Article

Soil Moisture Content Estimation Based on Sentinel-1 SAR Imagery Using an Artificial Neural Network and Hydrological Components

Jeehun Chung 1, Yonggwan Lee 1, Jinuk Kim 1, Chunggil Jung 2 and Seongjoon Kim 3,*

1 Department of Civil, Environmental and Plant Engineering, College of Engineering, Konkuk University, Seoul 05029, Korea; gop1519@konkuk.ac.kr (J.C.); leeyg@konkuk.ac.kr (Y.L.); saertt@konkuk.ac.kr (J.K.)
2 Forecast and Control Division, Yeongsan River Flood Control Office, 25, Jukbong-daero 22beon-gil, Gwangju 61934, Korea; wjd0823@konkuk.ac.kr
3 Division of Civil and Environmental Engineering, College of Engineering, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, Seoul 05029, Korea
* Correspondence: kimsj@konkuk.ac.kr; Tel.: +82-2-444-0186

Abstract: This study estimates soil moisture content (SMC) using Sentinel-1A/B C-band synthetic aperture radar (SAR) images and an artificial neural network (ANN) over a 40 × 50-km² area located in the Geum River basin in South Korea. The hydrological components characterized by the antecedent precipitation index (API) and dry days were used as input data as well as SAR (cross-polarization (VH) and copolarization (VV) backscattering coefficients and local incidence angle), topographic (elevation and slope), and soil (percentage of clay and sand)-related data in the ANN simulations. A simple logarithmic transformation was useful in establishing the linear relationship between the observed SMC and the API. In the dry period without rainfall, API did not decrease below 0, thus the Dry days were applied to express the decreasing SMC. The optimal ANN architecture was constructed in terms of the number of hidden layers, hidden neurons, and activation function. The comparison of the estimated SMC with the observed SMC showed that the Pearson’s correlation coefficient (R) and the root mean square error (RMSE) were 0.85 and 4.59%, respectively.

Keywords: antecedent precipitation index; artificial neural network; Sentinel-1; soil moisture content; synthetic aperture radar

1. Introduction

Soil moisture content (SMC) is an important hydrological factor that determines runoff and infiltration in the water cycle following rainfall and affects the global energy balance by influencing the distribution ratio of sensible and latent heat [1–3]. SMC was traditionally measured using observation equipment such as time-domain reflectometry (TDR), which allows a small quantity of point data to be obtained. This method has limitations in representing SMC for a wide area with heterogeneous characteristics. With the development of remote sensing technology and its advantages, which are low cost, high efficiency, real-time, and a wide range [4]. Research on monitoring SMCs across a large area continued over the past decades.

SMC monitoring using remote sensing is mainly divided into an indirect estimation method using an optical satellite and a direct estimation method using a microwave satellite [5–7]. The main differences between these two methods are the electromagnetic energy source, the wavelength region of the electromagnetic spectrum used, the response measured by the sensor and so on [8]. The former estimates SMC by analyzing correlations between SMC and various outputs from optical satellites, such as surface temperature and vegetation-related indices, and uses various statistical, empirical, or machine learning techniques [9–11]. The latter method directly estimates SMC using surface backscatter...
Remote Sens. 2022, 14, 465

and brightness temperature signals obtained from active or passive microwave satellites. Although many studies use microwave radiometers to monitor regional or global SMC [12–14], there is a limit to detailed SMC or drought monitoring in a local area because of the coarse spatial resolution (~10–40 km) of microwave radiometers [15,16]. In the case of South Korea, 64% of the land area is made up of forests, and there is great topographical variability [17], so SMC monitoring using microwave radiometers with a large spatial resolution is not suitable. Active microwave synthetic aperture radars (SARs), such as ERS-1/2, RADARSAT-1, and JERS-1, have high resolutions of up to a few meters, but they are greatly affected by the surface roughness and vegetation cover, and the revisit cycle is generally longer than 15 days [18]. Sentinel-1 is the first satellite launched by the European Space Agency (ESA) as part of the Copernicus Program for detailed global monitoring. Sentinel-1 consists of two satellites, Sentinel-1A, launched in April 2014 and Sentinel-1B, launched in April 2016, which share the same orbital plane and are equipped with C-band SAR with a center frequency of 5.405 GHz [19]. Each of the two satellites has a repeat cycle of 12 days, but when both satellites are used together, a high-resolution image of 10 m can be obtained with a repeat cycle of 6 days, making it a new alternative using SAR data.

Estimation of SMC that uses SAR over bare soil or sparsely vegetated soil was carried out by inverting the radar backscattering coefficient ($\sigma^0$) using physical [20] or statistical models [21,22]. Physical models only require radar configuration (radar wavelength, polarization, and incidence angle) and soil parameters such as correlation length ($L$) or the standard deviation of heights ($H_{rms}$) to simulate $\sigma^0$ [23]. Physical models were successfully applied to invert the SMC, but much information related to vegetation and soil is required, which might be difficult to acquire in the case of a wide and heterogeneous natural site, not a laboratory environment [24–26]. A representative example of a physical model is the integral equation model (IEM) [20]. This model was revealed to be inaccurate for estimating SMC with a large difference between simulated and observed $\sigma^0$ [27–30]. To improve the accuracy of the simulated $\sigma^0$ in the IEM, a semiempirical calibrated IEM (IEM_B) was proposed by Baghdadi et al. [28]. The IEM_B consists of replacing $L$ with a fitting parameter, the so-called $L_{opt}$, to reduce the IEM’s input soil parameters. Statistical models, such as the Bayesian approach, are based on creating a probability distribution function (PDF) from training samples of a radar sensor and soil surface measurement to estimate the unknown surface parameters [29]. The greatest advantage of the Bayesian method is that it is optimal for minimizing the expected error [31]. However, deriving the PDF as in Bayes’ theorem requires a large amount of experimental data, and in the inversion steps, it has the same limitation as forward models as it depends on the limited surface parameter conditions [32].

For vegetated soil, the water cloud model (WCM) proposed by Attema and Ulaby [33] is the main model used to invert SMC from $\sigma^0$. In the WCM, a total $\sigma^0$ is described as the summation of the contribution of the vegetation and soil, multiplied by the attenuation factor. The vegetation contribution and attenuation factor are estimated using vegetation descriptors, such as vegetation water content (VWC), leaf area index (LAI), and vegetation height [34]. However, similar to physical models, there is a limit in applicability because field measurements are required to determine these vegetation-related parameters [35]. Additionally, inaccurate results can be obtained in soil with well-developed vegetation cover because the vegetation contribution dominates the soil contribution in some vegetation types [23].

Recently, the performance of machine learning techniques in which artificial intelligence methods are used has developed rapidly. Various machine learning techniques (e.g., artificial neural networks (ANNs), support vector machines (SVRs), and random forests (RFs)) were used in many studies and showed better accuracy than traditional statistical models [32,36–42]. Among them, ANN is used for SAR-based SMC estimation, and its simulations are performed based on the radar signal, incidence angle, and soil parameters. The influence of vegetation was considered by using vegetation factors, such as the normalized difference vegetation index (NDVI) [4,38,39,43], leaf area index (LAI) [37], and normalized
difference water index (NDWI) [42], estimated by optical satellites or ground vegetation measurements [40].

