, M. J. Sanz, E. D. Schulze, T. Vesala, and R. Valentini. Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, 437:529–533, 2005. doi:10.1038/nature03972.

European Commission. Water scarcity and drought, first interim report. Available at: `http://ec.europa.eu/environment/water/pdf/1st_report.pdf`, 2006. Accessed: 2016-08-04.

European Commission. Communication from the commission to the european parliament and the council - addressing the challenge of water scarcity and droughts in the european union, com(2007) 414 final. Available at: http://ec.europa.eu/environment/water/quantity/eu_action.htm#2008_report, 2007. Accessed: 2016-11-13.

website Copernicus. Copernicus global land service website. Available at: `land.copernicus.eu/global/faq/how-convert-swi-hdf5-data-geotiff`. Visited 2017-02-11.

website Creative Commons. Creative commons website. Available at: `https://creativecommons.org/licenses/by-sa/3.0/`. Visited 2017-02-10.

Cat Downy. Esa cci webpage. Available at: `http://cci.esa.int/objective`, 04 2016. Accessed: 2016-11-01.

Bernd Fischer and Gregoire Pau. *rhdf5: HDF5 interface to R*. R package version 2.14.0.

Catriona M. K. Gardner, David Robinson, Ken Blyth, and David J. Cooper. Soil water content. In Keith A. Smith and Chris E. Mullins, editors, *Soil and Environmental Analysis: Physical Methods*, chapter Chapter 1, pages 1–64. Marcel Dekker, Inc., second edition edition, 2000.

GCOS. Implementation plan for the global observing system for climate in support of the unfccc (2010 update,gcos-138). Available at: http://library.wmo.int/opac/index.php?lvl=notice_display&id=449#.V0LD3fl97IU, 08 2010. Accessed: 2016-11-08.

website Globecover. Globecover website. Available at: `http://due.esrin.esa.int/page_globcover.php`. Visited 2017-02-10.

Robert J. Hijmans. *raster: Geographic Data Analysis and Modeling*, 2015. URL `https://CRAN.R-project.org/package=raster`. R package version 2.5-2.

R. Kidd, H. Makhmara, and C. Paulik. Gio global land component - lot i 'operation of the global land component' - product user manual - issue i1.30. Available at: `http://land.copernicus.eu/global/products/SWI`, 2014. Accessed: 2016-12-28.

Goran Lindstrom, Charlotta Pers, Jorgen Rosberg, Johan Stromqvist, and Berit Arheimer. Development and testing of the hype (hydrological predictions of the environment) water quality model for different spatial scales. *Hydrology Research*, 41.3-4:295–319, 2010.

Y. Y. Liu, R. M. Parinussa, W. A. Dorigo, W. De Jeu, R. A. M.and Wanger, A. I. J. M. van Dijk, A. I. Van Dijk, M. F. McCabe, and J. P. Evans. Developing an improved soil moisture datase by blending passive and active microwave satellite-based retreivals. *Hydrology and Earth System Sciences*, 15:425–436, 2011.

Y. Y. Liu, W. A. Dorigo, R. M. Parinussa, R. A. M. De Jeu, W. Wagner, M. F. McCabe, J. P. Evans, and A. I. J. M. Van Dijk. Trend-preserving blanding of passive and active microwave soil moisture retreivals. *Remote Sensing of Environment*, 123:280–297, 2012.

MARS. Mars bulletin. Available at: `https://bookshop.europa.eu/en/home/`, 2011a. ISSN: 1831 - 9793 and Catalogue number: LB-AM-11-08-EN-C.

MARS. Mars bulletin. Available at: `https://bookshop.europa.eu/en/home/`, 2011b. ISSN: 1831 - 9793 and Catalogue number: LB-AM-11-011-EN-C.

MARS. Mars bulletin. Available at: `https://bookshop.europa.eu/en/home/`, 2011c. ISSN: 1831 - 9793 and Catalogue number: LB-AM-11-022-EN-C.

Amy McNally, Shraddhanand Shukla, Kristi R. Arsenault, Shugong Wang, Christa D. Peters-Lidard, and James P. Verdin. Evaluating esa cci soil moisture in east africa. *International Journal of Applied Earth Observation and Geoinformation*, 48:96–109, 2016.

website OSGeo. Open source geospatial foundation website. Available at: `https://trac.osgeo.org/osgeo4w/`. Visited 2017-02-11.

C. Paulik. Gio global land component - lot i 'operation of the global land component' - algorithm theoretical basis document - product user manual - i2.40. Available at: `http://land.copernicus.eu/global/products/SWI`, 2016. Accessed: 2016-12-28.

George P. Petropoulos, Hywel M. Griffiths, Wouter Dorigo, Angelika Xave, and Alexander Gruber. Surface soil moisture estimation: Significance, controls, and conventional measurement techniques. In Georpe P. Petrolpoulos, editor, *Remote Sensing of Energy Fluxes and Soil Moisture Content*, chapter Chapter 2, pages 29–45. Taylor and Francis Group, 6000 Broken Sound Parkway NW, Suite 300, first edition edition, 2014.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. URL `https://www.R-project.org/`.

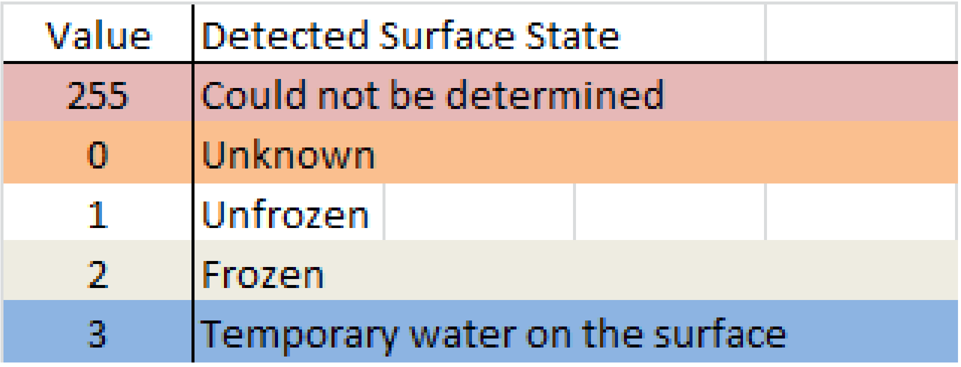
M. Robinson and R. C. Ward. *Principles of Hydrology*, volume 4th edition. McGraw Hill Publishing Company, Berkshire, England, 2000.

G. Sepulcre-Canto, S. Horin, A. Singleton, H. Carrao, and J. Vogt. Development of a combined drought indicator to detext agricultural drought in europe. *Natural Hazards and Earth System Sciences*, 12:3519–3531, 2012.

website SMHI. Smhi website. Available at: http://www.smhi.net/hype/wiki/doku.php?id=start:hype_model_description:hype_land. Visited 2017-02-10.

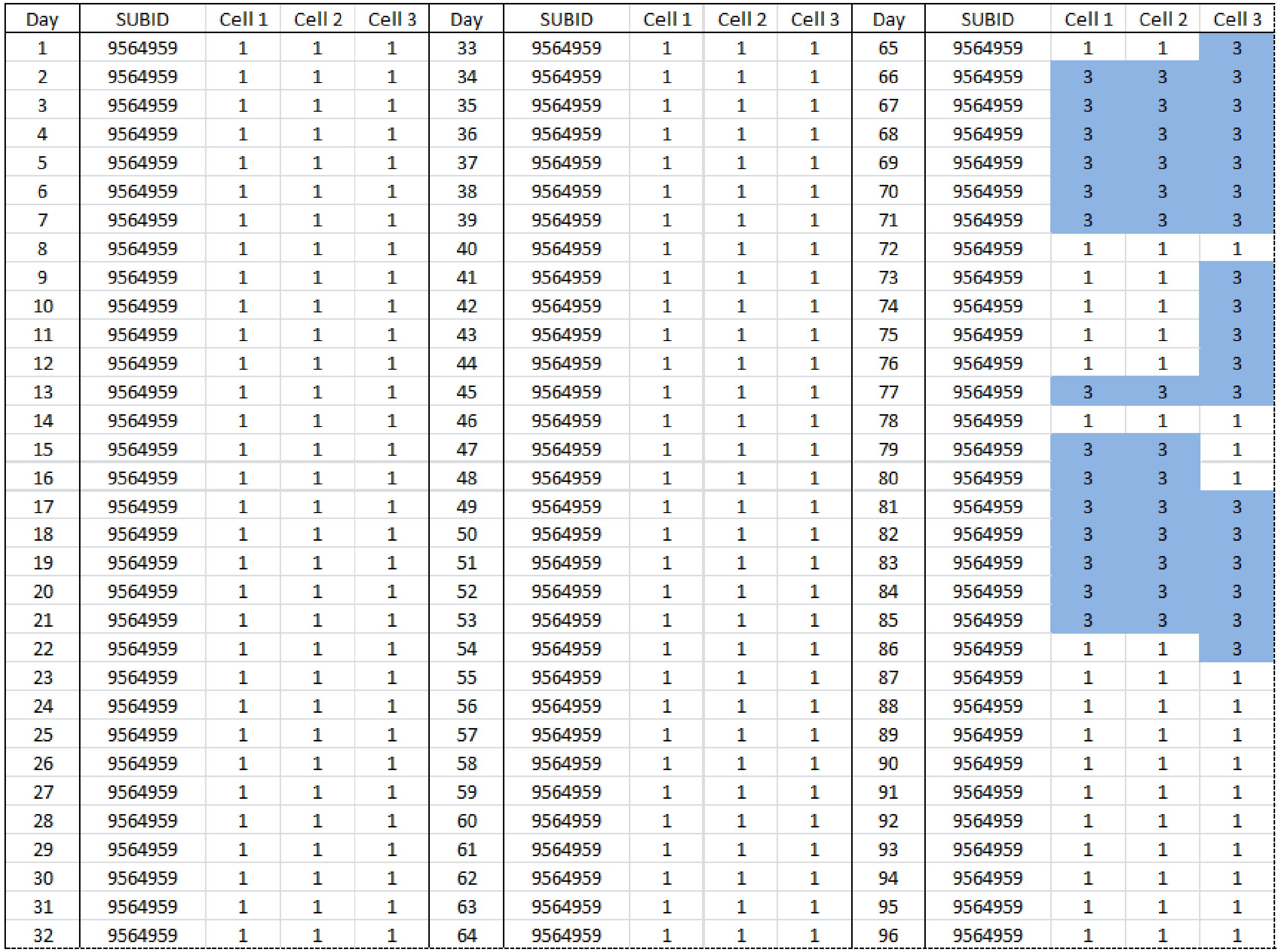
W. Wagner, W. Dorigo, R. de Jeu, D. Fernandez, J. Benveniste, E. Haas, and M. Ertl. Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture. *ISPRS Ann. Remote Sens. Spatial Inf. Sci.*, I-7:315–321, 2012. doi:10.5194/isprsannals-I-7-315-2012.

