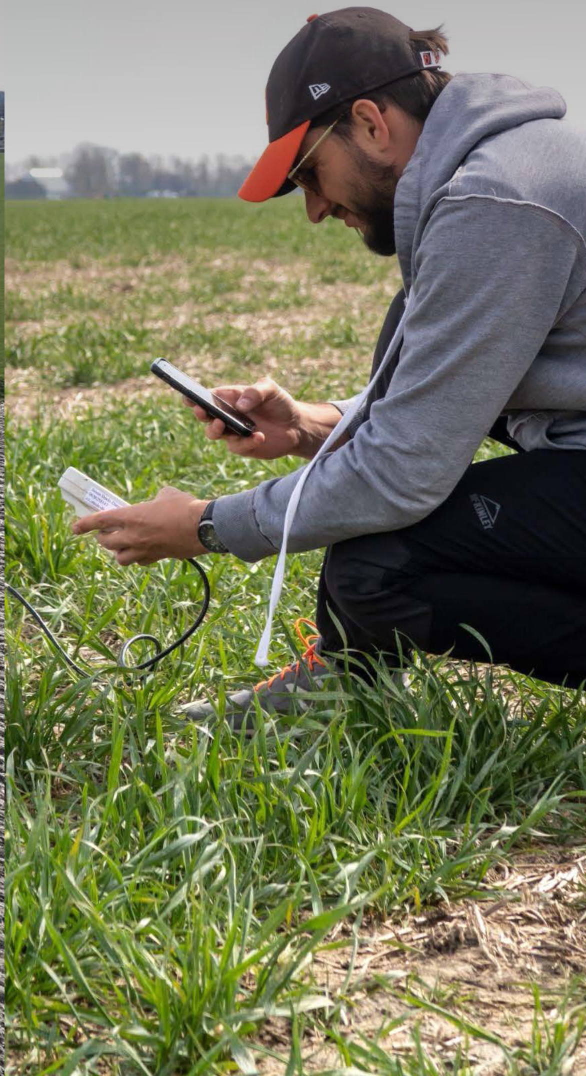


PLANETARY VARIABLES: SOIL WATER CONTENT (SWC)

SPATIAL AND TEMPORAL VALIDATION

Authors:
Anne De Nijs
Richard De Jeu

PLANET.COM

TABLE OF CONTENTS

Contributors:
Stephen Levay, Robert Simmon, Owen Troy, Benjamin Aouizerats, Yoann Malbeteau, Arjan Geers, Jaap Schellekens, Amber Mulder, Stephanie Gijsbers, Thomas Frederikse, Michel Van Klink, Sorosh Shoaie, Tessa Kramer, Mendy Van Der Vliet, Teije Van Der Horst, Joost Beckers, Filipe Catarino, Greg Robson, Jasper Dijkstra, Margot Ridderikhoff, Ashfak Mahmud

GENERAL INTRODUCTION

Determining the quality of a Planetary Variable like the Soil Water Content (SWC) is an important and challenging task. Like any satellite observed data product, the SWC products have uncertainties, but it is not always easy to quantify this uncertainty. Validation with locations where SWC values are known can help to determine the right quality metrics. By quantifying the reliability of SWC values over certain regions, the SWC products will prove more valuable for a range of applications.

Over the years, multiple validation methods for satellite observed SWC have been developed by the scientific community. Gruber et al., 2020[1] provide a good overview of these activities. Validation with calibrated ground sensor networks is considered to be the most common approach, but it also has its limitations. Observations from ground sensors are often only representative for a small area around the device, where a satellite-derived SWC value represents a whole field (depending on the pixel size). Moreover, the sensing depth is also different; a ground sensor is installed at a certain depth and represents the SWC at this particular depth. Satellite-derived SWC represents the first centimeters of the soil including the surface. Lastly, ground sensors always need to be calibrated (see Appendix A for a more detailed description) which creates an additional uncertainty that is not always considered. Due to these reasons, comparisons between satellite-derived SWC and ground stations do not always reveal high correlations.

But under the right conditions, validation with independent ground sensors can provide great insights and can be used to evaluate the quality of the Planet SWC products, which is the goal of this report.

The evaluation is done in two parts:

1) Part 1 focuses on the spatial evaluation of the SWC. It contains the results of fieldwork that was completed in April 2022. A 2-day field campaign was conducted in the Netherlands to get more insights in the spatial quality of both ground sensors and our satellite-derived SWC.

2) Part 2 contains a temporal validation study where we compare time series of our SWC products to several ground sensors on different places on Earth (the Netherlands, Spain and California).

[1] Gruber, A., et al. (2020). Validation practices for satellite soil moisture retrievals: What are (the) errors?. Remote sensing of environment, 244, 111806.

PART 1
SATELLITE-DERIVED VERSUS IN-SITU MEASUREMENTS: WHAT IS GROUND TRUTH?

INTRODUCTION

In April 2022, Planet conducted a field campaign to spatially validate our products, for which no useful ground data for this purpose was available. Additionally, Planet wanted to get more insight into the difference between ground and satellite data and the 'truthfulness' of ground observations. The main goal of the fieldwork was to compare spatial patterns of satellite SWC with in-situ SWC, with specific aims to discover how many ground measurements are needed to be within the same accuracy as satellite observations and to evaluate the spatial quality.

The field campaign was divided into two components. A sub-field experiment with many ground samples was performed to gain more insight into the spatial variability and representation of ground measurements as compared to satellite observations. This was done on an agricultural field in Warmenhuizen, the Netherlands. In addition, a multi-field scale experiment around Haarlem, the Netherlands, was done to assess the spatial variability of SWC.

DATA AND METHODS

In-situ Equipment

Ground measurements were taken with six hand-held FieldScout time-domain reflector (TDR) sensors with 7.5 cm pins, one Decagon GS3 sensor with 5 cm pins, and two Hydra Probes with 5 cm pins. To obtain consistent measurements using multiple sensors, all sensors were calibrated in the lab with gravimetric soil samples. A short description of the sensors, the calibration methodology and derived equations are described in Appendix A.

Satellite Data

The ground measurements were compared with the Planetary Variable L-band 100 m SWC product. This 100 m resolution volumetric SWC product is derived from JAXA's AMSR2 and NASA's SMAP passive microwave brightness temperatures in combination with observations from the infrared bands of Sentinel-2. The overpass time of SMAP is 6:00 a.m. and the sensing depth is ±5 cm. More information about this product can be found here. Especially for this study, we also produced a downscaled version at 10 m resolution. We made use of a similar approach as the 100 m product but with an extra constraint to resolve the data to 10 m. This process is computationally expensive but necessary to make an adequate spatial comparison with the ground sensors and to reduce the scale mismatch between the L-band SWC spatial resolution and the in-situ measurements.

Multi-field scale experiment site measurements

The locations of the multi-field scale experiment were decided on the morning that the experiment took place (12 April 2022). This was done based on the latest SWC map of the Haarlem area. In total, 14 fields were pre-selected to visit by bicycle. The selection consisted of several fields with high SWC values and fields with low SWC values. Ground measurements with a TDR sensor were taken where possible, whereby three measurements were taken within an area of 10 x 10 cm point and averaged afterwards.

Sub-field experiment site measurements

The sub-field experiment was done on an agricultural field in Warmenhuizen, the Netherlands (lat: 52.726, lon: 4.727; Fig. 1). It is mostly used for growing pumpkins and sunflowers. On the day of the fieldwork (April 11, 2022), the field was recently plowed. Some parts were covered with grass and the rest of the field was bare.

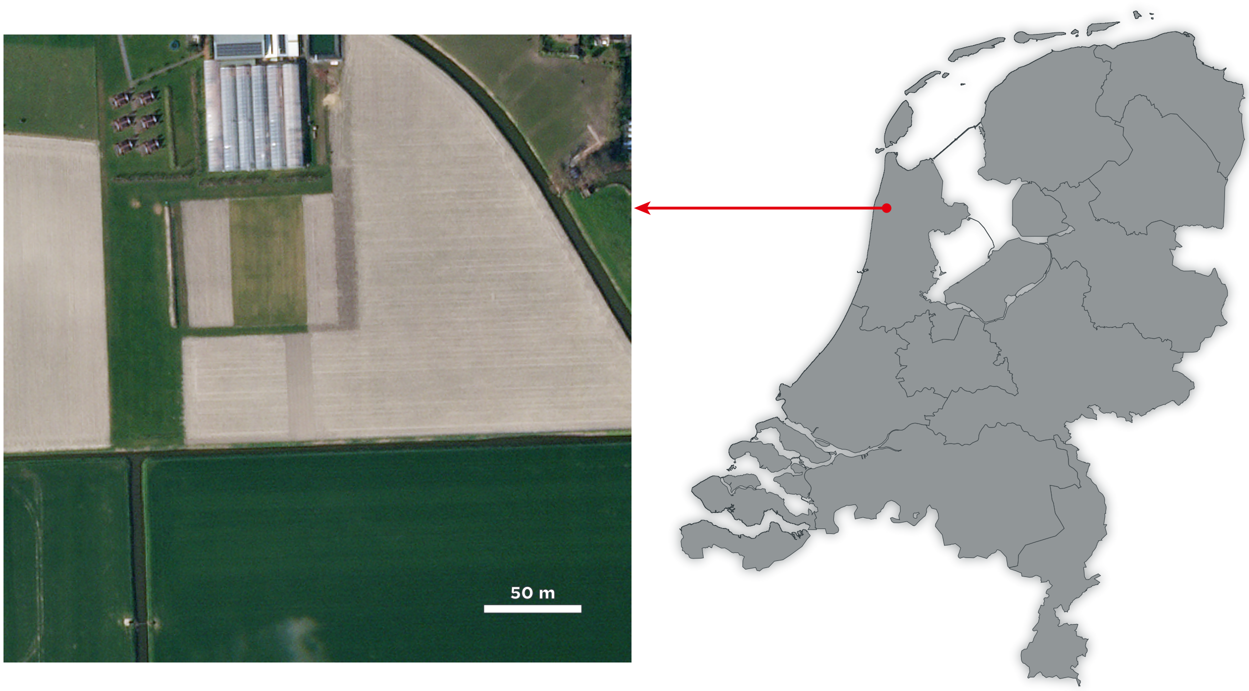


Figure 1: Location of the sub-field experiment; The Marlequi, Warmenhuizen, The Netherlands (lat: 52.726, lon: 4.727); Skysat image of 13 April 2021. Source shapefile of The Netherlands: Jan-Willem van Aalst — www.imergis.nl

Using GPS, measurement tape, and ropes, we created a grid in the field that corresponds to the boundaries of two 100 m SWC pixels (Fig. 2, yellow and orange boxes). Subgrids of 10 x 10 m were made within the grids. In six groups of two persons, in-situ SWC measurements were taken using the same type TDR sensor (TDR100).

Each group started in a subgrid at the North of the grid and worked in a line towards the South. The measurements were taken randomly within the subgrid (Fig. 2, white dots). In Pixel 1, we took five measurements per subgrid and in Pixel 2, we took ten measurements per subgrid. Additional measurements were taken with the GS3 and Hydro Probe sensors. In total, 1140 measurements were collected. With a drone, we made aerial photos of the full grid. These photos were used for georeferencing, so we could obtain the exact geolocation of the grid.

Figure 2: Pixel 1 (yellow) and 2 (orange) with subgrids and example of random sampling within the subgrids (white dots).