As an alternative to vegetation parameters, precipitation data were applied by borrowing the concept of antecedent precipitation from the Soil Conservation Service-Curve Number (SCS-CN) method in some studies [9–11,44–46]. The SCS-CN method was developed by the U.S. Soil Conservation Service (SCS) to create the synthetic unit hydrograph [47]. This method considers the soil cover, land use, and antecedent precipitation to create the peak flood volume and runoff hydrograph of the watershed. As with runoff, SMC is also greatly affected by precipitation, and the antecedent precipitation concept of SCS-CN can be applied to SMC estimation. Previous papers showed that the accuracy increased as the number of days of applied antecedent precipitation increased. These results may be influenced by preceding precipitation, but the possibility of overfitting as the independent variable increases cannot be excluded [48]. If the influence of antecedent precipitation can be considered without increasing the independent variables, the estimation performance of SMC can be improved while reducing the effect of overfitting. The antecedent precipitation index (API) is a widely used factor to infer the soil moisture condition, and it is calculated as a weighted sum of daily antecedent precipitation before a given time [49]. Since this index can provide information on water consumption due to surface runoff and evaporation as well as information on preceding precipitation [50,51], it can be an efficacious solution to overcome the overfitting problem.

This study aims to evaluate the coupling of SAR and hydrological components and their applicability in SMC estimation using ANNs. For hydrological components, not only API but also dry days used for agricultural drought analysis [52] were used, and the study was conducted in consideration of topographical factors and soil factors. Also, the optimal configuration of ANNs was tested in terms of the number of hidden layers, hidden neurons, and activation function.

2. Materials and Methods

In this study, SMC was estimated via an ANN by using the hydrological components represented by the API and dry days based on Sentinel-1 C-band SAR images. Sentinel-1A and Sentinel-1B images taken at 12-day intervals each over a total of 5 years, from 2015 to 2019, were collected. During the same period, SMC and precipitation observation data were collected daily. The API and dry days were estimated using precipitation data. The parameters of the API were optimized based on SMC and precipitation data. At the same location of the SMC station, the elevation and slope were extracted from the digital elevation model (DEM), and the percentage of sand and clay was extracted from the soil map. The Sentinel Application Platform (SNAP) provided by the European Space Agency (ESA) was used to preprocess the satellite images, and the data were converted into backscattering coefficients and the incidence angle. The ANN model for estimating SMC was estimated and verified through a comparison with the observed SMC data (Figure 1).

2.1. Study Area

The study area of 40 × 50 km² is located in the province of eastern Jeollabuk-do, South Korea (127°20′ E to 127°45′ E and 35°35′ N to 36°00′ N) and includes the Yongdam watershed in the upper Geum River basin (Figure 2). This area is mountainous, with an average elevation of approximately 500 m, and is composed of 72% forestland, 11% agricultural land, and 9% grassland. The Yongdam watershed is a testbed in which the United Nations Educational, Scientific and Cultural Organization International Hydrological Programme (UNESCO-IHP) conducts various studies on the scientific measurement of hydrological factors and the development and verification of observation equipment to ensure that quantitative water resource analyses and interpretations are highly reliable and that the associated models can be verified [53].
2.1. Study Area
The study area of 40 × 50 km² is located in the province of eastern Jeollabuk-do, South Korea (127°20′ E to 127°45′ E and 35°35′ N to 36°00′ N) and includes the Yongdam watershed in the upper Geum River basin (Figure 2). This area is mountainous, with an average elevation of approximately 500 m, and is composed of 72% forestland, 11% agricultural land, and 9% grassland. The Yongdam watershed is a testbed in which the United Nations Educational, Scientific and Cultural Organization International Hydrological Programme (UNESCO-IHP) conducts various studies on the scientific measurement of hydrological factors and the development and verification of observation equipment to ensure that quantitative water resource analyses and interpretations are highly reliable and that the associated models can be verified [53].

A total of 9 SMC stations are located in the study area; six of these stations are managed by K-water (Korea Water Resources Corporation), while the other three are managed by the RDA (Rural Development Administration). Each of the stations represents one pixel in the SAR image. These stations provide SMC data daily with TDR sensors located at depths
of 10, 20, 40, 60, and 80 cm. In this study, SMC data measured at 10-cm depths were used. Table 1 shows the general information of each SMC station.

Table 1. Relevant information of soil moisture content stations in study area.

| Station     | Latitude (Degree) | Longitude (Degree) | Elevation (m) | Land Cover  | * Max. (%) | * Min. (%) | * Mean (%) | * Std. (%) |
|-------------|-------------------|-------------------|--------------|-------------|------------|------------|------------|------------|
| Ancheon     | 35.8689           | 127.5433          | 313          | Upland crop | 50.8       | 11.8       | 38.3       | 5.8        |
| Bugwi       | 35.8622           | 127.3992          | 396          | Bare field  | 21.9       | 3.6        | 14.3       | 3.4        |
| Cheoncheon  | 35.6847           | 127.5114          | 409          | Grass       | 34.8       | 11.4       | 24.7       | 5.1        |
| Gye Buk     | 35.8075           | 127.6294          | 453          | Upland crop | 41.0       | 13.0       | 30.0       | 5.5        |
| Jinan       | 35.9706           | 127.4242          | 362          | Upland crop | 47.0       | 3.2        | 21.0       | 7.5        |
| Jucheon     | 35.8053           | 127.4844          | 303          | Upland crop | 33.8       | 9.5        | 25.9       | 4.9        |
| Jangsu      | 36.0059           | 127.6780          | 478          | Grass       | 41.1       | 4.2        | 24.7       | 6.3        |
| Muju        | 35.6198           | 127.5120          | 214          | Upland crop | 39.0       | 5.4        | 24.1       | 6.7        |
| Sangseon    | 35.7603           | 127.4380          | 334          | Upland crop | 43.5       | 18.8       | 35.4       | 3.4        |

* Max., Min., Mean, and Std. represent the maximum, minimum, mean, and standard deviation of the observed soil moisture content, respectively.

The elevation and slope data were extracted and estimated from the 10 m spatial resolution DEM, which was provided by the Korea National Spatial Data Infrastructure Portal [54]. The percentages of clay and sand in the soil were constructed from a detailed soil map at a 1:25,000 scale provided by the National Institute of Agricultural Sciences [55]. The soil map provided in vector format was rasterized and resampled to 10 m resolution.

2.2. Sentinel-1 C-Band SAR

Sentinel-1 was the first satellite launched by the ESA as part of the Copernicus Program [56] for detailed global monitoring. Sentinel-1A was launched in April 2014, and Sentinel-1B was launched in June 2016; each satellite has a 12-day cycle [19]. SAR images are provided by satellites according to the transmission and reception directions of microwaves from satellites, comprising a total of 4 types: vertical transmit-vertical receive (VV), vertical transmit-horizontal receive (VH), horizontal transmit-horizontal receive (HH), and horizontal transmit-vertical receive (HV). The image acquisition mode of Sentinel-1 depends on the image resolution and scan width, and four image modes exist: stripmap (SM), interferometric wide swath (IW), extrawide swath (EW), and wave (WV). According to each image mode, a single look complex (SLC) containing phase and amplitude information and a ground range detected (GRD) product containing only amplitude information are provided. In this study, the Sentinel-1A and Sentinel-1B GRD products of the IW mode with 10 m resolution were collected over a period of 5 years at 12-day intervals for each satellite; overall, 93 and 95 images were obtained from Sentinel-1A and Sentinel-1B, respectively (Table 2).