A Appendix - Surface states year 2011



Value	Detected Surface State
255	Could not be determined
0	Unknown
1	Unfrozen
2	Frozen
3	Temporary water on the surface

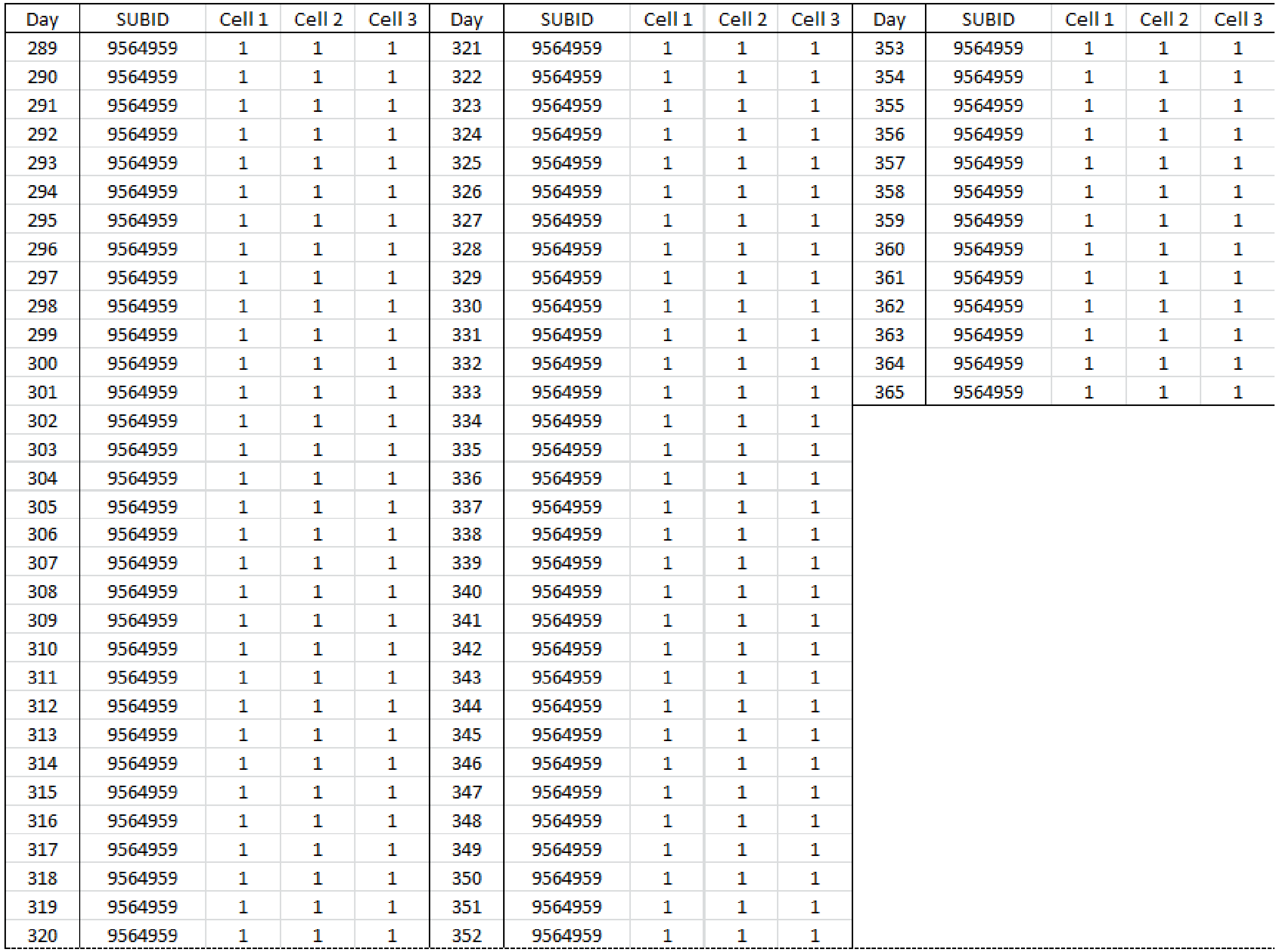
Figure 22: Definition of Surface State Flag values



Day	SUBID	Cell 1	Cell 2	Cell 3	Day	SUBID	Cell 1	Cell 2	Cell 3	Day	SUBID	Cell 1	Cell 2	Cell 3
1	9564959	1	1	1	33	9564959	1	1	1	65	9564959	1	1	3
2	9564959	1	1	1	34	9564959	1	1	1	66	9564959	3	3	3
3	9564959	1	1	1	35	9564959	1	1	1	67	9564959	3	3	3
4	9564959	1	1	1	36	9564959	1	1	1	68	9564959	3	3	3
5	9564959	1	1	1	37	9564959	1	1	1	69	9564959	3	3	3
6	9564959	1	1	1	38	9564959	1	1	1	70	9564959	3	3	3
7	9564959	1	1	1	39	9564959	1	1	1	71	9564959	3	3	3
8	9564959	1	1	1	40	9564959	1	1	1	72	9564959	1	1	1
9	9564959	1	1	1	41	9564959	1	1	1	73	9564959	1	1	3
10	9564959	1	1	1	42	9564959	1	1	1	74	9564959	1	1	3
11	9564959	1	1	1	43	9564959	1	1	1	75	9564959	1	1	3
12	9564959	1	1	1	44	9564959	1	1	1	76	9564959	1	1	3
13	9564959	1	1	1	45	9564959	1	1	1	77	9564959	3	3	3
14	9564959	1	1	1	46	9564959	1	1	1	78	9564959	1	1	1
15	9564959	1	1	1	47	9564959	1	1	1	79	9564959	3	3	1
16	9564959	1	1	1	48	9564959	1	1	1	80	9564959	3	3	1
17	9564959	1	1	1	49	9564959	1	1	1	81	9564959	3	3	3
18	9564959	1	1	1	50	9564959	1	1	1	82	9564959	3	3	3
19	9564959	1	1	1	51	9564959	1	1	1	83	9564959	3	3	3
20	9564959	1	1	1	52	9564959	1	1	1	84	9564959	3	3	3
21	9564959	1	1	1	53	9564959	1	1	1	85	9564959	3	3	3
22	9564959	1	1	1	54	9564959	1	1	1	86	9564959	1	1	3
23	9564959	1	1	1	55	9564959	1	1	1	87	9564959	1	1	1
24	9564959	1	1	1	56	9564959	1	1	1	88	9564959	1	1	1
25	9564959	1	1	1	57	9564959	1	1	1	89	9564959	1	1	1
26	9564959	1	1	1	58	9564959	1	1	1	90	9564959	1	1	1
27	9564959	1	1	1	59	9564959	1	1	1	91	9564959	1	1	1
28	9564959	1	1	1	60	9564959	1	1	1	92	9564959	1	1	1
29	9564959	1	1	1	61	9564959	1	1	1	93	9564959	1	1	1
30	9564959	1	1	1	62	9564959	1	1	1	94	9564959	1	1	1
31	9564959	1	1	1	63	9564959	1	1	1	95	9564959	1	1	1
32	9564959	1	1	1	64	9564959	1	1	1	96	9564959	1	1	1