RESULTS

Regional analysis

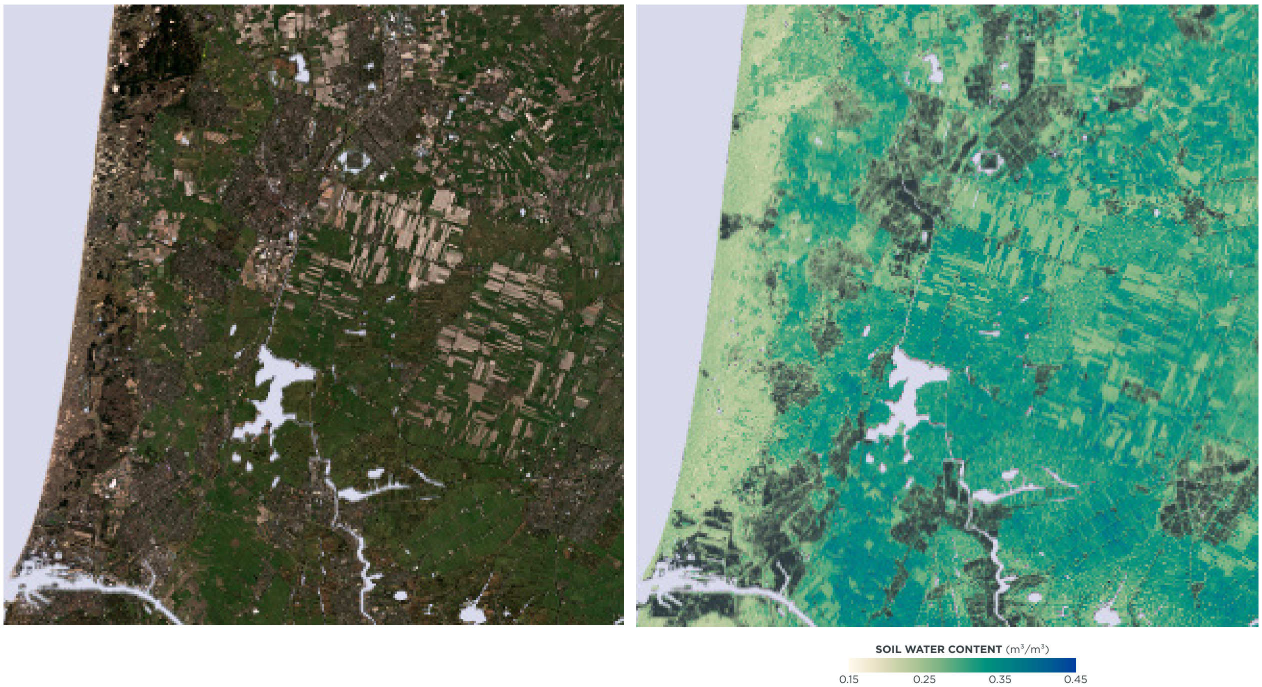


Figure 3: The Netherlands, Planet Monthly Basemap March, 2022 (left), 100 m L-band SWC on 22 March 2022 (right)

Figure 3 shows the 100 m L-band SWC product over a part of the Netherlands. Significant variations in SWC can be seen. On the west side of the image, the dune-area between Bergen aan Zee and Wijk aan Zee is clearly visible. As expected, SWC in this sandy-soil area is low. The wet parts are mainly agricultural fields. Differences between the fields are also visible. Such differences can also be seen in the example of Nebraska (Fig. 4), where the irrigation pivots stand out. In the example of Rwanda and Burundi (Fig. 5) the wet river beds are nicely shown. In the next section ('Multi-field scale experiment'), we will focus on the differences between several fields and compare in-situ measurements with satellite based SWC.

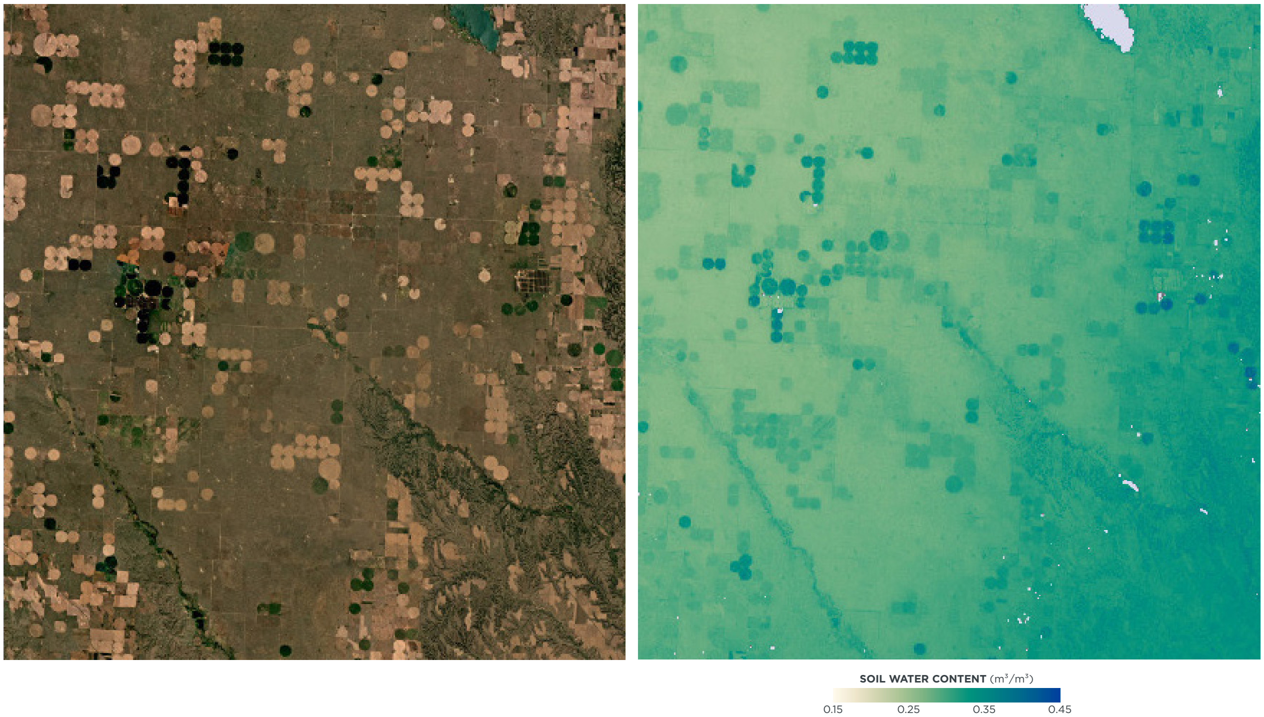


Figure 4: Southwestern Nebraska, Planet Monthly Basemap May 2022 (left) and 100 m L-band SWC of 11 May 2022 (right)

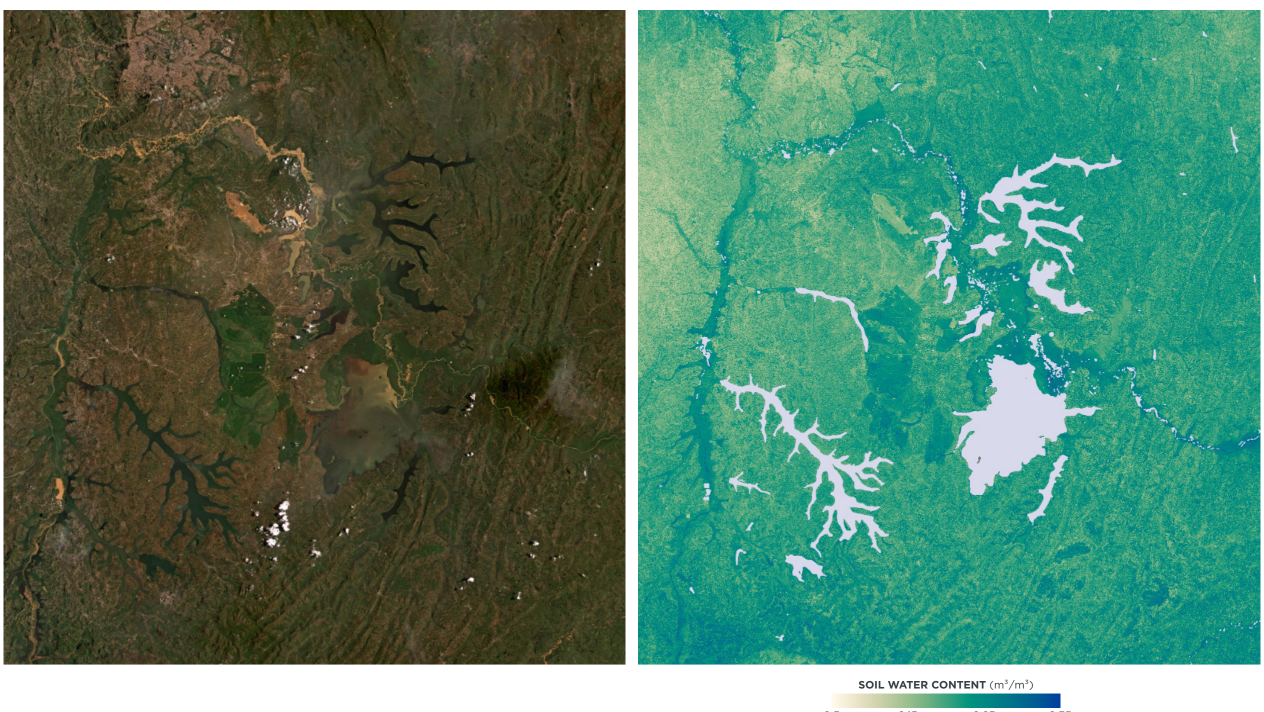


Figure 5: Part of Rwanda and Burundi, Planet Monthly Basemap March 2022 (left) and 100 m L-band SWC of 25 March 2022 (right)

Multi-field scale experiment

Only five of the 14 pre-selected fields were suitable and accessible for taking in-situ measurements: two were located in Vijfhuizen (fields 1 and 2) and three in Hoofddorp (fields 9, 10, and 'extra 10'), close to Schiphol Airport Amsterdam (Fig. 6). According to the field data of the Dutch Boer&Bunder[2] database; fields 1, 2, and 10 were used in 2022 for growing wheat, field 9 for pasture and 'Extra (10)' for Japanese oats.

Figure 6: Location of the multi-field scale experiment; close to Schiphol airport.
Left: 10 m L-band ER SWC of 11 April 2022, including TDR measurement points with field number.
Source shapefile of The Netherlands: Jan-Willem van Aalst — www.imergis.nl

In fields 1, 2 and 'Extra (10)', measurements were taken at a single point, while fields 9 and 10 were sampled at three different points. The right image of Figure 6 shows L-band SWC at 10 m resolution including the measurement points, which are labeled with the (pre-selected) field number. The field-averaged 10 m L-band SWC value was calculated for each field. The dotted lines in Figure 6 show the boundaries of each field. Figure 7 shows the TDR measurements against the mean satellite SWC values of the fields (R=0.79). The amount of points for this comparison is limited, but it can be seen that we found a strong relationship in the same direction between the satellite-derived and the in-situ SWC observations.