Table 2. Information on Sentinel-1 data used in this study.

| Platform | Sentinel-1A | Sentinel-1B |
|----------|-------------|-------------|
| Orbit direction | Descending | Ascending |
| Product type | GRD | GRD |
| Acquisition mode | IW | IW |
| Polarization mode | Dual (VH, VV) | Dual (VH, VV) |
| Incidence angle (°) | 36–54 | 29–45 |
| Revisit period | 12 days | 12 days |
| Spatial resolution | 10 m | 10 m |
| Acquisition period | Aug. 2015 to Dec. 2019 | Sep. 2016 to Dec. 2019 |
| Number of images | 93 | 95 |

Image preprocessing was performed using the SNAP. The initially provided Sentinel-1 images did not contain accurate orbit information. Therefore, orbit correction was
performed using precise orbit information provided approximately 3 weeks after the date of the first image provision. Additionally, a radiometric correction was performed to estimate the backscattering coefficient from the intensity of each pixel. The thermal noise of each image was corrected using the thermal noise removal function of the SNAP. The Lee sigma filter was used for speckle filtering, according to reference [4], which revealed good applicability for filtering SAR images in SMC estimation. A terrain correction was performed using the DEM output by the Shuttle Radar Topography Mission (SRTM) and a range-Doppler terrain correction. Finally, the backscattering coefficient was estimated using the following dB-scale conversion expression:

\[
\sigma^0_{dB} = 10 \cdot \log_{10} \sigma^0
\]

where \( \sigma^0 \) is sigma naught and \( \sigma^0_{dB} \) is the backscattering coefficient of the image.

2.3. Antecedent Precipitation Index

The API is based on the hydrological concept that the effect of precipitation on the current soil wetness decreases as the number of days since the previous precipitation event increases. This index is defined as the weighted sum of daily precipitation, as proposed by Koehler and Linsely (1951) [49]. The API (in mm) is calculated as follows:

\[
API = \sum_{t=-i}^{-1} P_t k^{-t}
\]

where \( i \) is the number of antecedent days, \( k \) is the decay constant, and \( P_t \) is the precipitation during day \( t \). The value \( i \) is generally set as 5, 7, or 14 days. The value \( k \) depends on the regional characteristics and season and ranges from 0.80 to 0.98 [57].

The API parameters \( (i \text{ and } k) \) were optimized based on SMC and precipitation data observed at the 9 stations in the study area. The optimal parameter combination was found by varying each parameter corresponding to the ranges and increments listed in Table 3.

Table 3. Ranges and increments of Antecedent Precipitation Index (API) parameters.

| Parameter | Description | Range       | Increment |
|-----------|-------------|-------------|-----------|
| \( i \)   | Number of days | 5–14        | 1         |
| \( k \)   | Decay constant | 0.80–0.98   | 0.02      |

2.4. Dry Days

Since the API is expressed as the summation of precipitation during the antecedent days, it is expected to be 0 if there is no precipitation in the corresponding period. During dry conditions, the characteristic related to the maximum decrease in API, which remains at a value of 0, may be insufficient to represent the actual SMC decreasing behavior.

The concept of dry days was applied to supplement the use of APIs. A dry day is a state without or with precipitation below a certain threshold [52], and a sequence of consecutive dry days is often used in agricultural drought analysis [58–61]. The threshold used to determine the dry day is usually greater than 0 to account for the very little amount of rainfall that cannot penetrate the soil due to interception or evaporation [61,62].

Several studies determined this threshold differently in a range of 1 to 10 mm of precipitation [59,61,63–65], and among them, the threshold of 1 mm suggested by Douguedroit (1987) [59] is the most widely used. In South Korea, based on the study of Hershfield (1972) [66], Kim (1995) [67] suggested the 5 mm threshold as a dry-day criterion. This threshold was used in several studies [68–70] and is currently used as a standard dry day in the agricultural drought management system (ADMS) of the rural development administration.

In this study, the dry day was calculated with a 5 mm threshold. As the input data of ANN, the accumulated number of the dry days was also calculated. If the 0, 3, 4, 10, 2 mm
of rainfall for 5 days, the inputs are 1, 2, 3, 0, 1. The accumulation of dry days increases as the dry periods longer and based on this characteristic, it was applied to express the decrease in SMC.

2.5. Artificial Neural Network

An ANN is an information processing system that replicates the behavior of the human brain by imitating the operations and connectivity of biological neurons [71]. In an ANN, complex and nonlinear functions with many parameters are adjusted in the training stage to cause the ANN outputs to be similar to the observed dataset. Additionally, ANNs are not very sensitive to noise in the input data; therefore, ANNs can serve as a universal and flexible function approximator for any kind of data [72].

A multilayer feedforward neural network (MLP-ANN) with a backpropagation learning rule was used to estimate SMC in this study. The MLP-ANN comprises one input layer, one or more hidden layers, and one output layer. In the hidden and output layers, the collections of nodes called artificial neurons are interconnected with weights. These weights are adaptively calibrated and adjusted in each iteration of training by the backpropagation rule to minimize the loss between input and output data. Each artificial neuron has an activation function that ranges between $-1$ and $1$. The tangent sigmoid and logarithmic sigmoid functions are the representative activation functions [73].

In this work, the Python language-based open-source machine learning library TensorFlow was used to build the ANN model. The ANN simulation flow is shown in Figure 3. As input datasets, four categories of data, which are SAR (VH and VV backscattering coefficients ($\sigma_0^{VH}$ and $\sigma_0^{VV}$) and incidence angle), topography (elevation and slope), soil (percentage of sand and clay), and precipitation (API and dry days)-related data, were used (Table 4). SAR, topography, and soil data were used by extracting the value at the same location as the SMC station from the preprocessed Sentinel-1 image, DEM, and soil map, respectively. The precipitation-related data were temporally unified according to the period and time of Sentinel-1 data obtained (i.e., the 6-day interval from 2015 to 2019). As a total, 15,111 data were used for ANN modeling.

| Groups   | Parameter | Source          | Temporal Resolution |
|----------|-----------|-----------------|---------------------|
| SAR      | $\sigma_0^{VH}$ | Sentinel-1      | 6-day               |
| SAR      | $\sigma_0^{VV}$ | Local incidence angle |
| Topography | Elevation | Digital elevation model |
| Topography | Slope   |                |
| Soil     | Sand percentage | Soil map          |
| Soil     | Clay percentage |                |
| PCP      | API       | Observed precipitation data |
| PCP      | Dry days  |                |

$\sigma_0^{VH}$: Backscattering coefficient in dB scale, API: Antecedent Precipitation Index.

![Figure 3. Flow of ANN simulation.](image-url)
Table 4. Brief information of input data used in artificial neural network (ANN) modeling.

| Groups | Parameter | Source | Temporal Resolution |
|--------|-----------|--------|---------------------|
| SAR    | $\sigma^0_{dB}$ VH | Sentinel-1 |                     |
|        | $\sigma^0_{dB}$ VV |         |                     |
|        | Local incidence angle |         |                     |
| Topography | Elevation | Digital elevation model | 6-day |
|         | Slope     |         |                     |
| Soil   | Sand percentage | Soil map |                     |
|        | Clay percentage |         |                     |
|        | API       | Observed precipitation data |     |
| PCP    | Dry days |         |                     |

* $\sigma^0_{dB}$: Backscattering coefficient in dB scale, API: antecedent precipitation index.