Day	SUBID	Cell 1	Cell 2	Cell 3	Day	SUBID	Cell 1	Cell 2	Cell 3	Day	SUBID	Cell 1	Cell 2	Cell 3
97	9564959	1	1	3	129	9564959	1	1	1	161	9564959	1	1	1
98	9564959	1	1	3	130	9564959	1	1	1	162	9564959	1	1	1
99	9564959	3	3	3	131	9564959	1	1	1	163	9564959	1	1	1
100	9564959	3	3	3	132	9564959	1	1	1	164	9564959	1	1	1
101	9564959	3	3	3	133	9564959	1	1	1	165	9564959	1	1	1
102	9564959	3	3	3	134	9564959	1	1	1	166	9564959	1	1	1
103	9564959	3	3	3	135	9564959	1	1	1	167	9564959	1	1	1
104	9564959	3	3	3	136	9564959	1	1	1	168	9564959	1	1	1
105	9564959	3	3	3	137	9564959	1	1	1	169	9564959	1	1	1
106	9564959	3	3	3	138	9564959	1	1	1	170	9564959	1	1	1
107	9564959	3	3	3	139	9564959	1	1	1	171	9564959	1	1	1
108	9564959	3	3	3	140	9564959	1	1	1	172	9564959	1	1	1
109	9564959	3	3	3	141	9564959	1	1	1	173	9564959	1	1	1
110	9564959	3	3	3	142	9564959	1	1	1	174	9564959	1	1	1
111	9564959	3	3	3	143	9564959	1	1	1	175	9564959	1	1	1
112	9564959	3	3	3	144	9564959	1	1	1	176	9564959	1	1	1
113	9564959	3	3	3	145	9564959	1	1	1	177	9564959	1	1	1
114	9564959	3	3	3	146	9564959	1	1	1	178	9564959	1	1	1
115	9564959	3	3	3	147	9564959	1	1	1	179	9564959	1	1	1
116	9564959	3	3	3	148	9564959	1	1	1	180	9564959	1	1	1
117	9564959	1	1	1	149	9564959	1	1	1	181	9564959	1	1	1
118	9564959	1	1	1	150	9564959	1	1	1	182	9564959	1	1	1
119	9564959	1	1	1	151	9564959	1	1	1	183	9564959	1	1	1
120	9564959	1	1	1	152	9564959	1	1	1	184	9564959	1	1	1
121	9564959	1	1	1	153	9564959	1	1	1	185	9564959	1	1	1
122	9564959	1	1	1	154	9564959	1	1	1	186	9564959	1	1	1
123	9564959	1	1	1	155	9564959	1	1	1	187	9564959	1	1	1
124	9564959	1	1	1	156	9564959	1	1	1	188	9564959	1	1	1
125	9564959	1	1	1	157	9564959	1	1	1	189	9564959	1	1	1
126	9564959	1	1	1	158	9564959	1	1	1	190	9564959	1	1	1
127	9564959	1	1	1	159	9564959	1	1	1	191	9564959	1	1	1
128	9564959	1	1	1	160	9564959	1	1	1	192	9564959	1	1	1

Day	SUBID	Cell 1	Cell 2	Cell 3	Day	SUBID	Cell 1	Cell 2	Cell 3	Day	SUBID	Cell 1	Cell 2	Cell 3
193	9564959	1	1	1	225	9564959	1	1	1	257	9564959	1	1	1
194	9564959	1	1	1	226	9564959	1	1	1	258	9564959	1	1	1
195	9564959	1	1	1	227	9564959	1	1	1	259	9564959	1	1	1
196	9564959	1	1	1	228	9564959	1	1	1	260	9564959	1	1	1
197	9564959	1	1	1	229	9564959	1	1	1	261	9564959	1	1	1
198	9564959	1	1	1	230	9564959	1	1	1	262	9564959	1	1	1
199	9564959	1	1	1	231	9564959	1	1	1	263	9564959	1	1	1
200	9564959	1	1	1	232	9564959	1	1	1	264	9564959	1	1	1
201	9564959	1	1	1	233	9564959	1	1	1	265	9564959	1	1	1
202	9564959	1	1	1	234	9564959	1	1	1	266	9564959	1	1	1
203	9564959	1	1	1	235	9564959	1	1	1	267	9564959	1	1	1
204	9564959	1	1	1	236	9564959	1	1	1	268	9564959	1	1	1
205	9564959	1	1	1	237	9564959	1	1	1	269	9564959	1	1	1
206	9564959	1	1	1	238	9564959	1	1	1	270	9564959	1	1	1
207	9564959	1	1	1	239	9564959	1	1	1	271	9564959	1	1	1
208	9564959	1	1	1	240	9564959	1	1	1	272	9564959	1	1	1
209	9564959	1	1	1	241	9564959	1	1	1	273	9564959	1	1	1
210	9564959	1	1	1	242	9564959	1	1	1	274	9564959	1	1	1
211	9564959	1	1	1	243	9564959	1	1	1	275	9564959	1	1	1
212	9564959	1	1	1	244	9564959	1	1	1	276	9564959	1	1	1
213	9564959	1	1	1	245	9564959	1	1	1	277	9564959	1	1	1
214	9564959	1	1	1	246	9564959	1	1	1	278	9564959	1	1	1
215	9564959	1	1	1	247	9564959	1	1	1	279	9564959	1	1	1
216	9564959	1	1	1	248	9564959	1	1	1	280	9564959	1	1	1
217	9564959	1	1	1	249	9564959	1	1	1	281	9564959	1	1	1
218	9564959	1	1	1	250	9564959	1	1	1	282	9564959	1	1	1
219	9564959	1	1	1	251	9564959	1	1	1	283	9564959	1	1	1
220	9564959	1	1	1	252	9564959	1	1	1	284	9564959	1	1	1
221	9564959	1	1	1	253	9564959	1	1	1	285	9564959	1	1	1
222	9564959	1	1	1	254	9564959	1	1	1	286	9564959	1	1	1
223	9564959	1	1	1	255	9564959	1	1	1	287	9564959	1	1	1
224	9564959	1	1	1	256	9564959	1	1	1	288	9564959	1	1	1