[2] https://boerenbunder.nl/

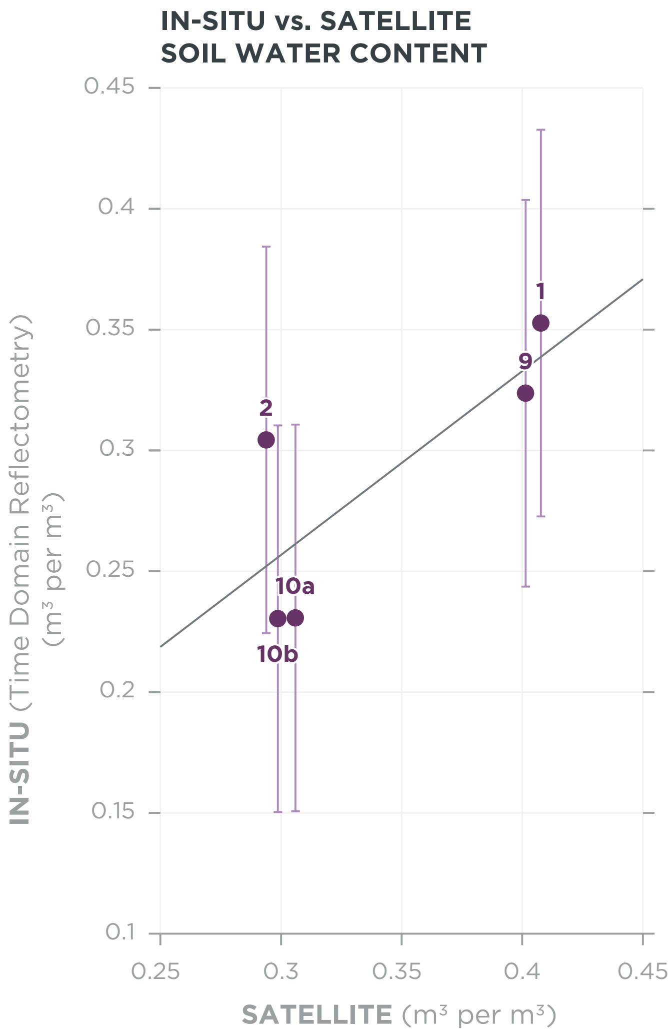


Figure 7: L-band SWC 10 m vs. in-situ TDR SWC, including field numbers

Sub-field experiment

After the fieldwork, the in-situ data was gathered and calibrated. Figure 8 shows maps of the calibrated TDR measurements and the 10 m L-band SWC. Each subgrid of Pixel 1 shows the average of 5 TDR measurements, while each subgrid of Pixel 2 shows the average of 10 TDR measurements. It can be seen that the general spatial patterns in the field correspond very well between the two maps.

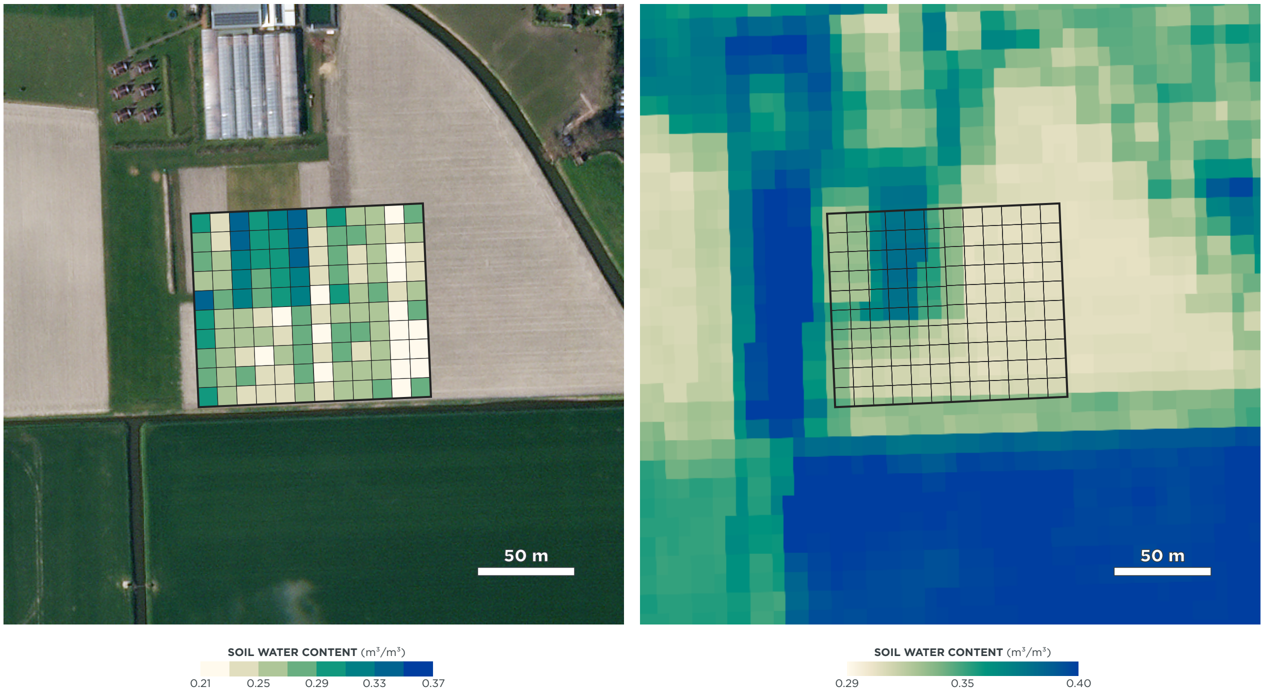


Figure 8: Spatial maps of SWC from TDR sensors(left),
representing more than 1,100 measurements vs. 10 m L-band ER (right).

Uncertainty analysis

One of the goals of the fieldwork was to find out how many ground measurements are needed to achieve the same uncertainty as satellite observations. As described in the SMAP handbook[3], the SMAP baseline requirement for SWC is specified as “estimates of soil moisture in the top 5 cm of soil with an error of no greater than 0.04 m^3/m^3 volumetric.” This number is generally taken as the uncertainty of satellite-derived SWC values. Figure 9 shows the results of the uncertainty analysis of the in-situ measurements within a 100 m pixel. These results were retrieved by averaging an increased number of randomly-selected measurements (sets of 2 measurements, sets of 3 measurements, sets of 4 measurements, etc) and calculating the variability within each set of measurements. For instance, with 10 measurements within our region of interest, 45 sets of 2 measurements are aggregated. The next iteration will have 120 sets of 3 measurements, then 210 sets of 4 measurements. For each iteration, the variability was estimated by computing the difference between the 99.7% percentile (3σ, considering a gaussian distribution) and the average of the given set. This was done for each 100 m pixel separately. Both graphs show that at least 7 measurements are needed to be within the satellite uncertainty. The uncertainty of the in-situ measurements decreases with the amount of measurements, and around 15 measurements within one pixel are needed to reach the accuracy of the sensor (0.02 m^3/m^3).

[3] Entekhabi, D., et al. (2014) SMAP Handbook Soil Moisture Active Passive, Mapping Soil Moisture Freeze/Thaw from Space

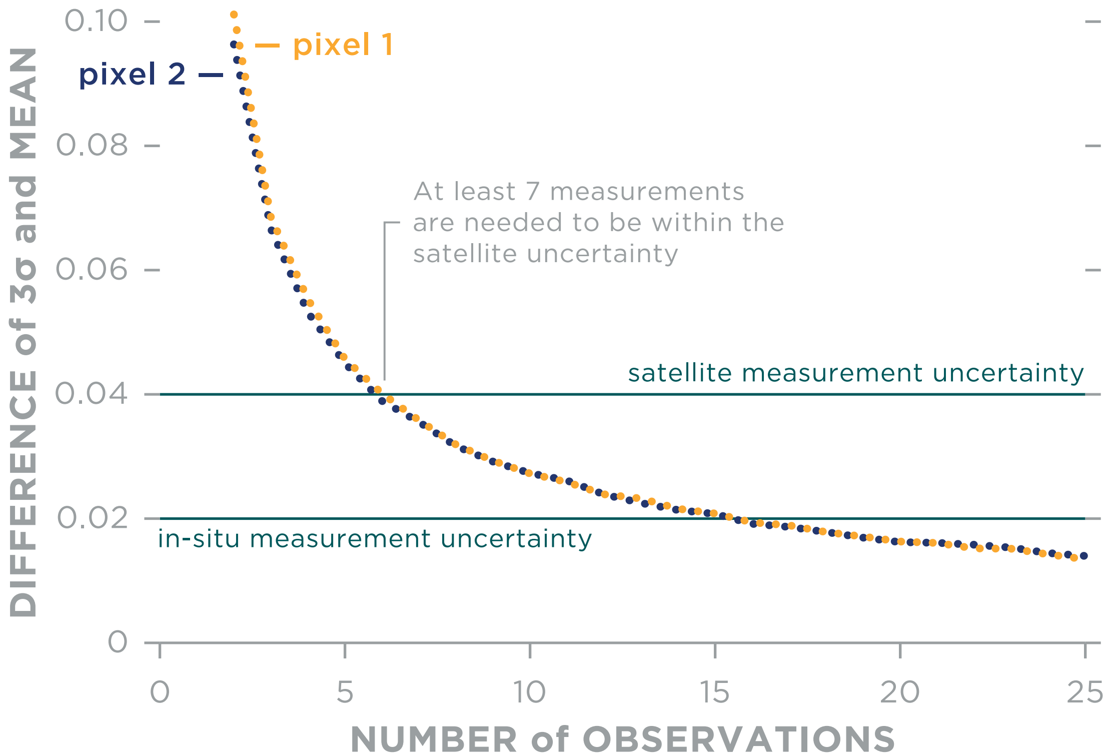


Figure 9: Difference between 3σ and mean of the observation pairs with a certain number of in-situ observations

DISCUSSION AND CONCLUSIONS

Aside from several studies with airplane data and field campaigns[4,5], research on the spatial validation of satellite-derived SWC products is sparse. The need for a validation strategy for downscaled soil moisture is also stated in Peng et al. (2017)[6]. For spatial validation of our satellite SWC products, a spatial validation field campaign was set up in the Netherlands. The goal was to get more insight into the quality, spatial variability and representation of our data compared to in-situ data.

---

[4] Panciera, R., et al. (2008). The NAFE'05/CoSMOS data set: Toward SMOS soil moisture retrieval, downscaling, and assimilation. IEEE Transactions on Geoscience and Remote Sensing, 46(3), 736-745.
[5] Ye, N., et al. (2020). The soil moisture active passive experiments: Validation of the SMAP products in Australia. IEEE Transactions on Geoscience and Remote Sensing, 59(4), 2922-2939.
[6] Peng, J., Loew, A., Merlin, O., & Verhoest, N. E. (2017). A review of spatial downscaling of satellite remotely sensed soil moisture. Reviews of Geophysics, 55(2), 341-366.

This 2-day exercise gave us multiple valuable insights. The in-situ SWC measurements of the multiscale field experiment were in line with the field averages for five agricultural fields. The spatial variability in the most recent satellite SWC map agrees with the ground sensor observations. The extensive sub-field experiment, which consists of more than 1100 measurements in two 100 m pixels, shows good agreement with same-day 10 m satellite SWC observations. The sub-field experiment also gave us a better understanding of uncertainties and the spatial representation of ground measurements compared to satellite observations. At least 7 ground measurements are needed for one 100 m pixel to be within the same uncertainty as satellite data. In practice, this would mean that it takes quite some time and money to achieve the same spatial and temporal quality as satellite observations. Although the few in-situ SWC measurements of the multiscale field experiment were in the right direction, we showed that just putting one or two sensors in a single field is not enough to say something about the field with confidence. The measured in-situ value of a certain point in the field can be quite different when you measure a few meters away. Using ground sensors is therefore labor intensive for this purpose due to the uncertainties and the potential variation within the field.

So, what is ground truth? It depends on the purpose. In-situ measurements are relatively precise since the instrument uncertainty is about 0.02 m^3/m^3. A single measurement is close to the 'truth' for that specific location. However, to obtain the conditions of a larger area or field, it takes a lot of measurements to get an accurate insight. This must be kept in mind when using (single) in-situ sensors as 'the truth' when validating satellite-derived data.

Proper sensor calibration is important, especially when multiple different sensors are used for the same dataset. Therefore, it was interesting to see during the fieldwork that there are multiple ways to simply take measurements. Some people carefully put the TDR sensor in the ground, making sure that the pins were completely in the ground. Others paid a bit less attention to that and obtained measurements more quickly. Furthermore, part of the field was covered with some small hills and gullies, which raised the question of how that variance should be accounted for. This shows that while measuring in the field, people may have different opinions on how to measure in the 'right' way.