All features in the input datasets were normalized into the 0–1 range using the minimum and maximum values of the entire period of each feature. The input datasets were randomly divided into a 5:5 ratio to obtain training and test datasets. The training datasets were randomly subdivided into 60% and 40% subsets. The first subset of training datasets is used to adjust the weights of the ANN using backpropagation, and the second subset is used to validate each training step. For the backpropagation rule, the mean square error (MSE) was selected as a loss function, and the Adam optimizer [74] was used to minimize the losses.

An iterative process was constructed to find the optimal ANN architecture regarding the number of neurons, hidden layer, and activation function. First, a simple ANN consisted of one hidden layer with the number of neurons the same as the number of inputs. After the training, ANN was applied to the test datasets, and its performance was evaluated using Pearson’s correlation coefficient (R) and root mean square error (RMSE). When the optimal input combination is found, the number of hidden layers and the number of neurons are increased and tested again with R and RMSE.

To prevent overfitting, dropout and early stopping were adopted in the ANN architecture. The dropout method involves random drop units from the ANN at the training state to prevent the units from co-adapting [75]. Dropout is generally set in the range of 20–50%. If the dropout value is too high, the simulation performance of the ANN is reduced, and if the dropout value is too low, overfitting occurs [76]. Therefore, in this study, dropout was set to 0.5. The early stopping function automatically stopped training if the model performance stopped improving for more than a certain number of epochs. In this study, when there was no performance improvement over one-fifth of the maximum epochs, the training terminated and was saved as the best model. After training, SMC estimated by the best model was validated using the remaining test datasets.

As an activation function, the rectified linear unit (ReLU) function [77], which was developed to solve the gradient vanishing problem in the sigmoid and hyperbolic tangent, was applied in previous works [4,78–80]. After the proposal of ReLU, new activation functions that improved ReLU, such as LeakyReLU [81], exponential linear units (ELU) [82], and scaled exponential linear unit (SELU) [83], were developed. In this work, the performance of the 4 activation functions, ReLU, LeakyReLU, ELU, and SELU, was also tested.

3. Results

3.1. Relationship between the Backscattering Coefficient and Observed Soil Moisture Content

The comparison results characterized by the R values obtained between the $\sigma^0_{dB}$ and observed SMC values of the Lee sigma filter at 9 stations in the whole period are shown in Table 5. In the Sentinel-1A images, $\sigma^0_{dB}$ VV has a higher correlation with the observed SMC than $\sigma^0_{dB}$ VH, while the Sentinel-1B images show the opposite result. This tendency results from the orbital direction of the satellites [84]. The agricultural area surface, which comprises the majority of the land cover surrounding the SMC stations in the study area,
is generally anisotropic [85], causing differences in the surface roughness values to be indicated by the SAR images depending on the orbital direction. In a previous study [86], when comparing $\sigma_{db}^0$ in soils with isotropic and anisotropic surfaces, it was revealed that $\sigma_{db}^0$ on anisotropic surfaces showed a large difference depending on the angle. Depending on the orbit of the satellite, the soil with an anisotropic surface may differ significantly from $\sigma_{db}^0$ on soil with an isotropic surface.

Table 5. Pearson’s correlation coefficients between backscattering coefficients of cross-polarization (VH) and copolarization (VV) and observed soil moisture content.

| Station | Sentinel-1A |  | Sentinel-1B |  |
|---------|-------------|  |-------------|  |
|         | VH Polarization | VV Polarization | VH Polarization | VV Polarization |
| Ancheon | 0.48 | 0.42 | 0.47 | 0.39 |
| Bugwi   | 0.42 | 0.56 | 0.52 | 0.42 |
| Cheoncheon | 0.26 | 0.36 | 0.31 | 0.32 |
| Gye Buk | 0.46 | 0.55 | 0.53 | 0.44 |
| Jinan   | 0.41 | 0.50 | 0.27 | 0.39 |
| Jucheon | 0.15 | 0.42 | 0.34 | 0.41 |
| Jangsu  | 0.19 | 0.51 | 0.21 | 0.22 |
| Muju    | 0.18 | 0.11 | 0.17 | 0.21 |
| Sangeon | 0.37 | 0.48 | 0.54 | 0.44 |
| Mean    | 0.32 | 0.43 | 0.37 | 0.36 |

The characteristics of the $\sigma_{db}^0$ VH and $\sigma_{db}^0$ VV of the Sentinel-1A images were analyzed by land-use types (Upland crop: Ancheon, Gye Buk, Jinan, Jucheon, Jangsu and Sangeon stations; Grass: Cheoncheon and Muju stations; Bare field: Bugwi station) over the whole period (Figure 4). In $\sigma_{db}^0$ VV, moderate correlations (R = 0.47 and 0.56) are observed in the bare field and upland crop areas (Figure 4a,c). Grassland shows a weak correlation (R = 0.11) due to the multiple scattering effects resulting from the presence of a vegetation layer (Figure 4b). In contrast, the $\sigma_{db}^0$ VH shows more dispersion over the 3 land cover types, which leads to less correlation. However, R is higher at 0.31 in grassland than at $\sigma_{db}^0$ VV due to the depolarization effect on the vegetation [87]. In the same context, the depolarization effect by vegetation may appear in upland crops, but the sensitivity is not high compared to that of VV polarization. Since the soil in the upland crop area lies in a bare state after harvest, it seems that the sensitivity of SMC and $\sigma_{db}^0$ VH over the postharvest period is low.

3.2. API Optimization

Figure 5 summarizes the results of the sensitivity analysis conducted to determine the optimal API parameter combinations. The estimated API and observed SMC values were evaluated using R, and the optimal parameter combination was determined based on maximizing these R values. The R values displayed in Figure 5 are the zonal average values of the study area, and a darker shade indicates a higher R. As a result, the combination of the decay constant $k = 0.98$ and an antecedent period of 5 days ($i = 5$) is found to be the best, with an R value of 0.39.

Scatter plots of the observed SMC and parameter-optimized API values at Gye Buk station, which is the most correlated, are shown in Figure 6. The linear relationship between the observed SMC and the modeled API is uncertain (Figure 6a), and an exponential function is found to be the best fit for the data. A logarithmic transformation was applied to the API to establish a linear relationship between the two datasets. To train the ANN using as much data as possible, we left it at 0 instead of excluding the 0 API. As shown in Figure 6b, the transformed API shows a significant linear relationship with the observed SMC, with an increase in R from 0.41 to 0.54. This result indicates that a simple data transformation can improve the performance of the API as an SMC index.
The characteristics of the $\sigma_0^{VV}$ and $\sigma_0^{VH}$ of the Sentinel-1A images were analyzed by land-use types (Upland crop: Ancheon, Gyebuk, Jinan, Jucheon, Jangsu and Sangjeon stations; Grass: Cheoncheon and Muju stations; Bare field: Bugwi station) over the whole period (Figure 4). In $\sigma_0^{VV}$, moderate correlations ($R = 0.47$ and $0.56$) are observed in the bare field and upland crop areas (Figure 4a, c). Grassland shows a weak correlation ($R = 0.11$) due to the multiple scattering effects resulting from the presence of a vegetation layer (Figure 4b). In contrast, the $\sigma_0^{VH}$ shows more dispersion over the 3 land cover types, which leads to less correlation. However, $R$ is higher at 0.31 in grassland than at $\sigma_0^{VV}$ due to the depolarization effect on the vegetation [87]. In the same context, the depolarization effect by vegetation may appear in upland crops, but the sensitivity is not high compared to that of VV polarization. Since the soil in the upland crop area lies in a bare state after harvest, it seems that the sensitivity of SMC and $\sigma_0^{VH}$ over the postharvest period is low.