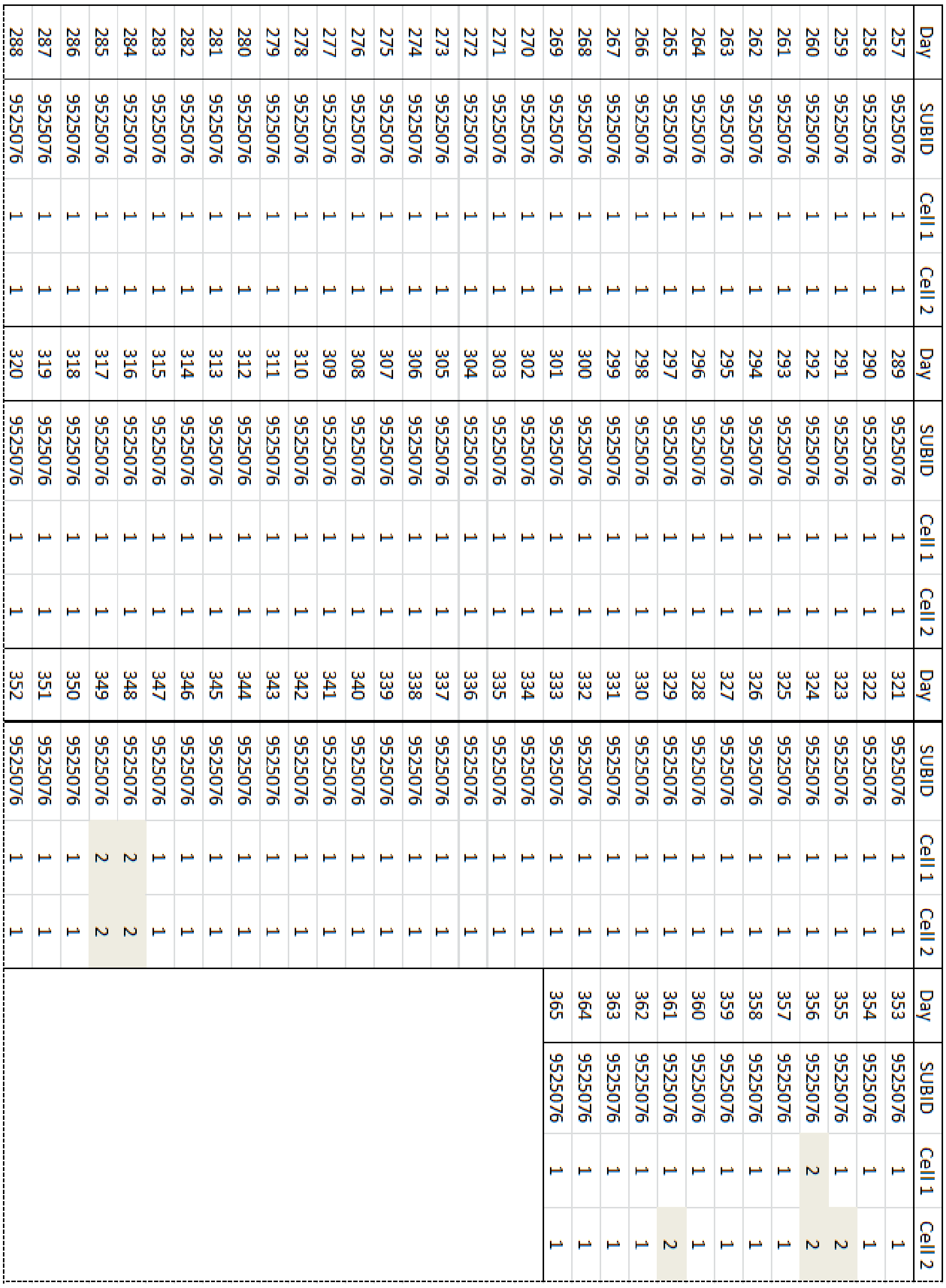


Day	SUBID	Cell 1	Cell 2	Cell 3	Day	SUBID	Cell 1	Cell 2	Cell 3	Day	SUBID	Cell 1	Cell 2	Cell 3
289	9564959	1	1	1	321	9564959	1	1	1	353	9564959	1	1	1
290	9564959	1	1	1	322	9564959	1	1	1	354	9564959	1	1	1
291	9564959	1	1	1	323	9564959	1	1	1	355	9564959	1	1	1
292	9564959	1	1	1	324	9564959	1	1	1	356	9564959	1	1	1
293	9564959	1	1	1	325	9564959	1	1	1	357	9564959	1	1	1
294	9564959	1	1	1	326	9564959	1	1	1	358	9564959	1	1	1
295	9564959	1	1	1	327	9564959	1	1	1	359	9564959	1	1	1
296	9564959	1	1	1	328	9564959	1	1	1	360	9564959	1	1	1
297	9564959	1	1	1	329	9564959	1	1	1	361	9564959	1	1	1
298	9564959	1	1	1	330	9564959	1	1	1	362	9564959	1	1	1
299	9564959	1	1	1	331	9564959	1	1	1	363	9564959	1	1	1
300	9564959	1	1	1	332	9564959	1	1	1	364	9564959	1	1	1
301	9564959	1	1	1	333	9564959	1	1	1	365	9564959	1	1	1
302	9564959	1	1	1	334	9564959	1	1	1					
303	9564959	1	1	1	335	9564959	1	1	1					
304	9564959	1	1	1	336	9564959	1	1	1					
305	9564959	1	1	1	337	9564959	1	1	1					
306	9564959	1	1	1	338	9564959	1	1	1					
307	9564959	1	1	1	339	9564959	1	1	1					
308	9564959	1	1	1	340	9564959	1	1	1					
309	9564959	1	1	1	341	9564959	1	1	1					
310	9564959	1	1	1	342	9564959	1	1	1					
311	9564959	1	1	1	343	9564959	1	1	1					
312	9564959	1	1	1	344	9564959	1	1	1					
313	9564959	1	1	1	345	9564959	1	1	1					
314	9564959	1	1	1	346	9564959	1	1	1					
315	9564959	1	1	1	347	9564959	1	1	1					
316	9564959	1	1	1	348	9564959	1	1	1					
317	9564959	1	1	1	349	9564959	1	1	1					
318	9564959	1	1	1	350	9564959	1	1	1					
319	9564959	1	1	1	351	9564959	1	1	1					
320	9564959	1	1	1	352	9564959	1	1	1					

Figure 23: Surface state flags for SUBID 9564959 (in the Centre region)

Day	SUBID	Cell 1	Cell 2	Day	SUBID	Cell 1	Cell 2	Day	SUBID	Cell 1	Cell 2	Day	SUBID	Cell 1	Cell 2
1	9525076	1	1	33	9525076	1	1	65	9525076	1	1	97	9525076	1	1
2	9525076	1	1	34	9525076	1	1	66	9525076	1	1	98	9525076	1	1
3	9525076	1	1	35	9525076	1	1	67	9525076	1	1	99	9525076	1	1
4	9525076	2	2	36	9525076	1	1	68	9525076	1	1	100	9525076	1	1
5	9525076	2	2	37	9525076	1	1	69	9525076	1	1	101	9525076	1	1
6	9525076	1	1	38	9525076	1	1	70	9525076	1	1	102	9525076	1	1
7	9525076	1	1	39	9525076	1	1	71	9525076	1	1	103	9525076	1	1
8	9525076	1	1	40	9525076	1	1	72	9525076	1	1	104	9525076	1	1
9	9525076	1	1	41	9525076	1	1	73	9525076	1	1	105	9525076	1	1
10	9525076	1	1	42	9525076	1	1	74	9525076	1	1	106	9525076	1	1
11	9525076	1	1	43	9525076	1	1	75	9525076	1	1	107	9525076	1	1
12	9525076	1	1	44	9525076	1	1	76	9525076	1	1	108	9525076	1	1
13	9525076	1	1	45	9525076	1	1	77	9525076	1	1	109	9525076	1	1
14	9525076	1	1	46	9525076	1	1	78	9525076	1	1	110	9525076	1	1
15	9525076	1	1	47	9525076	1	1	79	9525076	1	1	111	9525076	1	1
16	9525076	1	1	48	9525076	1	1	80	9525076	1	1	112	9525076	1	1
17	9525076	1	1	49	9525076	1	1	81	9525076	1	1	113	9525076	1	1
18	9525076	1	1	50	9525076	1	1	82	9525076	1	1	114	9525076	1	1
19	9525076	1	1	51	9525076	1	1	83	9525076	1	1	115	9525076	1	1
20	9525076	1	1	52	9525076	1	1	84	9525076	1	1	116	9525076	1	1
21	9525076	1	1	53	9525076	1	1	85	9525076	1	1	117	9525076	1	1
22	9525076	1	1	54	9525076	1	1	86	9525076	1	1	118	9525076	1	1
23	9525076	1	1	55	9525076	1	1	87	9525076	1	1	119	9525076	1	1
24	9525076	1	1	56	9525076	1	1	88	9525076	1	1	120	9525076	1	1
25	9525076	1	1	57	9525076	1	1	89	9525076	1	1	121	9525076	1	1
26	9525076	1	1	58	9525076	1	1	90	9525076	1	1	122	9525076	1	1
27	9525076	1	1	59	9525076	1	1	91	9525076	1	1	123	9525076	1	1
28	9525076	1	1	60	9525076	1	1	92	9525076	1	1	124	9525076	1	1
29	9525076	1	1	61	9525076	1	1	93	9525076	1	1	125	9525076	1	1
30	9525076	1	1	62	9525076	1	1	94	9525076	1	1	126	9525076	1	1
31	9525076	1	1	63	9525076	1	1	95	9525076	1	1	127	9525076	1	1
32	9525076	1	1	64	9525076	1	1	96	9525076	1	1	128	9525076	1	1

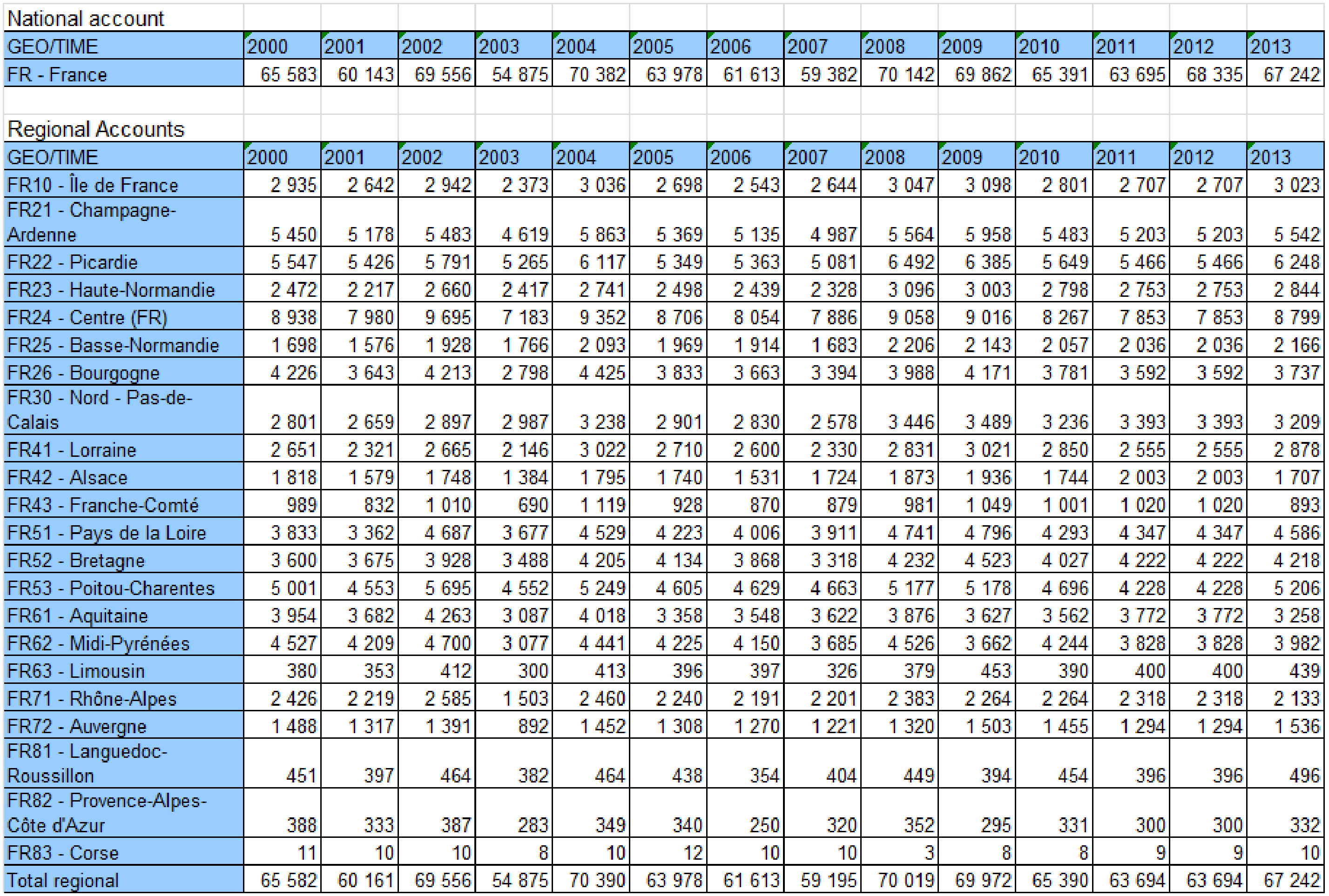
Day	SUBID	Cell 1	Cell 2	Day	SUBID	Cell 1	Cell 2	Day	SUBID	Cell 1	Cell 2	Day	SUBID	Cell 1	Cell 2
129	9525076	1	1	161	9525076	1	1	193	9525076	1	1	225	9525076	1	1
130	9525076	1	1	162	9525076	1	1	194	9525076	1	1	226	9525076	1	1
131	9525076	1	1	163	9525076	1	1	195	9525076	1	1	227	9525076	1	1
132	9525076	1	1	164	9525076	1	1	196	9525076	1	1	228	9525076	1	1
133	9525076	1	1	165	9525076	1	1	197	9525076	1	1	229	9525076	1	1
134	9525076	1	1	166	9525076	1	1	198	9525076	1	1	230	9525076	1	1
135	9525076	1	1	167	9525076	1	1	199	9525076	1	1	231	9525076	1	1
136	9525076	1	1	168	9525076	1	1	200	9525076	1	1	232	9525076	1	1
137	9525076	1	1	169	9525076	1	1	201	9525076	1	1	233	9525076	1	1
138	9525076	1	1	170	9525076	1	1	202	9525076	1	1	234	9525076	1	1
139	9525076	1	1	171	9525076	1	1	203	9525076	1	1	235	9525076	1	1
140	9525076	1	1	172	9525076	1	1	204	9525076	1	1	236	9525076	1	1
141	9525076	1	1	173	9525076	1	1	205	9525076	1	1	237	9525076	1	1
142	9525076	1	1	174	9525076	1	1	206	9525076	1	1	238	9525076	1	1
143	9525076	1	1	175	9525076	1	1	207	9525076	1	1	239	9525076	1	1
144	9525076	1	1	176	9525076	1	1	208	9525076	1	1	240	9525076	1	1
145	9525076	1	1	177	9525076	1	1	209	9525076	1	1	241	9525076	1	1
146	9525076	1	1	178	9525076	1	1	210	9525076	1	1	242	9525076	1	1
147	9525076	1	1	179	9525076	1	1	211	9525076	1	1	243	9525076	1	1
148	9525076	1	1	180	9525076	1	1	212	9525076	1	1	244	9525076	1	1
149	9525076	1	1	181	9525076	1	1	213	9525076	1	1	245	9525076	1	1
150	9525076	1	1	182	9525076	1	1	214	9525076	1	1	246	9525076	1	1
151	9525076	1	1	183	9525076	1	1	215	9525076	1	1	247	9525076	1	1
152	9525076	1	1	184	9525076	1	1	216	9525076	1	1	248	9525076	1	1
153	9525076	1	1	185	9525076	1	1	217	9525076	1	1	249	9525076	1	1
154	9525076	1	1	186	9525076	1	1	218	9525076	1	1	250	9525076	1	1
155	9525076	1	1	187	9525076	1	1	219	9525076	1	1	251	9525076	1	1
156	9525076	1	1	188	9525076	1	1	220	9525076	1	1	252	9525076	1	1
157	9525076	1	1	189	9525076	1	1	221	9525076	1	1	253	9525076	1	1
158	9525076	1	1	190	9525076	1	1	222	9525076	1	1	254	9525076	1	1
159	9525076	1	1	191	9525076	1	1	223	9525076	1	1	255	9525076	1	1
160	9525076	1	1	192	9525076	1	1	224	9525076	1	1	256	9525076	1	1