As a conclusion, we can say that quantifying the spatial quality of satellite-derived SWC is a difficult and time-consuming task. Ground sensors are a good tool to use for validation, but because of the difference in spatial support there are many in-situ measurements needed to say something about a whole field. Within the scientific community however, there are not many examples of comparable experiments for spatial validation of SWC. We have shown comparable spatial patterns between in-situ and satellite measurements with this short validation measurement. It would be beneficial if similar experiments were performed to get even more insights in the spatial quality and obtain more ground data for validation purposes. We will continue to work on this topic and have already planned a follow-up fieldwork experiment to measure the spatial variability in a dune area. Also, we aim to set up partnerships to perform collaborative research on the spatial validation of SWC.

PART 2
TEMPORAL VALIDATION

INTRODUCTION

In the second part of this white paper we evaluate time series of satellite-derived SWC estimates, using in-situ observations at four locations: two in the Netherlands, one in Spain, and one in the USA. Time series and correlation coefficients are presented to show the performance of our SWC products compared to in-situ data over time. We express SWC in volumetric units, namely m^3/m^3. In this analysis we focus on temporal variations and not on absolute values. As already mentioned in the introduction, there are differences in what exactly is measured by the satellite or ground sensors (for example due to the spatial support and depth) and how well the sensor is calibrated. To assess the temporal variations, the in-situ SWC was scaled using a linear regression toward the satellite-derived SWC values. More about this procedure can be found in Appendix C.

DATA AND METHODS

PWN sensor - the Netherlands

Planet possesses a data logger with five SWC in-situ sensors at various depths in the dunes near Castricum, the Netherlands (52.537 11°N, 4.623 13° E). These dunes are managed by the drinking water company Provinciaal Waterleidingbedrijf Noord-Holland (PWN). The in-situ sensors were placed in April 2018 for validation of our satellite-derived SWC products. The data logger measures every hour and the depths of the sensors vary from surface level to 40 cm depth. For this analysis, we use the sensor that was placed at 5 cm depth.

RAAM network - the Netherlands

The Raam river and its catchment area are located in the Southeast Netherlands. The river is a tributary of the Meuse River. In April 2016, a network of in-situ SWC sensors in this region was established by a consortium of the University of Twente, Wageningen University & Research, and the regional water management authority Aa en Maas. The aim of this network is to collect data for calibration and validation of satellite-derived data, the assessment of land process models, the understanding of processes affected by SWC, and for improving regional water management. The network consists of SWC and soil temperature sensors placed in 14 agricultural fields and 1 natural grass field within the Raam catchment and the Hooge Raam sub-catchment basins.

The Decagon 5TM sensors have been installed at 5, 10, 20, 40, and 80 cm depth. Detailed information of the network and the station can be found in Benninga et al. (2018)[7].

We will also use this network to show how a time series looks when the values of multiple well-performing stations and the corresponding satellite-derived data within one network are averaged and combined into one plot. A study towards the spatial-temporal variability of SWC is described in Mittelbach et al. (2012)[8].

REMEDHUS - Spain

The REMEDHUS network in Spain is located in an area of 35 x 40 km in the central part of the Duero Basin, close to Salamanca. SWC data at 5 cm depth is recorded hourly for 22 stations with HydroProbes (Stevens Water Monitoring System, Inc.). The land is primarily used for agriculture and the soil texture is mainly (loamy) sand[9]. Data from the REMEDHUS network is (among other things) used for calibration, validation campaigns and evaluation of satellite-derived SWC data[10]. More information about the data and network can be found in the referenced papers and the website of the network[11].

USCRN - USA

The U.S. Climate Reference Network (USCRN) is an extensive network for climate monitoring developed by the National Oceanic and Atmospheric Administration (NOAA)[12]. The primary goal of the USCRN is to provide long-term high quality temperature, precipitation, SWC, and soil temperature observations. There are in total 139 USCRN sites (114 in the contiguous U.S., 22 in Alaska, 2 in Hawaii, and 1 in Canada). For measuring SWC, the stations are equipped with instruments to measure at 5, 10, 20, 50, and 100 cm depths when possible. For all depths, hourly averages derived from 5-minute observations are saved to the datalogger. The 5-minute observations are stored for the 5 cm depth as well. For this whitepaper, two sites in California were selected for evaluation: 1) Yosemite Village (37.7592, -119.8208), and 2) Merced (37.2381, -120.8825).

---

[7] Benninga, H. F., et al. (2018), The Raam regional soil moisture monitoring network in the Netherlands. Earth System Science Data,10, 61–79, 2018 https://doi.org/10.5194/essd-10-61-2018

[8] Mittelbach, H., et al. (2012). A new perspective on the spatio-temporal variability of soil moisture: temporal dynamics versus time-invariant contributions. Hydrology and Earth System Sciences, 16(7), 2169-2179.

[9] Sánchez, N. et al. (2020): Spatial averages of in-situ measurements versus remote sensing observations: a soil moisture analysis. Journal of Spatial Science, DOI: 10.1080/14498596.2020.1833769

[10] Sánchez, N. et al. (2012): Validation of the SMOS L2 Soil Moisture Data in the REMEDHUS Network (Spain). IEEE Transactions on Geoscience and Remote Sensing, 50:1602–1611. DOI:10.1109/TGRS.2012.2186971

[11] campus.usal.es/~hidrus/

[12] www.ncei.noaa.gov/access/crn/overview.html

Satellite products

As for the spatial validation, we use our L-band SWC products. In this study we evaluate Planet's SWC 1000 m (V4) and its previous version (V3). Both versions use passive microwave observations from JAXA's AMSR2 and NASA's SMAP satellite. The rationale of showing both versions is to evaluate the improvements of subsequent versions.

Statistics

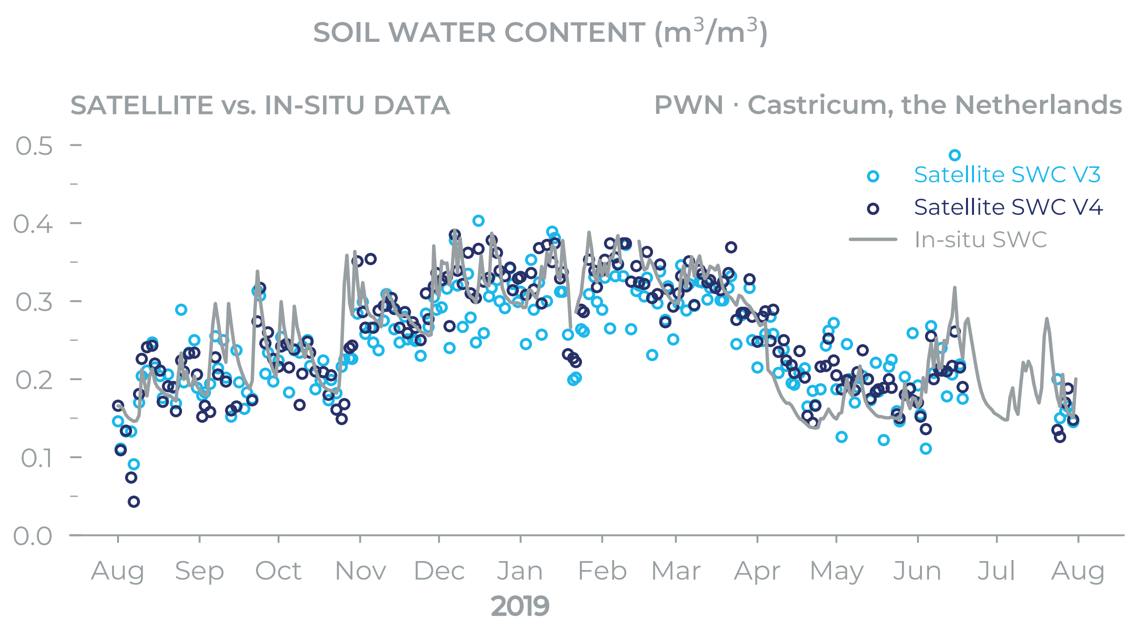
The satellite-derived SWC products are validated against the four in-situ networks using the Pearson's R correlation coefficient. This is a measure of the linear association between two variables. R is +1 or -1 if the dataset from the two variables constitutes exactly a straight line when they are plotted. In this case, the closer to +1, the higher the agreement between both variables.

RESULTS

This section shows the results of the temporal analysis for each network. For the networks with multiple ground sensors, the correlation coefficients are presented in a graph that shows Pearson's R of the L-band SWC version 4 product for each sensor and the difference of R version 3 compared to the R version 4. A positive number in the correlation coefficient plots indicate a higher correlation for version 4 than version 3. Several time series are shown as well. For tables with the correlation coefficients can be found in Appendix B.

PWN sensor - the Netherlands

1 August 2018 - 1 August 2021 (4yr data)



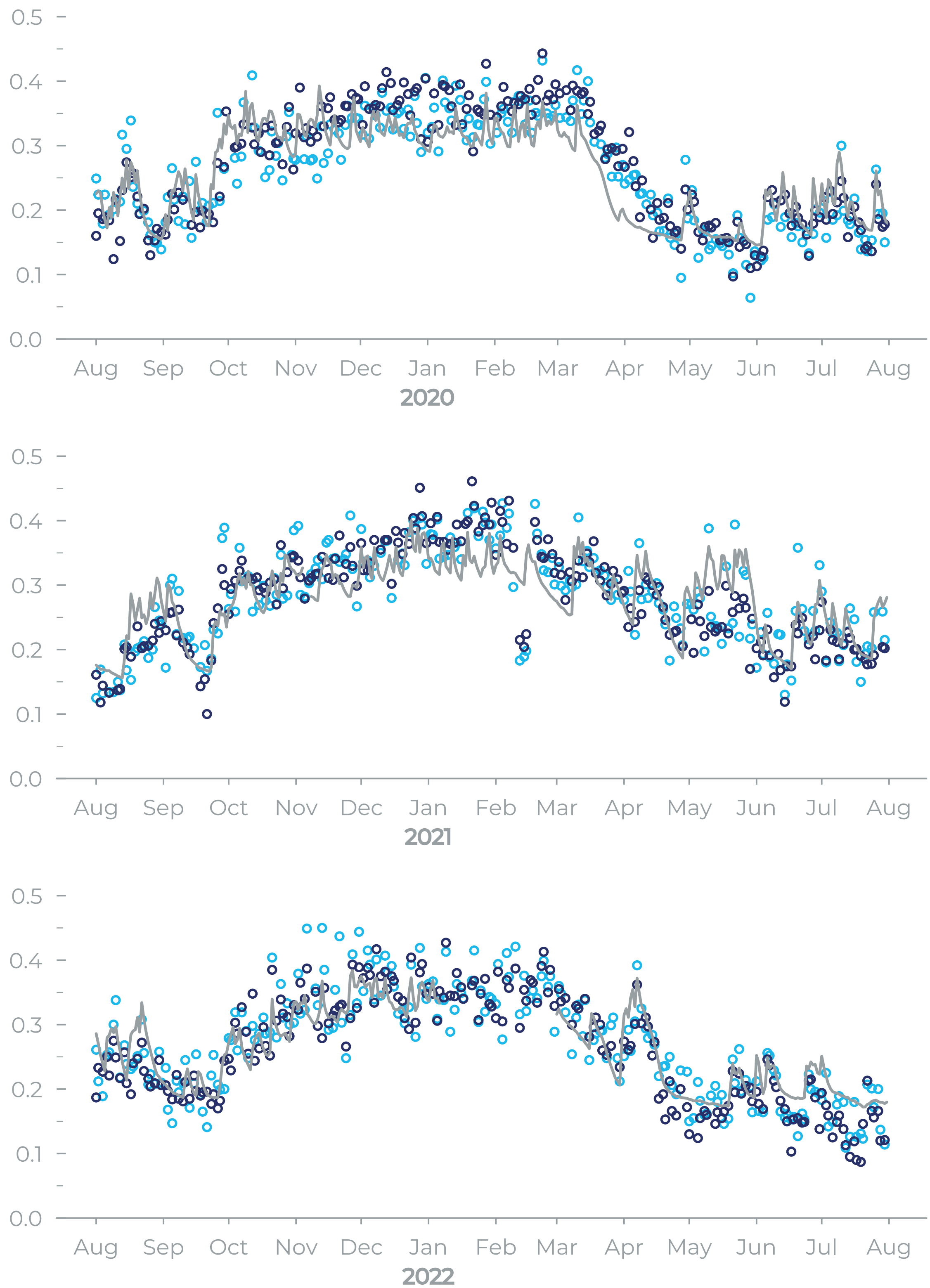


Figure 10: Time series of the PWN sensor in the Netherlands and Planet SWC.