3.2. API Optimization

Figure 5 summarizes the results of the sensitivity analysis conducted to determine the optimal API parameter combinations. The estimated API and observed SMC values were evaluated using $R$, and the optimal parameter combination was determined based on maximizing these $R$ values. The $R$ values displayed in Figure 5 are the zonal average values of the study area, and a darker shade indicates a higher $R$. As a result, the combination of the decay constant $k = 0.98$ and an antecedent period of 5 days ($i = 5$) is found to be the best, with an $R$ value of 0.39.

3.3. Estimating Soil Moisture Content Using ANN

After optimizing the API parameters, the SMC was estimated using the ANN based on the dataset covering the period from 2015 to 2019. First, simulation results according to the combination of input data are compared in Table 6. In the case of simulating ANN using only SAR-derived input, $R$ 0.3 or less and RMSE 8% or more show that additional data are needed to estimate SMC. As a result of the simulation, by adding topographical factors, the correlation significantly increases to $R$ 0.59 and RMSE 7.37% during training and $R$ 0.56 and RMSE 7.39% during the test, but the RMSE is still as high as 7%. The addition of soil attributes increases the performance of ANN by $R$ 0.67 and RMSE 6.57% in the test phase. Finally, the simulation using all data features shows the best result of $R$ 0.76 and RMSE 5.79% in the test state. ANN simulation without SAR input configuration was also tested. As the input of API and dry days, a result of $R$ 0.61 RMSE 6.99% shows in the test state. However, this is a lower performance than the result of the simulation considering SAR configuration. SAR images may have limitations in SMC estimation due to the revisit cycle. Nevertheless, the fusion of SAR images with additional data serves as a powerful tool for SMC estimation.
Figure 5. Correlation matrix of various combinations of API parameters to correspond with observed soil moisture content.

Figure 6. Comparison between (a) antecedent precipitation index and (b) logarithmic-transformed antecedent precipitation index against observed soil moisture content at Gyebuk station.

Figure 7 shows the ANN simulation results of 4 activation functions using all data features. As seen in the figure, the simulation performance is highest in the order of LeakyReLU, ReLU, SELU, and ELU (R: 0.82, 0.81, 0.77, and 0.76, respectively, RMSE: 5.40% for LeakyReLU, 5.79% for ReLU, 6.19% for SELU, and 6.57% for ELU).

Table 6. ANN simulation results according to input data combinations.

| Inputs             | Training | Test   | Overall |
|--------------------|----------|--------|---------|
|                    | R        | RMSE (Vol.%) | R        | RMSE (Vol.%) | R        | RMSE (Vol.%) |
| SAR                | 0.24     | 8.82   | 0.19    | 8.68   | 0.21    | 8.75   |
| SAR + * Topography | 0.59     | 7.37   | 0.56    | 7.39   | 0.58    | 7.38   |
| SAR + Topography + * Soil | 0.74 | 6.19 | 0.67    | 6.57   | 0.70    | 6.38   |
| SAR + Topography + Soil + * PCP | 0.81 | 5.40 | 0.76    | 5.79   | 0.78    | 5.60   |
| Topography + Soil + PCP | 0.69     | 6.59   | 0.61    | 6.99   | 0.65    | 6.79   |

* SAR: backscattering coefficients of VH and VV polarization, incidence angle; topography: elevation, slope; * soil: percentage of clay and sand; PCP: antecedent precipitation index, dry days.
Table 6. ANN simulation results according to input data combinations.

| Inputs                        | Training          | Test             | Overall         |
|-------------------------------|-------------------|------------------|-----------------|
|                               | R     | RMSE (Vol.%) | R     | RMSE (Vol.%) | R     | RMSE (Vol.%) |
| * SAR                         | 0.24  | 8.82        | 0.19  | 8.68        | 0.21  | 8.75        |
| SAR + Topography              | 0.59  | 7.37        | 0.56  | 7.39        | 0.58  | 7.38        |
| SAR + Topography + Soil       | 0.74  | 6.19        | 0.67  | 6.57        | 0.70  | 6.38        |
| Topography + Soil + * PCP     | 0.81  | 5.40        | 0.76  | 5.79        | 0.78  | 5.60        |
| Topography + Soil + PCP       | 0.69  | 6.59        | 0.61  | 6.99        | 0.65  | 6.79        |

* SAR: backscattering coefficients of VH and VV polarization, incidence angle; topography: elevation, slope; * soil: percentage of clay and sand; PCP: antecedent precipitation index, dry days.

Figure 7 shows the ANN simulation results of 4 activation functions using all data features. As seen in the figure, the simulation performance is highest in the order of LeakyReLU, ReLU, SELU, and ELU (R: 0.82, 0.81, 0.77, and 0.76, respectively, RMSE: 5.17%, 5.59%, 5.66%, and 5.78%, respectively). ReLU and LeakyReLU, which show good simulation performance, show little difference in R (0.01) but more difference in RMSE (0.42%). This seems to originate from the phenomenon that the ANN simulation in ReLU converges at approximately 27% (red dotted line in Figure 7a). This phenomenon, the so-called “dying ReLU” [88], is the scenario when many ReLU neurons only output values of 0 in the backpropagation. This dying ReLU problem can be solved by reducing the learning rate, but in the rest of the activation functions, this is not found at the same learning rate. In LeakyReLU, the simulation proceeds normally, and the gradient is the most linear among the 4 activation functions. The simulation results using ELU and SELU are more dispersed than the simulation results of the top two activation functions, so the performance is not as good as that of ReLU.

The effects of the number of hidden layers and hidden neurons on ANN simulation were analyzed (Figure 8). The performance of ANN using LeakyReLU was evaluated by R and RMSE while increasing the number of hidden layers from 1 to 5. Based on the number of hidden layers showing the best performance, performance evaluation was performed by increasing the number of hidden neurons to 1 to 11 times the input layer (i.e., 9 to 99 hidden neurons). As a result, when the number of hidden layers is 3, the R is 0.84 and RMSE is 4.76%, which are most correlated to the observed SMC (Figure 8a). In terms of hidden neurons, R is the lowest, which is 0.84 for 9 hidden neurons and does not change further to 0.85 for more than 18 hidden neurons. RMSE is low at 4.61% and 4.59% when there are 27 and 99 hidden neurons, respectively. Although the RMSE is lowest when the number of hidden neurons is 99, in this study, an ANN model with 3 hidden layers and 27 hidden neurons was selected as the optimal model to design an ANN with a relatively simple architecture.
ReLU and LeakyReLU, which show good simulation performance, show little difference in R (0.01) but more difference in RMSE (0.42%). This seems to originate from the phenomenon that the ANN simulation in ReLU converges at approximately 27% (red dotted line in Figure 7a). This phenomenon, the so-called “dying ReLU” [88], is the scenario when many ReLU neurons only output values of 0 in the backpropagation. This dying ReLU problem can be solved by reducing the learning rate, but in the rest of the activation functions, this is not found at the same learning rate. In LeakyReLU, the simulation proceeds normally, and the gradient is the most linear among the 4 activation functions. The simulation results using ELU and SELU are more dispersed than the simulation results of the top two activation functions, so the performance is not as good as that of ReLU.

**Figure 7.** Scatter plots between observed and estimated soil moisture content results of test state using ANN with different activation functions: (a) rectified linear unit (ReLU), (b) LeakyReLU, (c) exponential linear units (ELU), and (d) scaled exponential linear unit (SELU).