Day	SUBID	Cell 1	Cell 2	Day	SUBID	Cell 1	Cell 2	Day	SUBID	Cell 1	Cell 2	Day	SUBID	Cell 1	Cell 2
257	9525076	1	1	289	9525076	1	1	321	9525076	1	1	353	9525076	1	1
258	9525076	1	1	290	9525076	1	1	322	9525076	1	1	354	9525076	1	1
259	9525076	1	1	291	9525076	1	1	323	9525076	1	1	355	9525076	1	2
260	9525076	1	1	292	9525076	1	1	324	9525076	1	1	356	9525076	2	2
261	9525076	1	1	293	9525076	1	1	325	9525076	1	1	357	9525076	1	1
262	9525076	1	1	294	9525076	1	1	326	9525076	1	1	358	9525076	1	1
263	9525076	1	1	295	9525076	1	1	327	9525076	1	1	359	9525076	1	1
264	9525076	1	1	296	9525076	1	1	328	9525076	1	1	360	9525076	1	1
265	9525076	1	1	297	9525076	1	1	329	9525076	1	1	361	9525076	1	2
266	9525076	1	1	298	9525076	1	1	330	9525076	1	1	362	9525076	1	1
267	9525076	1	1	299	9525076	1	1	331	9525076	1	1	363	9525076	1	1
268	9525076	1	1	300	9525076	1	1	332	9525076	1	1	364	9525076	1	1
269	9525076	1	1	301	9525076	1	1	333	9525076	1	1	365	9525076	1	1
270	9525076	1	1	302	9525076	1	1	334	9525076	1	1				
271	9525076	1	1	303	9525076	1	1	335	9525076	1	1				
272	9525076	1	1	304	9525076	1	1	336	9525076	1	1				
273	9525076	1	1	305	9525076	1	1	337	9525076	1	1				
274	9525076	1	1	306	9525076	1	1	338	9525076	1	1				
275	9525076	1	1	307	9525076	1	1	339	9525076	1	1				
276	9525076	1	1	308	9525076	1	1	340	9525076	1	1				
277	9525076	1	1	309	9525076	1	1	341	9525076	1	1				
278	9525076	1	1	310	9525076	1	1	342	9525076	1	1				
279	9525076	1	1	311	9525076	1	1	343	9525076	1	1				
280	9525076	1	1	312	9525076	1	1	344	9525076	1	1				
281	9525076	1	1	313	9525076	1	1	345	9525076	1	1				
282	9525076	1	1	314	9525076	1	1	346	9525076	1	1				
283	9525076	1	1	315	9525076	1	1	347	9525076	1	1				
284	9525076	1	1	316	9525076	1	1	348	9525076	2	2				
285	9525076	1	1	317	9525076	1	1	349	9525076	2	2				
286	9525076	1	1	318	9525076	1	1	350	9525076	1	1				
287	9525076	1	1	319	9525076	1	1	351	9525076	1	1				
288	9525076	1	1	320	9525076	1	1	352	9525076	1	1				

Figure 24: Surface state flags for SUBID 9525076 (in the Poitou-Charente region)

B Appendix - Total regional and national production of cereals in France between 2000-2013



National account

GEO/TIME	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
FR - France	65 583	60 143	69 556	54 875	70 382	63 978	61 613	59 382	70 142	69 862	65 391	63 695	68 335	67 242

Regional Accounts

GEO/TIME	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
FR10 - Île de France	2 935	2 642	2 942	2 373	3 036	2 698	2 543	2 644	3 047	3 098	2 801	2 707	2 707	3 023
FR21 - Champagne-Ardenne	5 450	5 178	5 483	4 619	5 863	5 369	5 135	4 987	5 564	5 958	5 483	5 203	5 203	5 542
FR22 - Picardie	5 547	5 426	5 791	5 265	6 117	5 349	5 363	5 081	6 492	6 385	5 649	5 466	5 466	6 248
FR23 - Haute-Normandie	2 472	2 217	2 660	2 417	2 741	2 498	2 439	2 328	3 096	3 003	2 798	2 753	2 753	2 844
FR24 - Centre (FR)	8 938	7 980	9 695	7 183	9 352	8 706	8 054	7 886	9 058	9 016	8 267	7 853	7 853	8 799
FR25 - Basse-Normandie	1 698	1 576	1 928	1 766	2 093	1 969	1 914	1 683	2 206	2 143	2 057	2 036	2 036	2 166
FR26 - Bourgogne	4 226	3 643	4 213	2 798	4 425	3 833	3 663	3 394	3 988	4 171	3 781	3 592	3 592	3 737
FR30 - Nord - Pas-de-Calais	2 801	2 659	2 897	2 987	3 238	2 901	2 830	2 578	3 446	3 489	3 236	3 393	3 393	3 209
FR41 - Lorraine	2 651	2 321	2 665	2 146	3 022	2 710	2 600	2 330	2 831	3 021	2 850	2 555	2 555	2 878
FR42 - Alsace	1 818	1 579	1 748	1 384	1 795	1 740	1 531	1 724	1 873	1 936	1 744	2 003	2 003	1 707
FR43 - Franche-Comté	989	832	1 010	690	1 119	928	870	879	981	1 049	1 001	1 020	1 020	893
FR51 - Pays de la Loire	3 833	3 362	4 687	3 677	4 529	4 223	4 006	3 911	4 741	4 796	4 293	4 347	4 347	4 586
FR52 - Bretagne	3 600	3 675	3 928	3 488	4 205	4 134	3 868	3 318	4 232	4 523	4 027	4 222	4 222	4 218
FR53 - Poitou-Charentes	5 001	4 553	5 695	4 552	5 249	4 605	4 629	4 663	5 177	5 178	4 696	4 228	4 228	5 206
FR61 - Aquitaine	3 954	3 682	4 263	3 087	4 018	3 358	3 548	3 622	3 876	3 627	3 562	3 772	3 772	3 258
FR62 - Midi-Pyrénées	4 527	4 209	4 700	3 077	4 441	4 225	4 150	3 685	4 526	3 662	4 244	3 828	3 828	3 982
FR63 - Limousin	380	353	412	300	413	396	397	326	379	453	390	400	400	439
FR71 - Rhône-Alpes	2 426	2 219	2 585	1 503	2 460	2 240	2 191	2 201	2 383	2 264	2 264	2 318	2 318	2 133
FR72 - Auvergne	1 488	1 317	1 391	892	1 452	1 308	1 270	1 221	1 320	1 503	1 455	1 294	1 294	1 536
FR81 - Languedoc-Roussillon	451	397	464	382	464	438	354	404	449	394	454	396	396	496
FR82 - Provence-Alpes-Côte d'Azur	388	333	387	283	349	340	250	320	352	295	331	300	300	332
FR83 - Corse	11	10	10	8	10	12	10	10	3	8	8	9	9	10
Total regional	65 582	60 161	69 556	54 875	70 390	63 978	61 613	59 195	70 019	69 972	65 390	63 694	63 694	67 242