Figure 10 shows the time series of the PWN sensor over a period of 4 years. The satellite-derived SWC products have high correlations with the ground sensor: 0.824 for v3 and 0.853 for v4. Measurements of the ground sensor and the satellite-derived products follow the same pattern and spikes as a result of rain events match quite well.

PWN sensor - the Netherlands, intertemporal analysis

In addition to the previous time series, an extended time series of v4 and the in-situ sensor is shown in the Figure 11 below. The SWC data of both the in-situ and satellite data is normalized in order to have a value between 0 and 1. A threshold value of 0.35 m^3/m^3 was set to indicate 'dry days'. In the bar plot, the percentage of 'dry days' of all the days in a month are shown. Only days when there is both satellite-derived and in-situ data were taken into account. The intensity and duration of dry periods in the different years can be put into perspective. It can be clearly seen that 2021 was less dry than the other years for example.

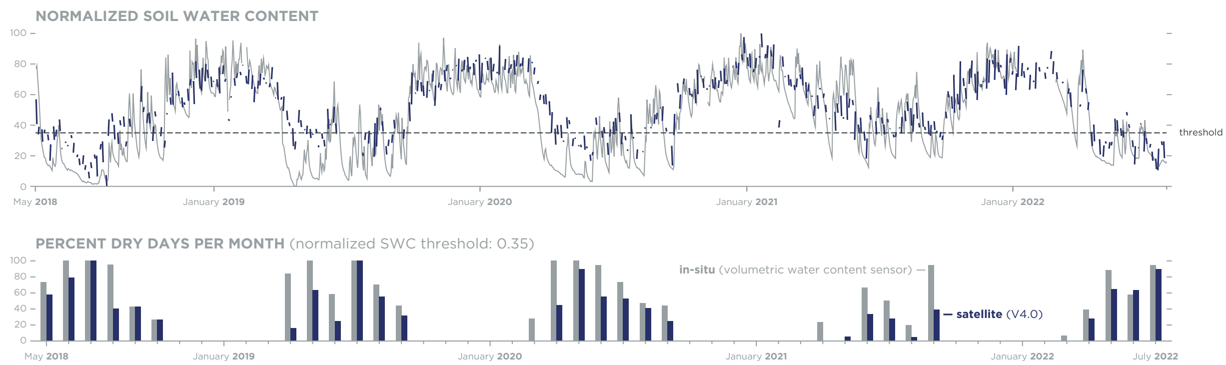


Figure 11:
Upper: Time series of the PWN sensor in the Netherlands and Planet SWC.
Lower: Percent dry days per month.

RAAM network - the Netherlands

1 April 2018 - 1 April 2019 (1yr data)

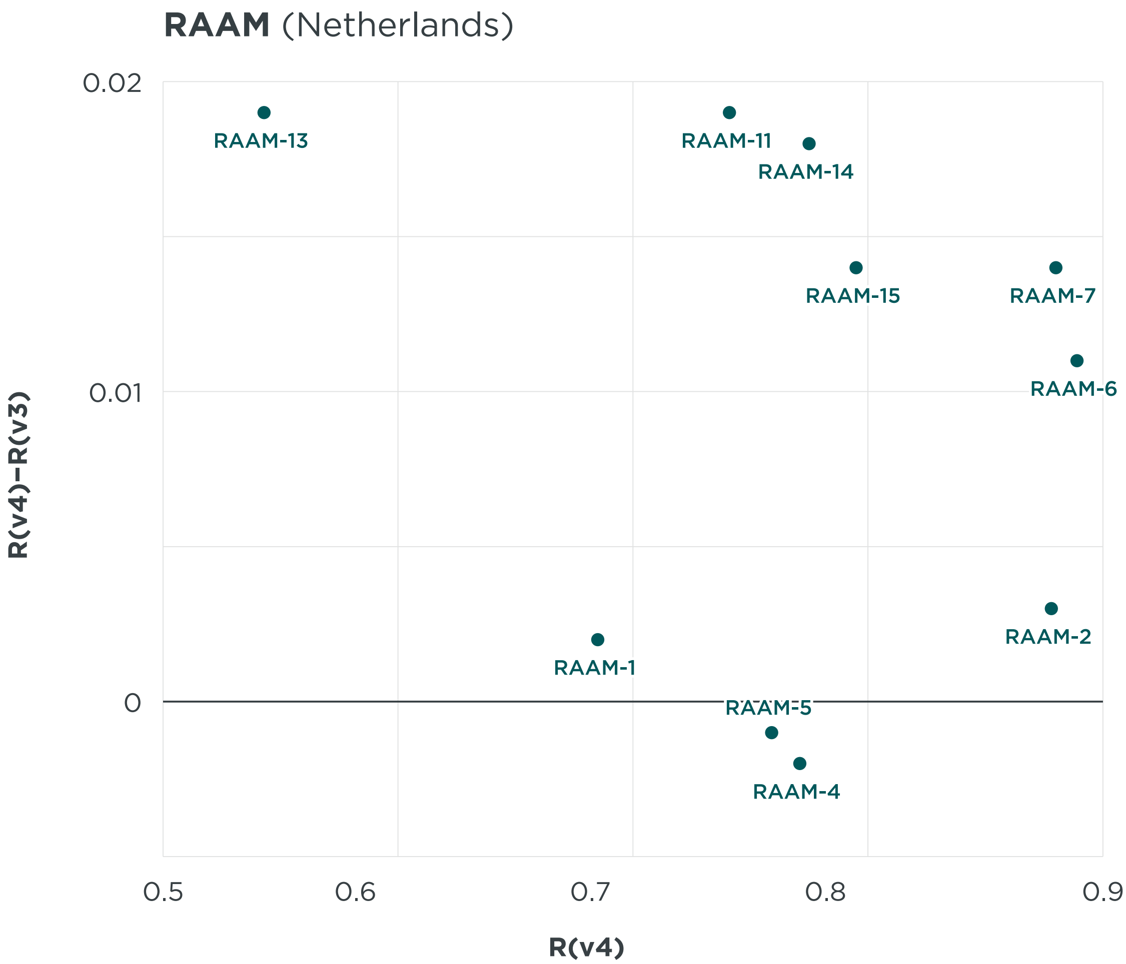


Figure 12: Correlation coefficients of L-band SWC version 4 and the difference with version 3 for each sensor in the RAAM network.

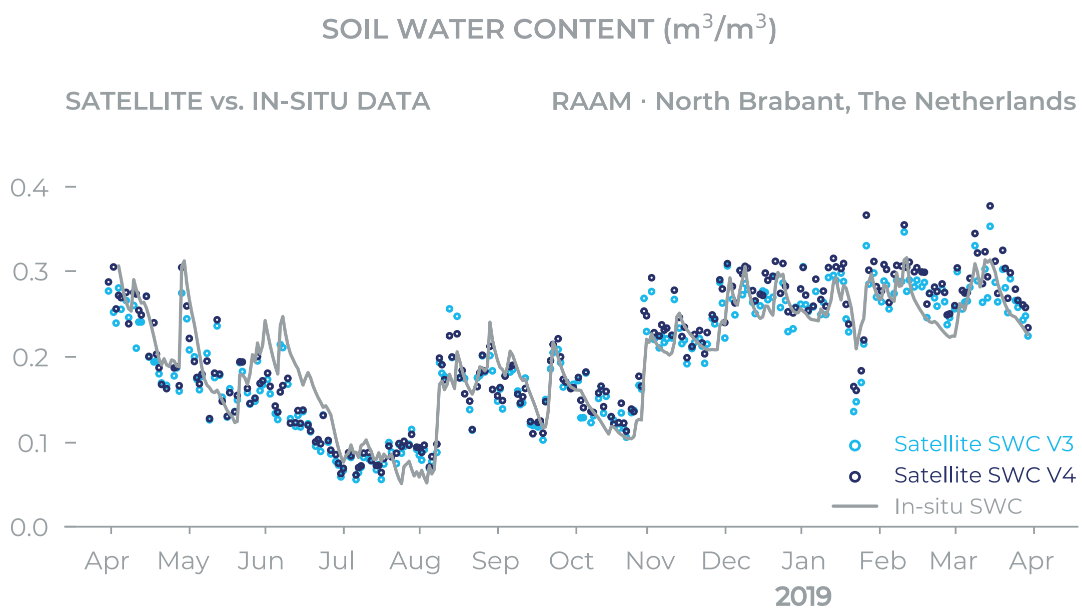


Figure 13: Time series of the average of 9 stations of the RAAM network, the Netherlands.

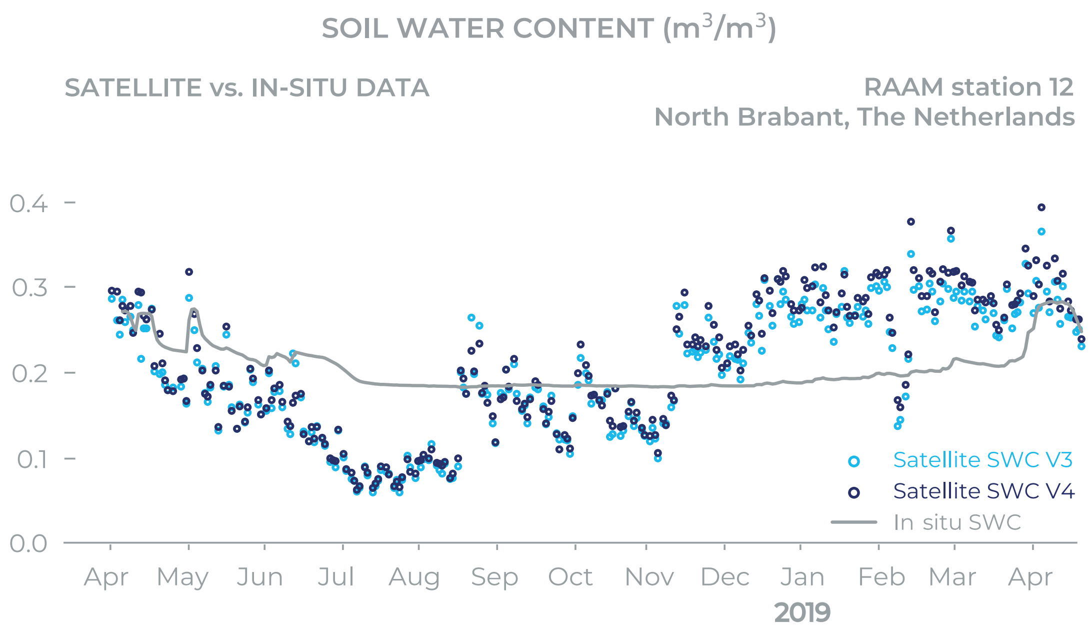


Figure 14: Time series of station 12 of the RAAM network, the Netherlands.

For the RAAM network, data of 1 year is used to create the time series and statistics. Figure 13 shows the time series of the average of 9 well-performing sensors (station 1, 2, 4-7, 11, 14 and 15). The Pearson correlation coefficient for v3 is 0.907 and for v4 it's 0.913. This high correlation shows that combining multiple sensors in a certain area leads to better results. This result is in line with our findings of part 1.