The simulated ANN results using optimal parameters are plotted in Figure 9. The estimated SMC is summarized as seasonal (spring: March to May; summer: June to August; autumn: September to November; and winter: December to February of the following year). A red solid line and black dotted line representing the 1:1 line and difference of ± one standard deviation (SD) of the observed SMC, respectively, are also included in the figure. The most stable results are seen in autumn, with an R of 0.90 and an RMSE of 3.60% (Figure 9c). However, some values are beyond the range of ± one SD during the summer and winter dry seasons, resulting in high RMSEs of 5.13% and 5.51% (Figure 9b,d), respectively. The SMC estimated in summer is linearly fitted except for the values estimated under dry conditions, while the winter results are more dispersed; therefore, the R values...
of the summer and winter datasets are 0.90 and 0.78, respectively. Due to the characteristics of the weather in South Korea, the amount of annual precipitation is highly concentrated in summer [46]; therefore, the soil surface has sufficient moisture in this season, expressed in the high correlation of SMC with $c_{db}$ [89]. Therefore, the summer results are the best in terms of the resulting R with the autumn result, but the summer RMSE is higher than that obtained in autumn, possibly because the active growth of vegetation along with rainfall increases the simulation uncertainty. Spring shows intermediate results between summer and winter, and the R and RMSE values are 0.88 and 3.83%, respectively (Figure 9a). Outliers are also found in some of the spring datasets, which possibly resulted from the remaining snow cover in the high-elevation regions of the study area in early spring. The simulated results for the whole period are R 0.85 and RMSE 4.59%.

![Figure 8](image)

**Figure 8.** ANN simulation performance according to (a) number of hidden layers and (b) number of hidden neurons with three hidden layers. Bar graph represents Pearson correlation coefficient (R), and red dot with the solid line represents root mean square error (RMSE).

![Figure 9](image)

**Figure 9.** Scatter plots between observed and estimated soil moisture contents using ANN. Each plot shows results of (a) spring (March to May), (b) summer (June to August), (c) autumn (September to November), (d) winter (December to February of the following year), and (e) whole period estimations. A dotted line represents difference of ±1 standard deviation (SD) of observed soil moisture content.
The effect of dry days on the SMC simulation results was analyzed. Figure 10 shows the time series of the observed SMC, estimated SMC, and dry days. As shown in the blue-shaded areas in the figure, both overestimated and underestimated simulations were obtained. Compared to the observed SMC between the 25% and 35% ranges in the winter of 2015 and 2016, the number of dry days is approximately 80 days. In the spring and winter of 2017 and the summer of 2018, the number of dry days is approximately 60 days compared to the sharply reduced SMC. As a result, the SMC estimations obtained herein are affected by dry days, especially when the expression of the decreasing SMC pattern was insufficient or excessive.

The temporal behavior of SMC was compared with the monthly precipitation of the study area. Figure 11 shows the temporal variation in monthly precipitation and SMC obtained by the ANN. During the entire period, depending on the Korean climatic characteristics, precipitation was concentrated mainly in June to August, and minimum precipitation appeared from December to February. The cumulative annual precipitation was 893.5 mm, 1334.4 mm, 1046.5 mm, 1500.0 mm, and 1180.0 mm for 2015, 2016, 2017, 2018, and 2019, respectively. In 2015, when there was relatively low rainfall among the five years, the SMC changes according to the precipitation and changes sensitively. Conversely, there is no significant change in SMC from 2018 to 2019 except during the winter of 2018 and the spring of 2019. This is related to the annual distribution of monthly precipitation. From 2018 to 2019, the monthly precipitation was evenly distributed at more than 100 mm. The uniform distribution of precipitation maintains the SMC at a certain level; therefore, the change in SMC is small [90]. The maximum value of SMC is approximately 25% in 2015, not in 2019, when the annual precipitation was the highest. This might be the result of additional irrigation for crop growth in the dry season, unlike in the normal or wet seasons. In addition, monthly precipitation and monthly mean SMC estimated by ANN have a logarithmic relationship with an R of 0.70.
4. Discussion

Sentinel-1 images taken under the same conditions were converted to $\sigma^0_B$ values using two speckle-filtering methods, and the Lee sigma filtering method was found to be the most suitable. After comparing various filtering techniques described in a previous study [4], the correlation of the estimated SMC with the observed SMC is found to be highest in both bare fields and vegetated fields.

As shown in Figure 4, the correlation between the $\sigma^0_B$ values of upland crops and SMC is intermediate between the grasslands and bare fields. The $\sigma^0_B$ properties in bare fields and vegetated fields have previously been researched [91,92]. In this study, the upland crop characteristics are interpreted as resulting from the type and planting density of the crops grown in each field or from the difference between fallow conditions and vegetated conditions in the fields.

Overall, the API and transformed API values show adequate performance in terms of the SMC index. However, the undiminished API characteristics in the dry period are reduced when surface moisture is sufficiently and homogeneously distributed, and stable SMC estimations are highly probable under these conditions [94]. However, in the opposite case, nonlinear SMC behavior is considered likely to be produced.

To explain the dispersion of the scatter plots displayed in Figure 9 and the inaccuracy of the resulting SMC estimations, a few arguments are proposed:

- Increased surface interference occurs under dry conditions. The surface interference is reduced when surface moisture is sufficiently and homogeneously distributed, and stable SMC estimations are highly probable under these conditions [94]. However, in the opposite case, nonlinear SMC behavior is considered likely to be produced.

- Multiple scattering effects result from the vegetation density. Generally, as temperatures increase, evaporation from the ground surface becomes more active, and the vegetation biomass increases, resulting in a decreased SMC [95]. As a result, errors occur due to the scattering of the radar signal above dense vegetation; these errors can be inferred to influence the SMC estimation.

- Soil freezes in cold weather. Soil freezing is a well-known problem related to the reliability of SMC measurements obtained using TDR or frequency domain reflectometry (FDR). Freezing soil causes a change in the dielectric constant, causing the $\sigma^0_B$ and the brightness temperature to change abruptly and significantly [94]. Filtering a soil freezing condition can be done using the soil temperature data when the soil temperature drops below 0 °C [96]. However, in this study, filtering soil freeze was impossible due to the non-existent soil temperature data at the SMC stations. In a further study, soil temperature nor surface temperature data should be collected to

Figure 11. Temporal variations in monthly precipitation and mean soil moisture content.
filter the soil freezing states directly or indirectly, and also compared the result of ANN simulation with/without filtering soil freeze.

- The dry-day threshold can affect outputs. The dry-day threshold was set to 5 mm in this work; however, under- and overestimates compared to the observed decreases in SMC occurred. In the study area, this trend was found to be especially prominent in summer and winter, but it cannot be generalized as a seasonal pattern. It is necessary to determine an optimum threshold value through seasonal analyses as well as through SMC variation pattern analyses.

One of the limitations of this study may be the SMC measurement method. The penetration depth of the C-band is 2 to 5 cm [97], but the observed SMC was measured at a depth of 10 cm, implying uncertainty. Since the topsoil layer is subjected to faster drying and rewetting, the SMC patterns that occur at a depth of 10 cm may be different from those that occur in the upper soil layer. The public SMC observation data provided in South Korea provide data only from 10 cm below the surface, and the data quality is not actively managed [98]. To obtain a better understanding in SMC studies, SMC measurements obtained at various depths and quality control are required.