Figure 25: Table over national and regional total production of cereals* in France. Source: Eurostat

**Cereals (excluding rice) for the production of grain (including seed)*

C Appendix - Presentation of soil moisture estimates from the E-HYPE model, the ESA CCI SM product and the Copernicus SWI for regions Centre and Poitou-Charente

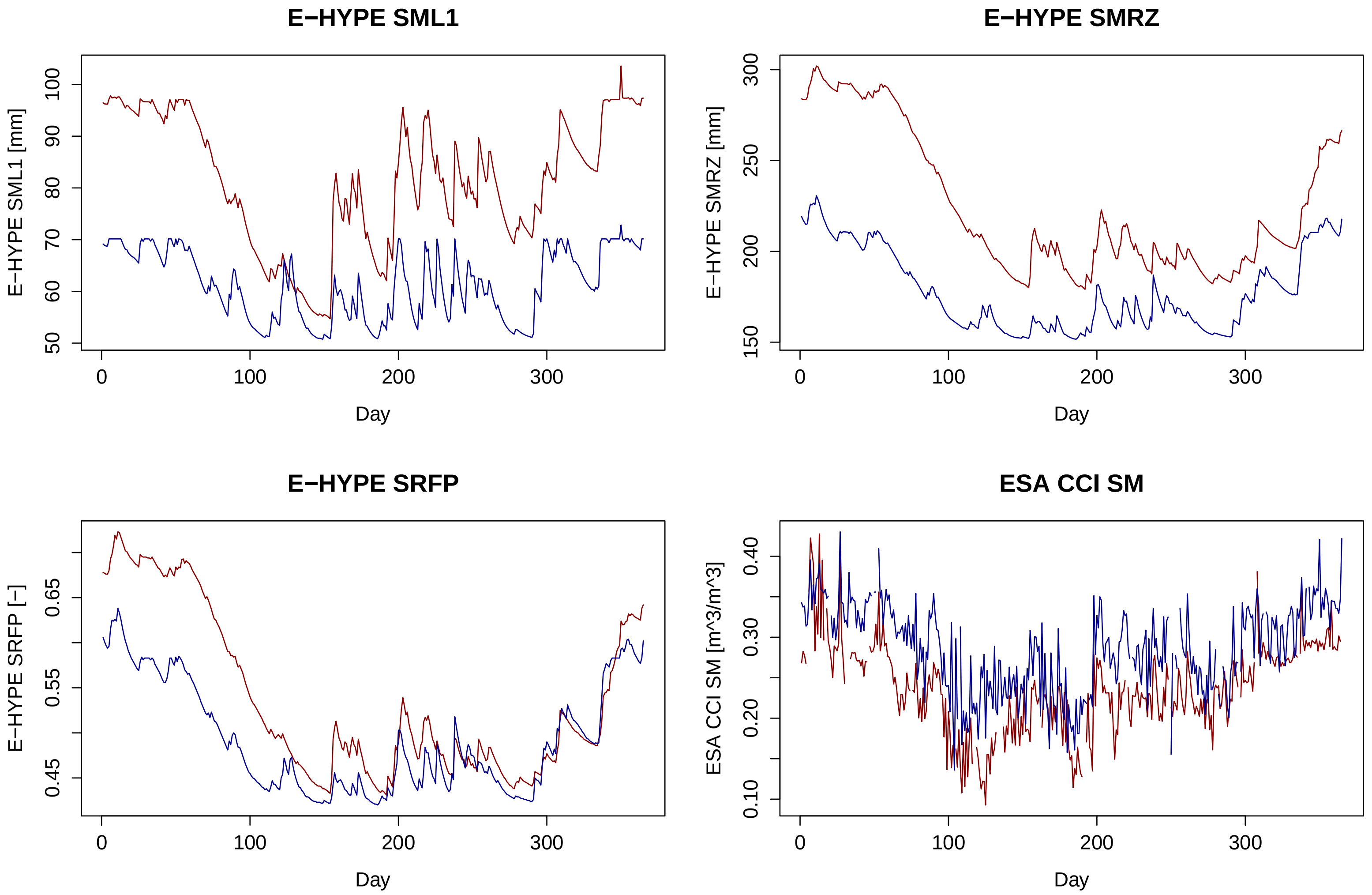


Figure 26: E-HYPE SML1, SMRZ, SRFP and the ESA CCI SM estimates for the two segments 9564959 (Centre, red color) and 9525076 (Poitou-Charente, blue color) for the 365 days during the year 2011

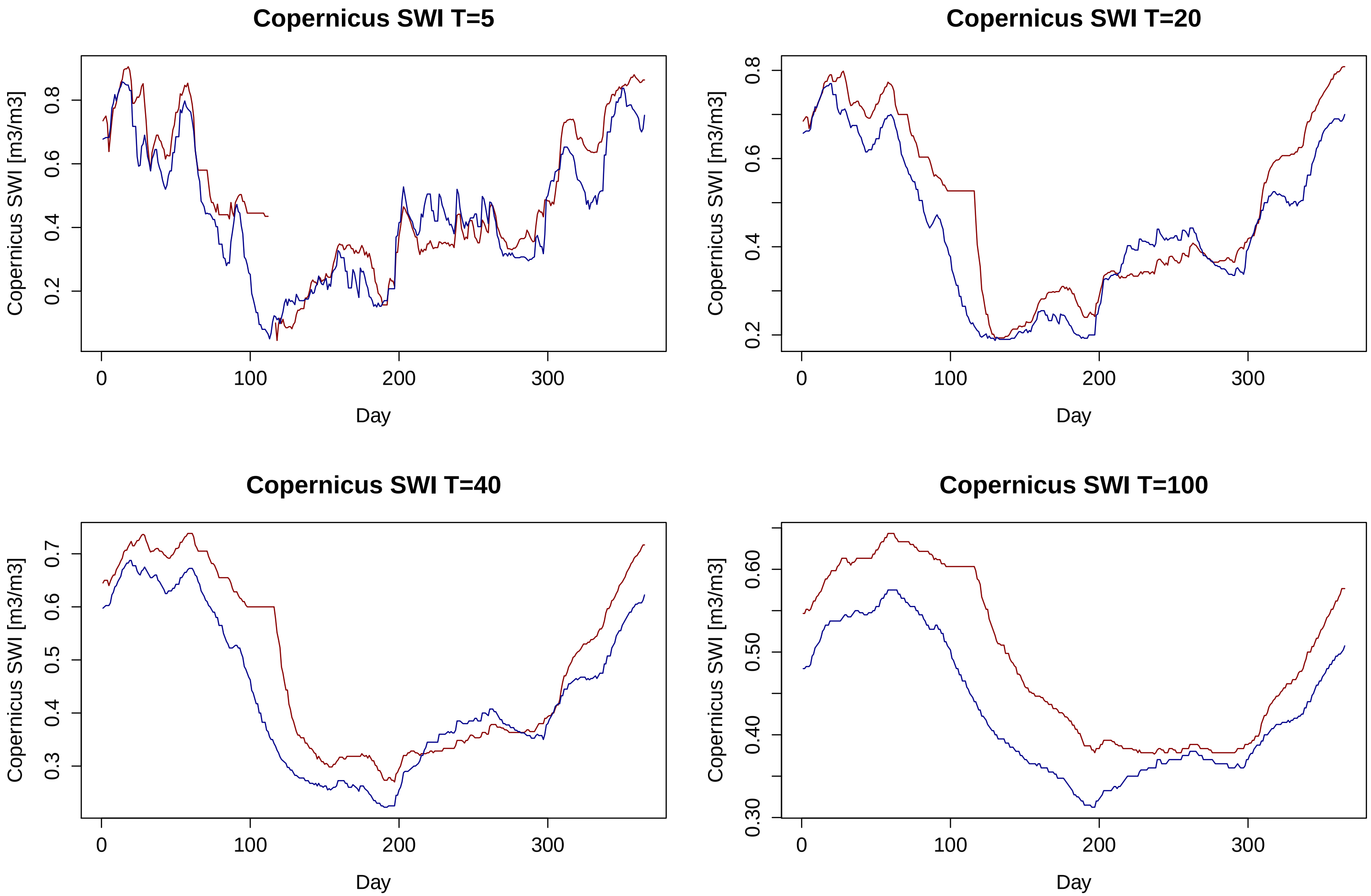


Figure 27: Copernicus SWI estimates with T factors 5, 20, 40 and 100 for the two segments 9564959 (Centre, red color) and 9525076 (Poitou-Charente, blue color) for the 365 days during the year 2011