Figure 14 shows an example of a poorly performing station. The low SWC values of the ground sensor in the autumn and winter are not realistic for this location. Including such a sensor in a validation study might influence the outcome of a study in a negative way, and should be avoided. We therefore recommend to thoroughly assess each in-situ dataset that is used in a comparison.

REMEDHUS network - Spain

1 August 2017 - 1 August 2020 (3yr data)

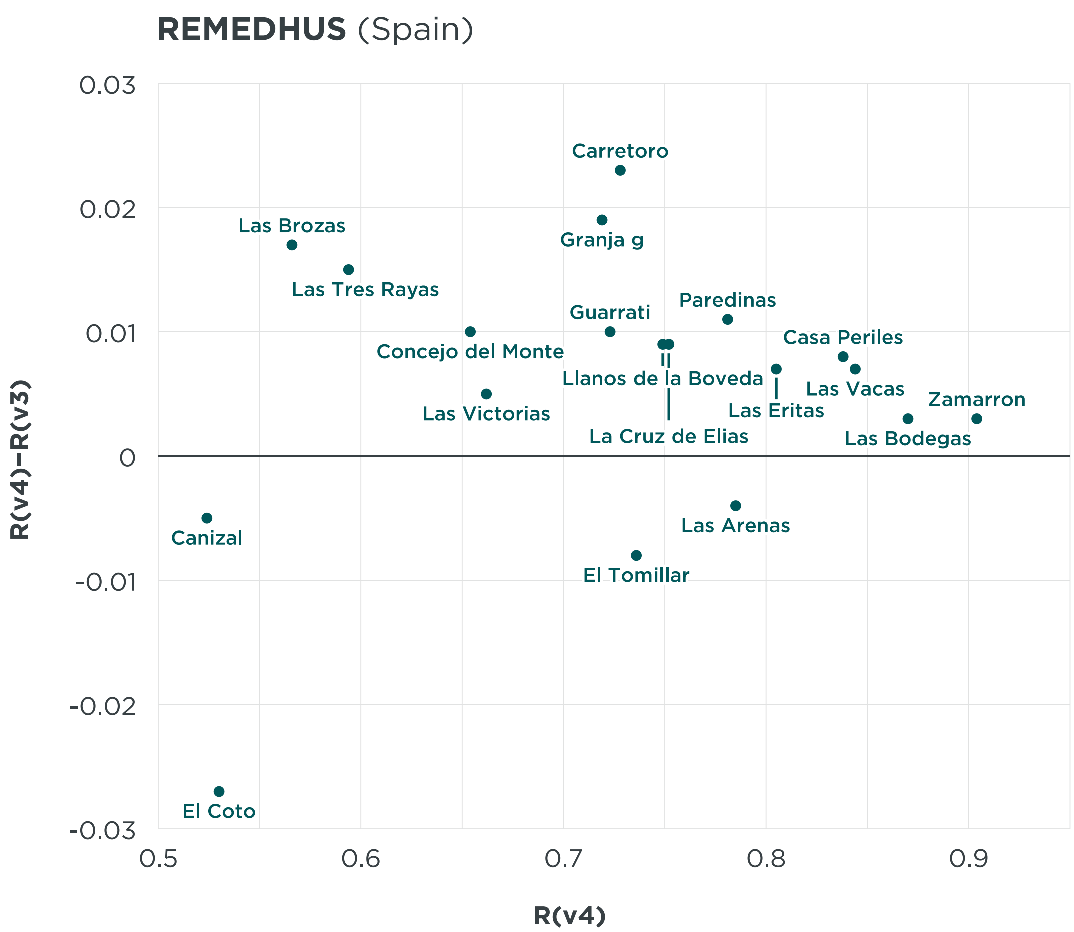


Figure 15: Correlation coefficients of L-band SWC version 4 and the difference with version 3 for each sensor in the REMEDHUS network.

For the analysis with the REMEDHUS network, 3 years of data was used. Figure 15 shows that the correlation coefficients vary between around 0.53 to 0.91. Most values are between 0.7 and 0.8. 15 out of 19 stations have a slightly higher correlation coefficient for L-band SWC version 4 than for version 3. Time series of Las Eritas (R version 4 = 0.804) and Las Vacas (R version 4 = 0.844) are shown in Figures 16 and 17. Both stations have a high correlation and from the time series it can be seen that despite a few spikes, the overall patterns are very similar.

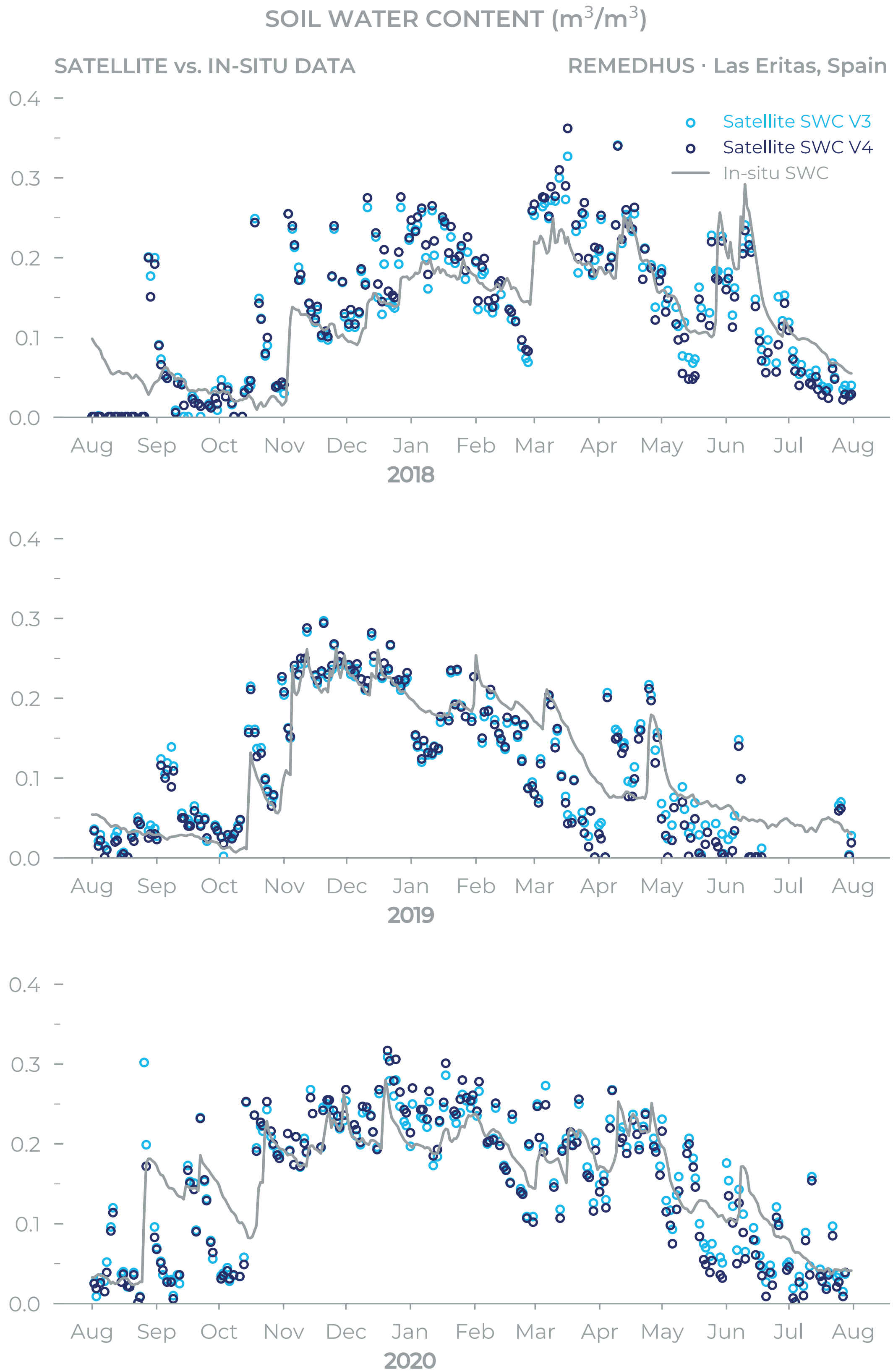


Figure 16: Time series of sensor Las Eritas of the REMEDHUS network, Spain.