In terms of the ANN simulation, the results of this study were obtained on a relatively small dataset, therefore it should be verified on a larger dataset. A characteristic of data-driven approaches such as ANNs is that since the datasets used are obtained from specific environmental conditions (i.e., site-dependent), they must be verified their representativeness of other environmental conditions [43]. These site dependencies can be applied the same for API and dry days. The characteristics and optimal parameters of the two features are based on precipitation, differ for each region. Reducing this site dependency will be possible by collecting large-scale representative datasets from various regions and updating the ANN. Utilizing global products such as the international soil moisture network (ISMN) [99] and global precipitation measurement mission (GPM) [100] will be of great assistance in the verification and generalization of the algorithm presented in this study.

5. Conclusions

In this study, the hydrological components characterized by the antecedent precipitation index (API) and dry days were proposed as auxiliary data to obtain Sentinel-1 synthetic aperture radar (SAR)-based soil moisture content (SMC) estimations in South Korea. SMC modeling was performed using an ANN, and the primary conclusions are summarized as follows:

1. In the SAR image preprocessing step, the technique that effectively reduced the speckle pattern and facilitated a high correlation of the outputs with the observed SMC was found to be the Lee sigma filtering method. The correlation between the derived $\sigma_{dB}^0$ VV (copolarization) and SMC values was highest in the bare fields (R = 0.56) and lowest in the grasslands (R = 0.18). The $\sigma_{dB}^0$ VH (cross-polarization) in grassland had a higher correlation than $\sigma_{dB}^0$ VV due to the depolarization effect. For the upland crop that remaining exposed after harvest, the correlation of $\sigma_{dB}^0$ VH was lower than that of grassland.

2. The API showed incomplete linearity with SMC; thus, a logarithmic transformation was performed to establish a linear relationship. As a result, the R value increased from 0.41 to 0.54. However, the API did not decrease below 0 even under dry conditions; therefore, dry days were introduced to reflect the decreased SMC.

3. The ANN performance increased when not only SAR data but also topographical, soil, and rainfall-related data were added. In terms of activation functions, LeakyReLU, ReLU, SELU, and ELU showed good performance in that order. Based on the LeakyReLU activation function, the best performance was achieved when the number of hidden layers and hidden neurons was 3 and 27, respectively.

4. As a result of estimating SMC using the whole dataset, the most stable result was obtained in autumn (R = 0.90 and RMSE = 3.60%), and outliers were found under
dry conditions in summer, winter, and spring. These results can be explained by the presence of snow cover in winter and remaining snow in early spring in the high-elevation regions, the increasing vegetation biomass with heavy rainfall in summer, and the freezing of soils in winter due to low temperatures. In addition, the estimated SMC values were overestimated and underestimated when dry days overexpressed and underexpressed SMC loss, respectively.

5. The average monthly SMC simulated through the Artificial Neural Network (ANN) changed well according to the monthly precipitation. When monthly precipitation was evenly distributed over 100 mm a year, the SMC did not change much, and when it was concentrated in a specific season, it changed sensitively. Moreover, monthly precipitation and monthly mean SMC showed a logarithmic relationship with an R value of 0.70.

In conclusion, the use of API and dry days showed good applicability in SAR-based SMC estimations. Nevertheless, limitations and uncertainties were presented in this study. To obtain better results and overcome these limitations, several future works can be performed. Using SMC observations collected at a depth of 5 cm may reduce the uncertainty according to the C-band penetration depth. A reliable API can be obtained by estimating an empirically independent \( k \) parameter from sufficient data by using the method described in a previous study [93]. The optimal threshold of dry days corresponding to the observed SMC values can be determined, thus reducing the number of outliers. Filtering the snow-covered and frozen soil states using snowfall data and soil temperature data may boost the SMC estimation performance. Overall, a lot of input data from various regions are required for ANN simulation and verification.

Author Contributions: Conceptualization, Y.L. and S.K.; data curation, J.C., J.K. and C.J.; software, J.C. and J.K.; formal analysis, J.C. and C.J.; writing—original draft preparation, J.C. and Y.L.; and writing—review & editing, Y.L. and S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was supported by Konkuk University in 2017.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References
1. Seneviratne, S.I.; Corti, T.; Davin, E.L.; Hirschi, M.; Jaeger, E.B.; Lehner, I.; Orlowsky, B.; Teuling, J. Investigating soil moisture–climate interactions in a changing climate: A review. Earth-Sci. Rev. 2010, 99, 125–161. [CrossRef]
2. Fallon, P.; Jones, C.D.; Ades, M.; Paul, K. Direct soil moisture controls of future global soil carbon changes: An important source of uncertainty. Glob. Biogeochem. Cycles 2011, 25, GB3010. [CrossRef]
3. Zhang, D.; Zhou, G. Estimation of Soil Moisture from Optical and Thermal Remote Sensing: A Review. Sensors 2016, 16, 1308. [CrossRef] [PubMed]
4. Liu, Y.; Qian, J.; Yue, H. Combined Sentinel-1A with Sentinel-2A to Estimate Soil Moisture in Farmland. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2021, 14, 1292–1310. [CrossRef]
5. Gillies, R.R.; Carlson, T.N. Thermal Remote Sensing of Surface Soil Water Content with Partial Vegetation Cover for Incorporation into Climate Models. J. Appl. Meteorol. Climatol. 1995, 34, 745–756. [CrossRef]
6. Sandholt, I.; Rasmussen, K.; Andersen, J. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. Remote Sens. Environ. 2002, 79, 213–224. [CrossRef]
7. Amazirh, A.; Merlin, O.; Er-Raki, S.; Gao, Q.; Rivalland, V.; Malbeteau, Y.; Khabba, S.; Escorihuela, M.J. Retrieving surface soil moisture at high spatio-temporal resolution from a synergy between Sentinel-1 radar and Landsat thermal data: A study case over bare soil. Remote Sens. Environ. 2018, 211, 321–337. [CrossRef]
8. Wang, L.; Qu, J.J. Satellite remote sensing applications for surface soil moisture monitoring: A review. Front. Earth Sci. China 2009, 3, 237–247. [CrossRef]
36. Lakhankar, T.; Ghedira, H.; Khanbilvardi, R. Soil moisture retrieval from RADARSAT data: A Neuro-Fuzzy approach. In Proceedings of the IEEE International Symposium on Geoscience and Remote Sensing, Denver, CO, USA, 31 July–4 August 2006; pp. 2228–2331.

37. Said, S.; Kothyari, U.C.; Arora, M.K. ANN-Based Soil Moisture Retrieval over Bare and Vegetated Areas Using ERS-2 SAR Data. J. Hydrol. Engineering 2008, 13, 461–475. [CrossRef]

38. Paloscia, S.; Pettinato, S.; Santi, E.; Notarnicola, C.; Pasolli, L.; Repucci, A. Soil moisture mapping using artificial neural networks: Algorithm and preliminary validation. Remote Sens. Environment. 2013, 134, 234–248. [CrossRef]

39. Hajdu, I.; Yule, I.; Dehghan-Shear, M.H. Modelling of Near-Surface Soil Moisture Using Machine Learning and Multi-Temporal Sentinel 1 Images in New Zealand. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGRASS), Valencia, Spain, 22–27 July 2018; pp. 1422–1425.