D Appendix - Correlation and rmse for different T factors in segment 9564959 and 9525076

Table 11: Correlation and RMSE between E-HYPE SML1, SMRZ, SRFP and Copernicus SWI 001, 005, 010, 020, 040, 060, 100, for segment 9564959 (located in the Centre region)

E-HYPE variable	T factor	Pearson correlation	RMSE
SML1	1	0.837	NA
SML1	5	0.856	NA
SML1	10	0.837	NA
SML1	15	0.808	NA
SML1	20	0.770	NA
SML1	40	0.592	NA
SML1	60	0.444	NA
SML1	100	0.265	NA
SMRZ	1	0.679	NA
SMRZ	5	0.748	NA
SMRZ	10	0.803	NA
SMRZ	15	0.841	NA
SMRZ	20	0.868	NA
SMRZ	40	0.903	NA
SMRZ	60	0.882	NA
SMRZ	100	0.816	NA
SRFP	1	0.699	0.192
SRFP	5	0.767	0.168
SRFP	10	0.818	0.153
SRFP	15	0.852	0.141
SRFP	20	0.875	0.131
SRFP	40	0.899	0.102
SRFP	60	0.872	0.086
SRFP	100	0.799	0.075

Table 12: Correlation and RMSE between E-HYPE SML1, SMRZ, SRFP and Copernicus SWI 001, 005, 010, 020, 040, 060, 100, for segment 9525076 (located in the Poitou-Charente region)

E-HYPE variable	T factor	Pearson correlation	RMSE
SML1	1	0.840	NA
SML1	5	0.823	NA
SML1	10	0.779	NA
SML1	15	0.745	NA
SML1	20	0.717	NA
SML1	40	0.622	NA
SML1	60	0.542	NA
SML1	100	0.428	NA
SMRZ	1	0.819	NA
SMRZ	5	0.877	NA
SMRZ	10	0.899	NA
SMRZ	15	0.908	NA
SMRZ	20	0.911	NA
SMRZ	40	0.884	NA
SMRZ	60	0.833	NA
SMRZ	100	0.739	NA
SRFP	1	0.820	0.190
SRFP	5	0.877	0.164
SRFP	10	0.899	0.151
SRFP	15	0.908	0.142
SRFP	20	0.911	0.136
SRFP	40	0.884	0.114
SRFP	60	0.833	0.097
SRFP	100	0.738	0.082

E Appendix - z-scores for the E-HYPE SMRZ and the ESA CCI SM product years 2002, 2003 and 2011

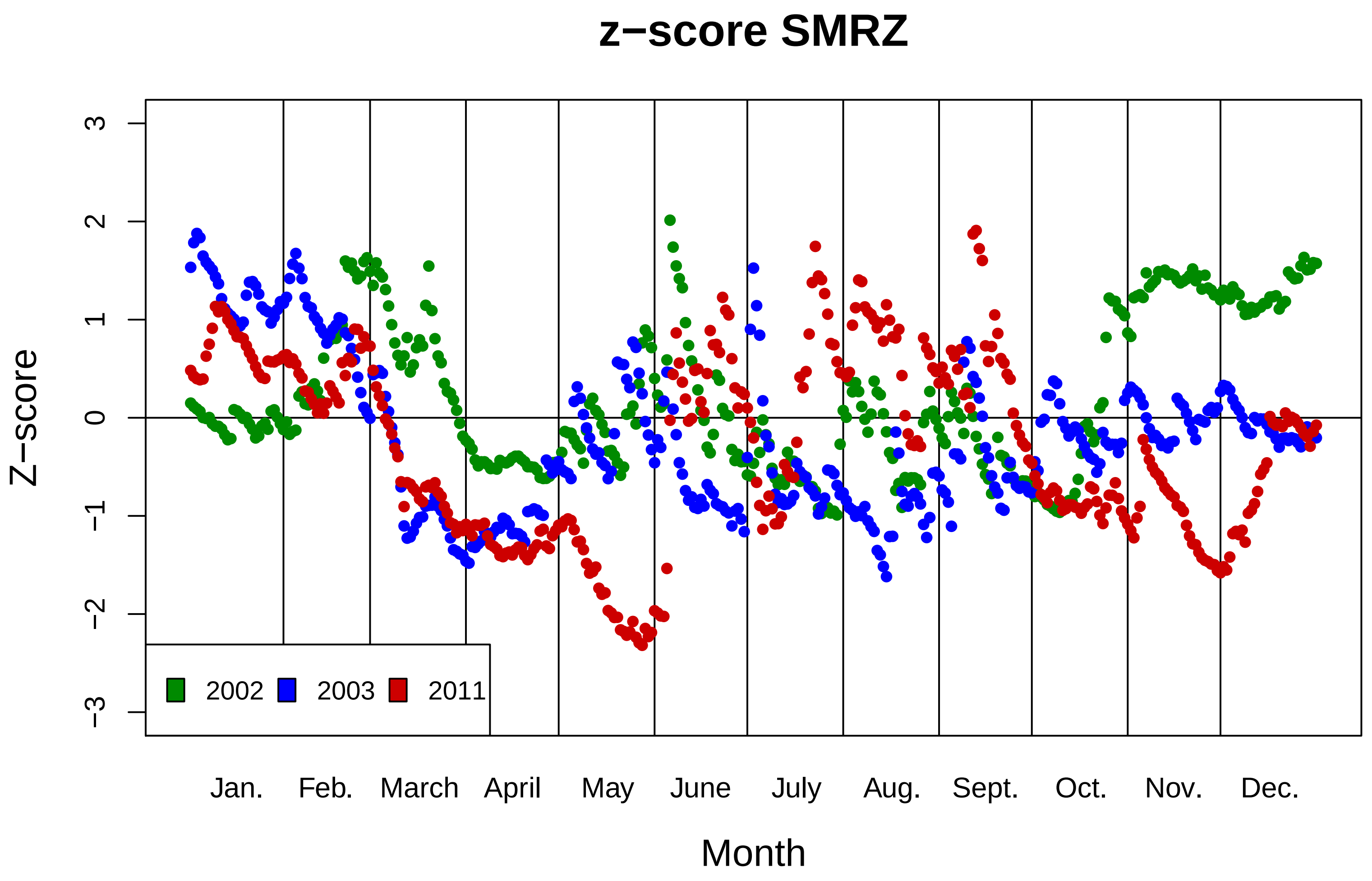


Figure 28: z-score of the E-HYPE SMRZ variable in segment 9525076 (located in the Centre region) for the 365 studied days years 2002, 2003 and 2011. Vertical lines separates the individual months and the horizontal line show the average soil moisture estimate.

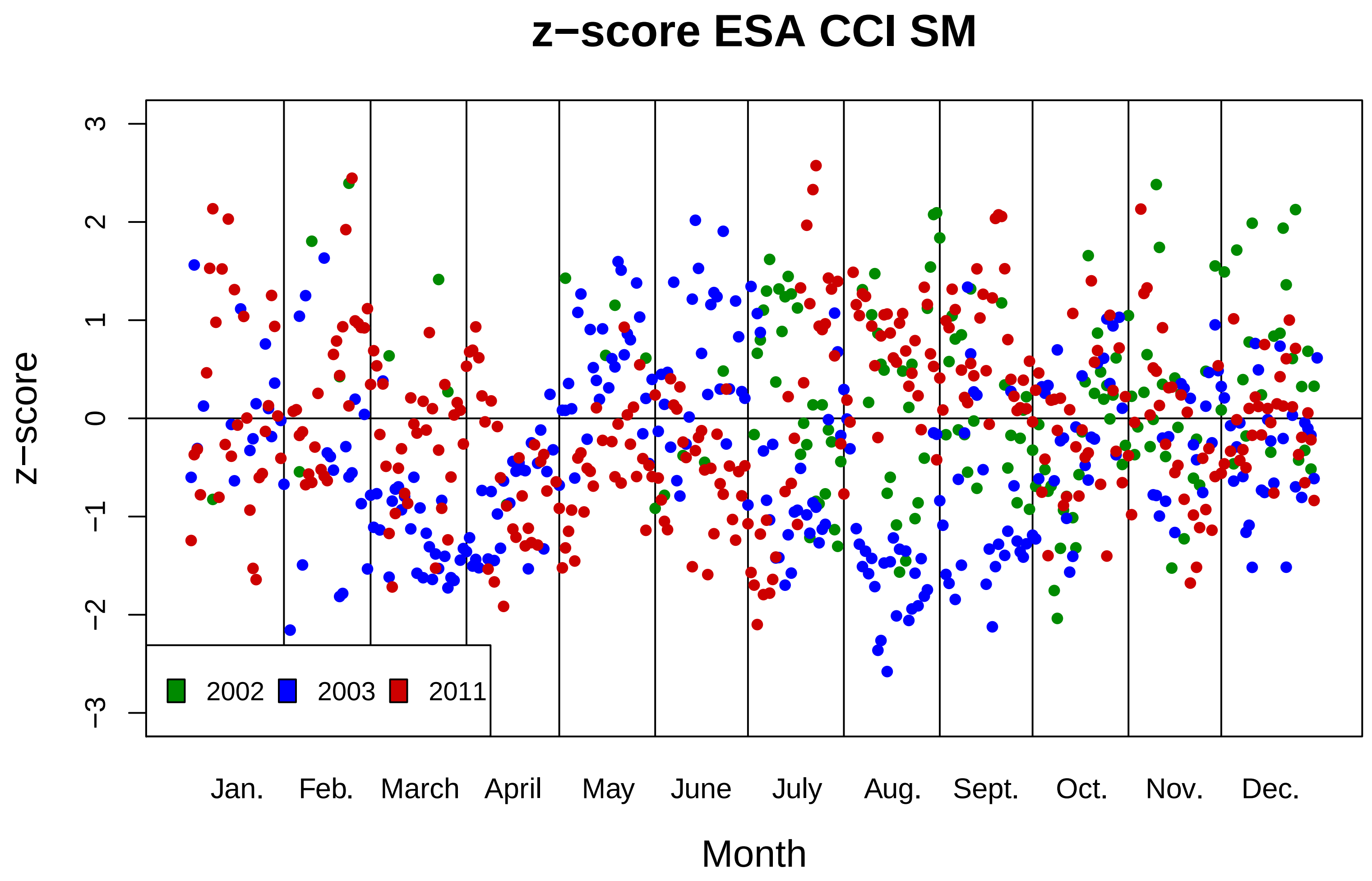


Figure 29: z-score of the ESA CCI SM product in segment 9525076 (located in the Centre region) for the 365 studied days years 2002, 2003 and 2011. Vertical lines separates the individual months and the horizontal line show the average soil moisture estimate.