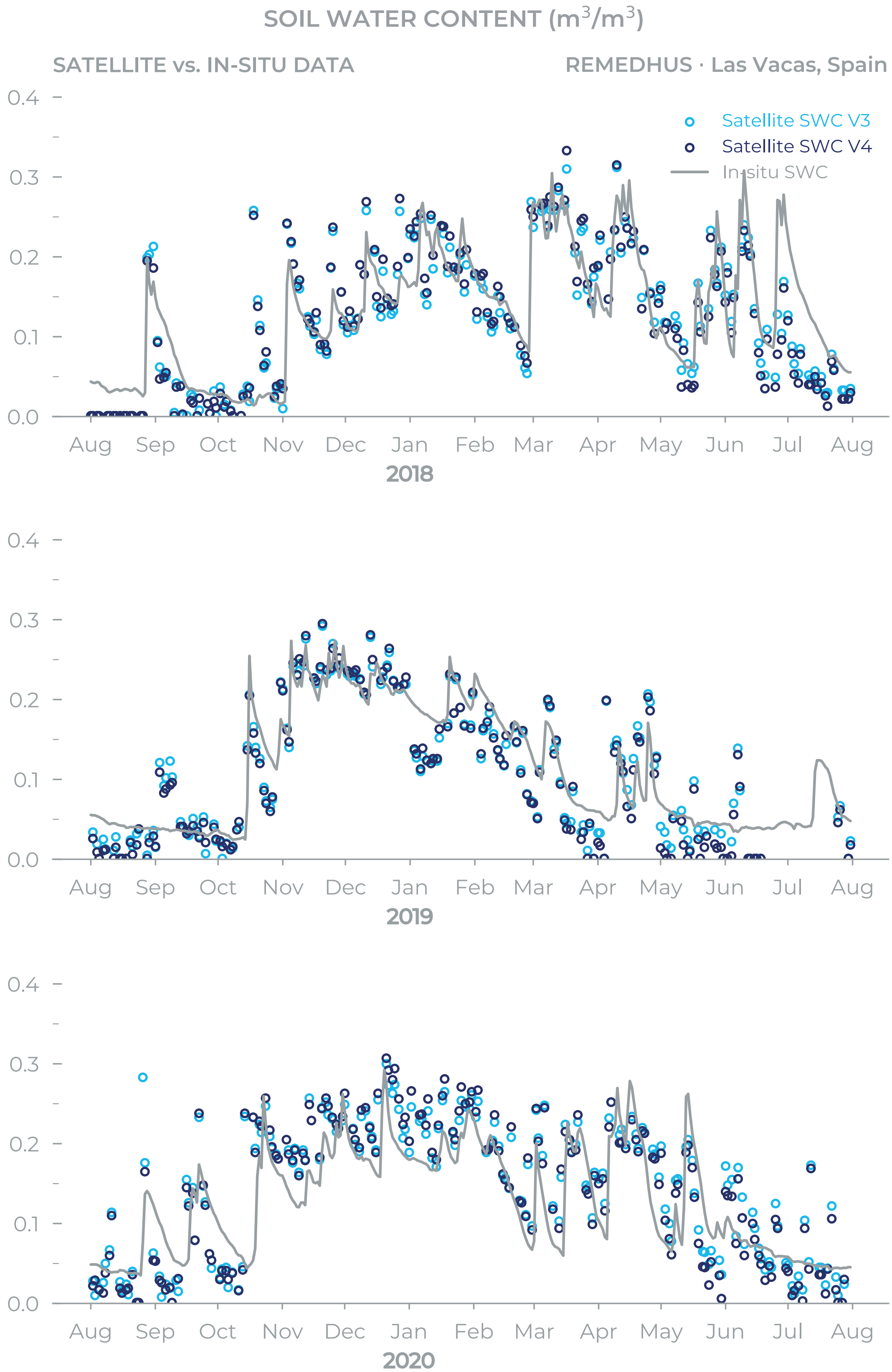
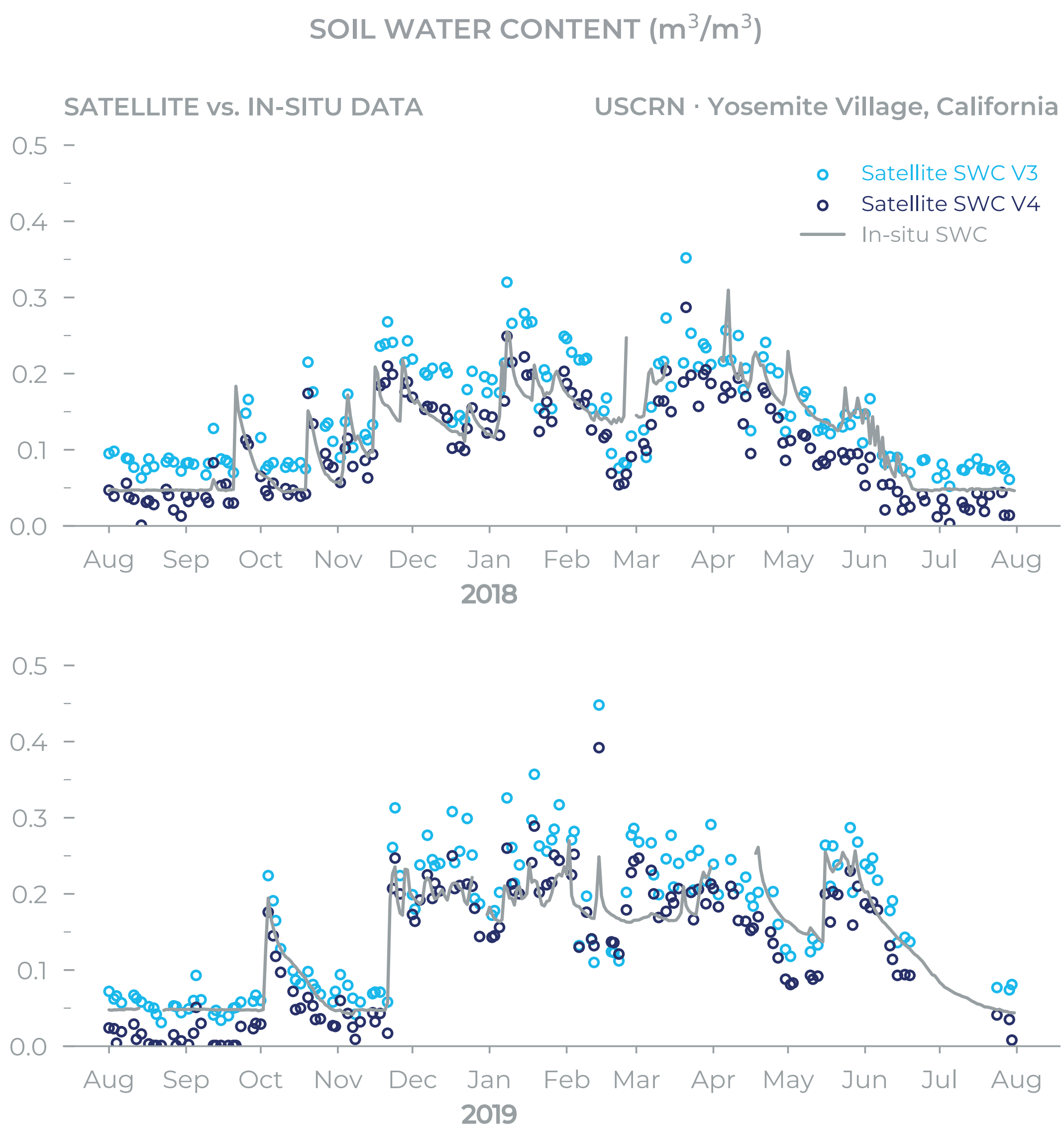


Figure 17: Time series of sensor Las Vacas of the REMEDHUS network, Spain.

USCRN - California, USA

1 August 2017 - 1 August 2020 (3yr data)

Both stations in California have high correlation coefficients. The Yosemite Village station shows a slightly higher correlation for L-band SWC version 4 (R=0.88) than version 3 (R=0.86). The station in Merced shows the opposite (R v3 = 0.89, R v4 = 0.86). The time series of Yosemite Village is shown in Figure 18. The patterns align nicely with the sensor data.



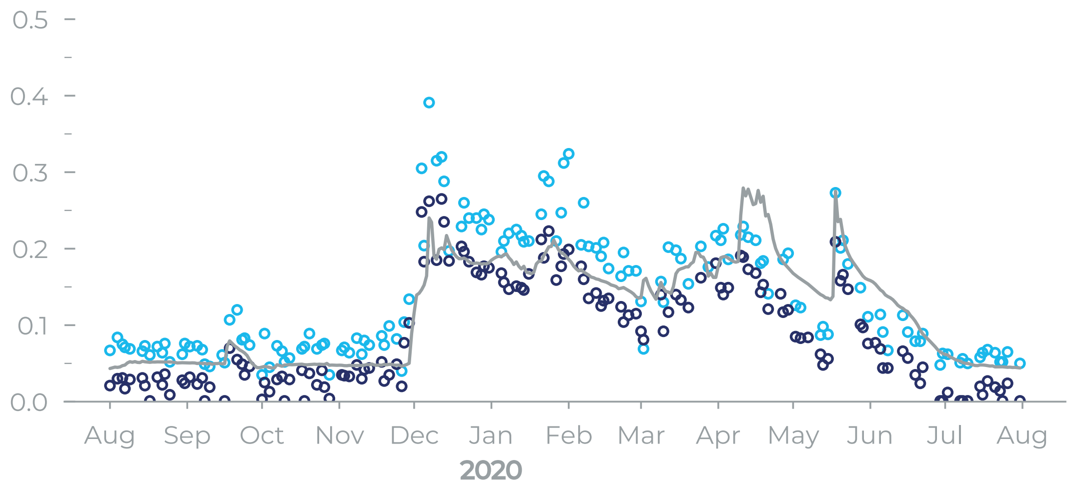


Figure 18: Time series of sensor Yosemite Village of the USCRN network, California.

DISCUSSION AND CONCLUSIONS

In this second part of the validation white paper we show how the validation of our satellite-derived SWC products can be done in a temporal way by cross-comparing time series of satellite and ground sensors. For this evaluation, we used several ground networks on different locations on Earth (the Netherlands, Spain and California). Time series were made to compare the temporal patterns of the ground sensors with the satellite-derived data. Pearson R correlation coefficients were calculated to measure the statistical relationship between the ground and satellite-derived data. The temporal relationships are good, especially for our sensor in the dunes of the Netherlands (Rv4 = 0.85) and for the USCRN stations in California (Rv4= ~0.88). A few stations of the REMEDHUS and RAAM network showed very low correlation coefficients, which can be attributed to malfunctioning in-situ sensors that show unrealistic temporal patterns or missing data. After removing these faulty sensors, the correlation coefficients of the remaining locations of the RAAM and REMEDHUS networks are in range of ~0.53 to 0.91 and mainly around 0.75-0.80. This range is in line with what has been published so far and what we can expect from the quality of satellite data. In NASA's "Soil Moisture Product Validation Good Practices Protocol[13], state-of-the-art L-band soil moisture retrieval algorithms from several L-band satellites have been validated using in-situ ground truth with a correlation of >0.8. By combining data of the sensors of a dense network, even higher correlation coefficients can be obtained. This is the case for the RAAM network where 9 well performing combined sensors result in a R for version 4 of 0.91. As an overall conclusion, our data show good temporal correlations with the selected networks and our result is consistent with previous work on the validation of L-band SWC products.

[13] Montzka, C., et al. (2020): Soil Moisture Product Validation Good Practices Protocol Version 1.0. In: C. Montzka, M. Cosh, J. Nickeson, F. Camacho (Eds.): Good Practices for satellite-derived Land Product Validation (p. 123), Land Product Validation Subgroup (WGCV/CEOS), doi:10.5067/doc/ceoswgcv/lpv/sm.001

OVERALL CONCLUSIONS

The goal of this white paper is to evaluate our satellite-derived SWC product by comparing it with ground data. This is done in two ways, spatially and temporally. Spatial ground data is sparse, so we organized 2 days of fieldwork in the Netherlands to collect in-situ data in an agricultural field. This fieldwork gave us multiple insights. Because of the differences between the spatial support of ground and satellite data, at least 7 ground measurements are needed to reach the same uncertainty as satellite data and to estimate a representative value for a 100 m field. In practice, it takes a lot of time and money to use ground sensors for larger areas. Furthermore, ground sensors need to be properly calibrated. With the 1100+ measurements, obtained within two 100 m pixels, we found similar spatial patterns for the ground and satellite-derived SWC data. In the future, we want to extend this work and provide more insights into the spatial quality of our SWC data.

Satellite-derived SWC time series have a high correlation coefficient with several ground stations on different parts on Earth. The correlation coefficients of the USCRN stations in California and the PWN stations are around 0.88 and 0.85 respectively, which is high for such an evaluation. The correlation coefficients of the other network (RAAM & REMEDHUS) vary between 0.53 and 0.90. In general, our findings are in line with previous studies into L-band SWC data.

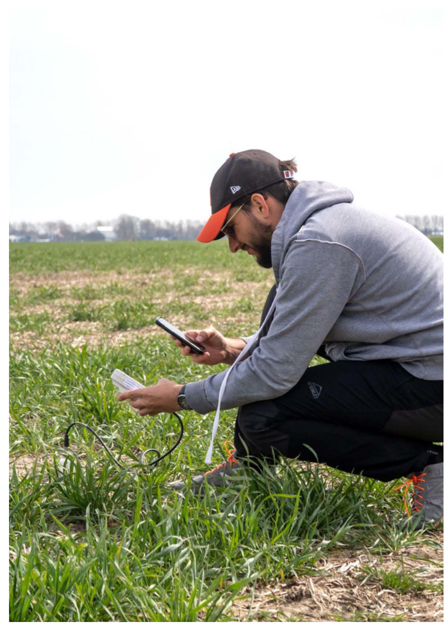
Using ground data as a reference is a suitable way to evaluate satellite-derived SWC observations. However, several things should be kept in mind. Ground sensors have a different spatial support and sensing depth than satellite-derived data. Data of the ground sensor shouldn't be simply considered as the truth, especially for spatial studies. Combining multiple sensors within a certain area also gives valuable insights in the performance of satellite-derived SWC data. We aim to do more research on this validation topic, in particular spatially.

APPENDIX A
GROUND SENSOR CALIBRATION

Background

There are multiple types of ground sensors that measure SWC. The most common type is the Frequency Domain Reflectrometry (FDR) sensor (also called capacitance probe) and the Time Domain Reflectrometry (TDR) sensor. The FDR sensor measures the capacitance as a change in frequency of a reflected radio wave. This method uses an oscillator to propagate an electromagnetic signal through a metal pin or other wave guide. The difference between the output wave and the return wave frequency is linked to the apparent permittivity of the soil and can be used to determine SWC (Dean et al., 1987)[14].

A TDR is similar to an FDR but the mechanics behind the measurement system are different. TDR sensors use parallel rods, acting as transmission lines. A voltage is applied to the rods and reflected back to the sensor for analysis. The speed or velocity of the given energy pulse along the rods is also related to the apparent permittivity of the soil, and can therefore also be used to measure SWC (Roth et al., 1992)[15].



Both the FDR and TDR need to have good contact with the soil since they are measuring with no air gaps. The field of influence is greatest at the sensor and declines rapidly from there. Generally, the field of influence is approximately 1 cm distance from the sensor.