40. Hachani, A.; Ouessar, M.; Paloscia, S.; Santi, E.; Pettinato, S. Soil moisture retrieval from Sentinel-1 acquisitions in an arid environment in Tunisia: Application of Artificial Neural Networks techniques. Int. J. Remote Sens. 2019, 40, 9159–9180. [CrossRef]

41. Santi, E.; Dabboor, M.; Pettinato, S.; Paloscia, S. Combining Machine Learning and Compact Polariometry for Estimating Soil Moisture from C-Band SAR Data. Remote Sens. 2019, 11, 2451. [CrossRef]

42. Ayehu, G.; Tadesse, T.; Gessesse, B.; Yigrem, Y.M.; Melesse, A. Combined Use of Sentinel-1 SAR and Landsat Polarmeters Products for Residual Soil Moisture Retrieval over Agricultural Fields in the Upper Blue Nile Basin, Ethiopia. Sensors 2020, 20, 3282. [CrossRef]

43. Santi, E.; Paloscia, S.; Pettinato, S.; Fontanelli, G. Application of artificial neural networks for the soil moisture retrieval from active and passive microwave spaceborne sensors. Int. J. Appl. Earth Obs. Geoinf. 2016, 51, 61–73. [CrossRef]

44. Lee, Y.G.; Jung, C.G.; Cho, Y.H.; Kim, S.J. Estimation of soil moisture using multiple linear regression model and COMS land surface temperature data. J. Korean Soc. Agric. Eng. 2017, 59, 11–20.

45. Jung, C.; Lee, Y.; Cho, Y.; Kim, S. A study of spatial soil moisture estimation using a multiple linear regression model and MODIS land surface temperature data corrected by conditional merging. Remote Sens. 2017, 9, 870. [CrossRef]

46. Chung, J.; Son, M.; Lee, Y.; Kim, S. Estimation of Soil Moisture Using Sentinel-1 SAR Images and Multiple Linear Regression Model Considering Antecedent Precipitations. Korean J. Remote Sens. 2021, 37, 515–530.

47. Soil Conservation Service. National Engineering Handbook, Section 4, Hydrology; Department of Agriculture: Washington, DC, USA, 1972.

48. Douglas, D.M. The Problem of Overfitting. J. Chem. Inf. Comput. Sci. 2004, 44, 1–12.

49. Koehler, M.A.; Linsley, R.K. Predicting the Runoff from Storm Rainfall, Research Paper n.34; Weather Bureau; US Department of Commerce: Washington, DC, USA, 1951.

50. Sittner, W.T.; Schauss, C.E.; Monro, J.C. Continuous hydrograph synthesis with an API-type hydrologic model. Water Resour. Res. 1969, 5, 1007–1022. [CrossRef]

51. Wei, F.; Hu, K.; Zhang, J.; Jiang, Y.; Chen, J. Determination of Effective Antecedent Rainfall for Debris Flow Forecast Based On Soil Moisture Content Observation in Jiangjia Gully, China. WIT Trans. Eng. Sci. 2008, 60, 13–22.

52. Rivoire, P.; Tramblay, Y.; Neppel, L.; Hertig, E.; Vicente-Serrano, S.M. Impact of the dry-day definition on Mediterranean extreme dry-spell analysis. Nat. Hazards Earth Syst. Sci. 2019, 19, 1629–1638. [CrossRef]

53. Ryu, G.S. UNESCO-IHP Yongdam study watershed. Water Future 2017, 50, 70–74.

54. National Spatial Data Infrastructure Portal. Available online: http://www.irdi.go.kr (accessed on 1 November 2018).

55. Korean Soil Information System. Available online: http://soil.rda.go.kr (accessed on 1 November 2021).

56. Aschbacher, J.; Milagro-Peréz, M.P. The European Earth monitoring (GMES) programme: Status and perspectives. Remote Sens. Environment. 2012, 120, 3–8. [CrossRef]

57. Viessman, W.; Klapp, J.W.; Lewis, G.L.; Harbaugh, T.E. Introduction to Hydrology; Harper and Row: New York, NY, USA, 1977.

58. Sivakumar, M.V.K. Empirical analysis of dry spells for agricultural applications in West Africa. Hydrol. Earth Syst. Sci. 2008, 12, 3–8. [CrossRef]

59. Hajdu, I.; Yule, I.; Dehghan-Shear, M.H. Modelling of Near-Surface Soil Moisture Using Machine Learning and Multi-Temporal Sentinel 1 Images in New Zealand. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGRASS), Valencia, Spain, 22–27 July 2018; pp. 1422–1425.

60. Mathugama, S.C.; Peiris, T.S.G. Critical Evaluation of Dry Spell Research. Int. J. Basic Appl. Sci. 2011, 11, 153–160.

61. Raymond, F.; Ullmann, A.; Camberlin, P.; Drobinski, P.; Chateau Smith, C. Extreme dry spell detection and climatology over the Mediterranean Basin during the wet season. Geophys. Res. Lett. 2016, 43, 7196–7204. [CrossRef]

62. Dougouedroit, A. The variations of dry spells in Marseilles from 1865 to 1984. Int. J. Climatol. 1987, 7, 541–551. [CrossRef]

63. Vicente-Serrano, S.M.; Beguñá-Portugués, S. Estimating extreme dry-spell risk in the middle Ebro valley (northeastern Spain): A comparative analysis of partial duration series with a general Pareto distribution and annual maxima series with a Gumbel distribution. Int. J. Climatol. 2006, 26, 1999–2021. [CrossRef]

64. Serra, C.; Lana, X.; Burgueño, A.; Serra, C.; Martín-Vide, J.; Gómez, L. Distributions of long dry spells in the Iberian Peninsula, years 1951–1990. Int. J. Climatol. 2009, 515–530. [CrossRef]

65. Mathugama, S.C.; Peiris, T.S.G. Critical Evaluation of Dry Spell Research. Int. J. Basic Appl. Sci. 2011, 11, 153–160.

66. Raymond, F.; Ullmann, A.; Camberlin, P.; Drobinski, P.; Chateau Smith, C. Extreme dry spell detection and climatology over the Mediterranean Basin during the wet season. Geophys. Res. Lett. 2016, 43, 7196–7204. [CrossRef]

67. Kim, H.Y. Analysis of the definition and characteristics of drought in Korea—What are the causes and countermeasures of drought? J. Korean Soc. Civ. Eng. 1995, 43, 23–32.
97. Ottinger, M.; Kuenzer, C. Spaceborne L-Band Synthetic Aperture Radar Data for Geoscientific Analyses in Coastal Land Applications: A Review. *Remote Sens.* 2020, *12*, 2228. [CrossRef]

98. Lee, Y.J.; Kim, G.Y.; Lee, Y.G.; Jeong, J.H.; Choi, M.H. Introduction and development direction of various soil moisture measurement methods. *Water Future* 2020, *53*, 126–134.

99. Dorigo, W.; Himmelbauer, I.; Aberer, D.; Schremmer, L.; Petrakovic, I.; Zappa, L.; Preimesberger, W.; Xaver, A.; Annor, F.; Ardö, J.; et al. The International Soil Moisture Network: Serving Earth system science for over a decade. *Hydrol. Earth Syst. Sci.* 2021, *25*, 5749–5804.

100. Hou, A.Y.; Kakar, R.K.; Neeck, S.; Azarbarzin, A.A.; Kummerow, C.D.; Kojima, M.; Oki, R.; Nakamura, K.; Iguchi, T. The Global Precipitation Measurement Mission. *Bull. Am. Meteorol. Soc.* 2014, *95*, 701–722. [CrossRef]