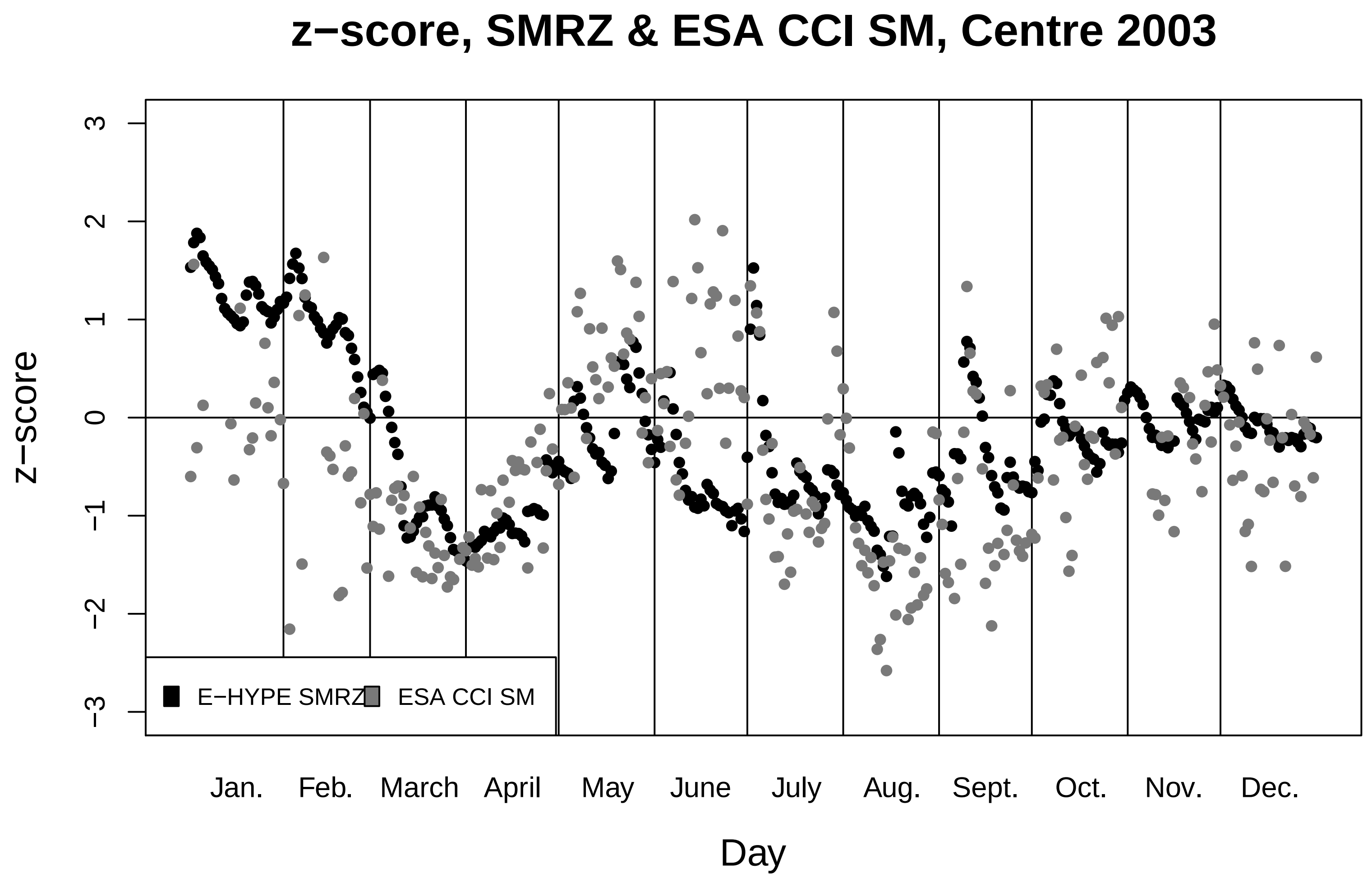


Figure 30: z-score of the E-HYPE SMRZ and the ESA CCI SM product in segment 9525076 (located in the Centre region) for the 365 studied days year 2003. Vertical lines separates the individual months and the horizontal line show the average soil moisture estimate.

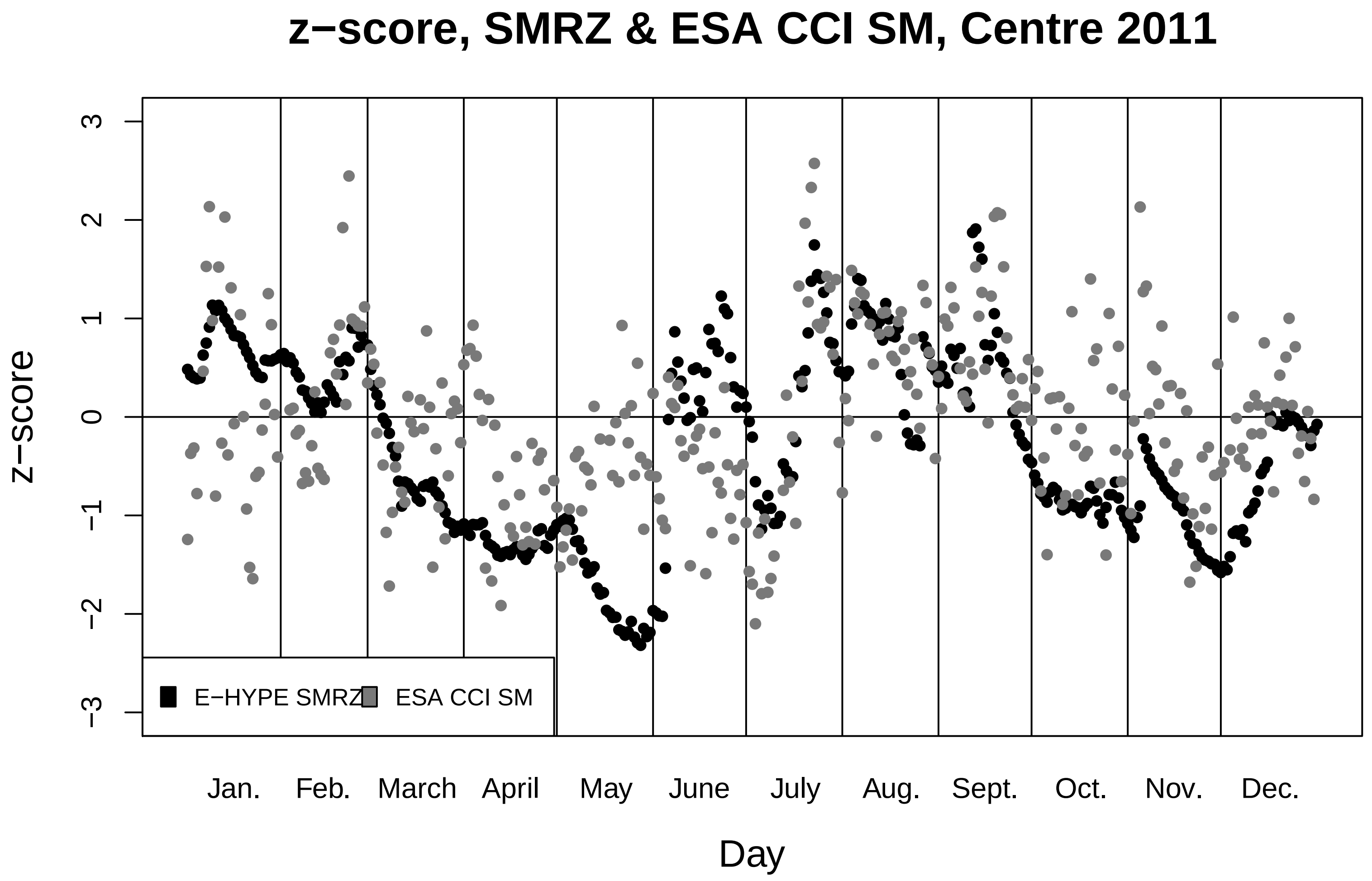


Figure 31: z-score of the E-HYPE SMRZ and the ESA CCI SM product in segment 9525076 (located in the Centre region) for the 365 studied days year 2011. Vertical lines separates the individual months and the horizontal line show the average soil moisture estimate.