The relationship between apparent permittivity and SWC is different for different soils. Therefore, sensor calibration is always needed with a TDR and FDR.

[14] Dean, T. J. et al. (1987). Soil moisture measurement by an improved capacitance technique, Part I. Sensor design and performance. Journal of Hydrology, 93(1-2), 67-78.

[15] Roth, C. H. et al.. (1992). Empirical evaluation of the relationship between soil dielectric constant and volumetric water content as the basis for calibrating soil moisture measurements by TDR. Journal of Soil Science, 43(1), 1-13.

Calibration procedure

For the field experiment (Part 1) all the in-situ sensors were calibrated. Here we describe the calibration procedure. Undisturbed soil samples were taken from the test location.

These samples were placed on a scale and simultaneously the sensor readings and the weight of the soil samples were recorded (see Fig. A.2). The samples were saturated two times and the dry down was measured with regular intervals over a period of two months. At the end the soil samples were oven dried for 24 hrs and weighed to determine the dry weight of the soil samples.

With this information, we can then compute volumetric SWC (in m^3/m^3) following:

SWC = (Weight Soil Sample - Weight Dry Soil Sample) / Volume Soil Sample

The gravimetric measurements of the soil samples were considered as the ground truth and the sensor readings are adjusted to these values using a linear regression. For each sensor this relationship needs to be determined.

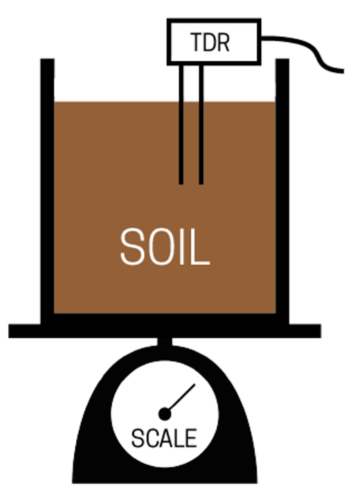
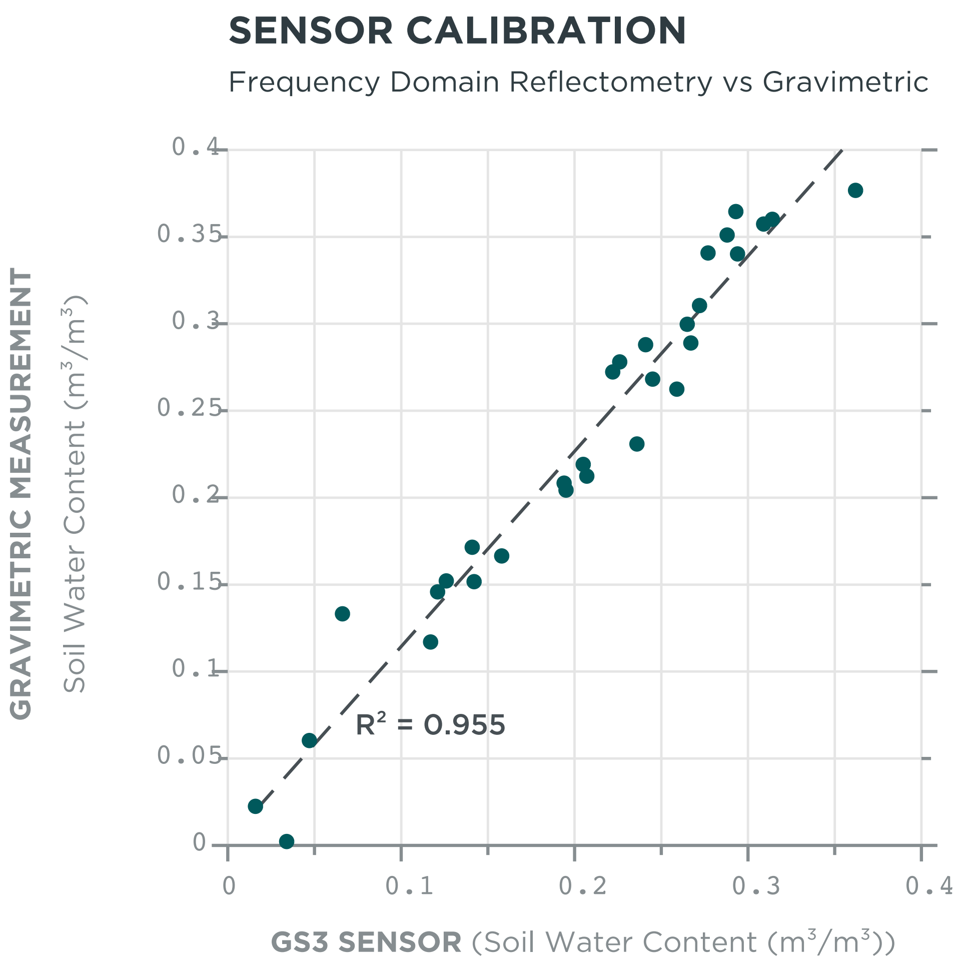


Figure A.2: A schematic representation of a setup were both the soil weight and the sensor reading can be obtained.

All sensors were calibrated and below one example is presented. Here FDR sensor SWC readings were plotted against the gravimetric measurements (based on the measured weights). There was a clear linear relationship, which we used to adjust the FDR sensor SWC readings to actual SWC readings for this given soil. The relationship was a bit off the 1:1 Line with a Pearsons squared correlation of 0.96 and standard error of 0.02 m^3/m^3. All the other sensors (9 in total, 3 FDRs and 6 TDRS) had similar relationships and similar errors. No distinct difference in performance was observed between the TDRs and FDRs used in our experiment. The errors as measured with this experiment were all higher than the numbers presented in the sensor user manuals. They all ranged around 0.02 m^3/m^3.



FigureA.3: An example of a scatter plot between SWC readings as given by the FDR sensor and the measured SWC based on the measured sample weights.

APPENDIX B
CORRELATION COEFFICIENTS

This appendix contains the Pearson R correlation coefficients for the stations in the different networks including the amount over observations (in brackets).

PWN sensor - the Netherlands

1 August 2018 - 1 August 2022 (4yr data)

Sensor	R SWC v3	R SWC v4
PWN	0.824 (835)	0.853 (818)

RAAM network - the Netherlands

1 April 2018 - 1 April 2019 (1yr data)

Station number	R SWC v3	R SWC v4
1	0.683 (271)	0.685 (270)
2	0.875 (271)	0.878 (269)
3	0.384 (271)	0.401 (269)
4	0.773 (271)	0.771 (269)
5	0.760 (271)	0.759 (269)
6	0.878 (271)	0.889 (270)
7	0.866 (271)	0.880 (270)
11	0.722 (271)	0.741 (271)
12	0.349 (271)	0.361 (271)
13	0.524 (271)	0.543 (271)
14	0.757 (271)	0.775 (270)
15	0.781 (271)	0.795 (270

REMEDHUS Network - Spain

1 August 2017 - 1 August 2020 (3yr data)

Station	R SWC v3	R SWC v4
Canizal	0.529 (585)	0.524 (585)
Carretoro	0.705 (585)	0.728 (586)
Casa Periles	0.830 (585)	0.838 (586)
Concejo del Monte	0.644 (575)	0.654 (577)
El Coto	0.557 (585)	0.530 (586)
El Tomillar	0.744 (573)	0.736 (575)
Granja g	0.700 (581)	0.719 (582)
Guarrati	0.713 (579)	0.723 (580)
La Cruz de Elias	0.743 (585)	0.752 (587)
Las Arenas	0.789 (577)	0.785 (579)
Las Bodegas	0.867 (558)	0.869 (558)
Las Brozas	0.549 (568)	0.566 (570)
Las Eritas	0.798 (585)	0.804 (586)
Las Tres Rayas	0.579 (585)	0.594 (586)
Las Vacas	0.837 (585)	0.844 (586)
Las Victorias	0.657 (582)	0.662 (584)
Llanos de la Boveda	0.740 (585)	0.749 (586)
Paredinas	0.770 (585)	0.780 (587)
Zamarron	0.901 (585)	0.904 (586)

USCRN - USA

1 August 2017 - 1 August 2020 (3yr data)

Station	R SWC v3	R SWC v4
Yosemite Village	0.856 (424)	0.880 (429)
Merced	0.885 (471)	0.862 (472)

APPENDIX C
LINEAR REGRESSION VALUES

The in-situ SWC values were scaled to the satellite SWC values for the time series plots. This was done to better be able to compare the patterns. Linear regressions were applied to the in-situ SWC values against the average of the v3 and v4 satellite SWC values (using only the data for the shown time period). The values can be found in this table:

1 April 2018 - 1 April 2019 (1yr data)

Time series	Slope	Intercept
PWN	1.327	0.137
RAAM, 9 STATIONS	1.160	-0.072
RAAM, STATION 12	0.214	0.165
REMEDHUS, Las Vacas	1.673	-0.002
REMEDHUS, Las Eritas	0.996	-0.135
USCRN, Yosemite Village	0.948	0.042

REFERENCES

[1] Gruber, A., et al. (2020). Validation practices for satellite soil moisture retrievals: What are (the) errors?. Remote sensing of environment, 244, 111806.

[2] https://boerenbunder.nl/

[3] Entekhabi, D., et al. (2014) SMAP Handbook Soil Moisture Active Passive, Mapping Soil Moisture Freeze/Thaw from Space

[4] Panciera, R., et al. (2008). The NAFE’05/CoSMOS data set: Toward SMOS soil moisture retrieval, downscaling, and assimilation. IEEE Transactions on Geoscience and Remote Sensing, 46(3), 736-745.

[5] Ye, N., et al. (2020). The soil moisture active passive experiments: Validation of the SMAP products in Australia. IEEE Transactions on Geoscience and Remote Sensing, 59(4), 2922-2939.

[6] Peng, J., Loew, A., Merlin, O., & Verhoest, N. E. (2017). A review of spatial downscaling of satellite remotely sensed soil moisture. Reviews of Geophysics, 55(2), 341-366.

[7] Benninga, H. F., et al. (2018), The Raam regional soil moisture monitoring network in the Netherlands. Earth System Science Data,10, 61–79, 2018 https://doi.org/10.5194/essd-10-61-2018

[8] Mittelbach, H., et al. (2012). A new perspective on the spatio-temporal variability of soil moisture: temporal dynamics versus time-invariant contributions. Hydrology and Earth System Sciences, 16(7), 2169-2179.

[9] Sánchez, N. et al. (2020): Spatial averages of insitu measurements versus remote sensing observations: a soil moisture analysis. Journal of Spatial Science, DOI: 10.1080/14498596.2020.1833769

[10] Sánchez, N. et al. (2012): Validation of the SMOS L2 Soil Moisture Data in the REMEDHUS Network (Spain). IEEE Transactions on Geoscience and Remote Sensing, 50:1602–1611. DOI:10.1109/TGRS.2012.2186971

[11] campus.usal.es/~hidrus/

[12] www.ncei.noaa.gov/access/crn/overview.html

[13] Montzka, C., et al. (2020): Soil Moisture Product Validation Good Practices Protocol Version 1.0. In: C. Montzka, M. Cosh, J. Nickeson, F. Camacho (Eds.): Good Practices for satellite-derived Land Product Validation (p. 123), Land Product Validation Subgroup (WGCV/CEOS), doi:10.5067/doc/ceoswgcv/lpv/sm.001

[14] Dean, T. J. et al. (1987). Soil moisture measurement by an improved capacitance technique, Part I. Sensor design and performance. Journal of Hydrology, 93(1-2), 67-78.

[15] Roth, C. H. et al.. (1992). Empirical evaluation of the relationship between soil dielectric constant and volumetric water content as the basis for calibrating soil moisture measurements by TDR. Journal of Soil Science, 43(1), 1-13.