ISPRS Journal of Photogrammetry and Remote Sensing 211 (2024) 244–261

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Forecasting corn NDVI through AI-based approaches using sentinel 2 image time series

A. Farbo [a,*], F. Sarvia [a], S. De Petris [a], V. Basile [b], E. Borgogno-Mondino [a]

[a] Department of Agriculture, Forestry and Food Sciences, University of Torino, Grugliasco (TO) 10095, Italy
[b] Computer Science Department, University of Turin, Via Verdi 8, 10124 Turin, Italy

ARTICLE INFO

Keywords:
LSTM
NDVI forecasting
Corn
Generalization Capability
Decision Support System

ABSTRACT

Precision Agriculture (PA) has revolutionized crop management by leveraging information technology, satellite positioning data, and remote sensing. One crucial component in PA applications is the Normalized Difference Vegetation Index (NDVI), which offers valuable insights into crop vigor and health. However, discontinuity of optical satellite acquisitions related to cloud cover and the huge load of the required processing time pose challenges to real-time applications. NDVI prediction emerges as an innovative solution to address these limitations. It allows for proactive decision-making by providing accurate estimates, enabling farmers and land managers to plan essential agronomic activities such as irrigation, fertilization, and pest control, based on anticipated future conditions. This study introduces an Artificial Neural Network (ANN) model incorporating NDVI, Normalized Difference Water Index (NDWI), temperatures, and precipitation as predictive variables. The model employs a novel time series slicing algorithm, Boosting Adaptive Time Series Slicer (BATS), to enhance the input training dataset's variability, presenting the model with a broader range of examples. A 2-Bidirectional Long Short-Term Memory (LSTM) forecasting model was developed to predict future NDVI values over short and medium-term horizons. The study area used to train, test and validate the ANN corresponds to a diverse landscape of cultivated corn fields located in Piemonte (NW-Italy). Results showed that NDVI future estimates were accurate; considering three time horizons for predictions (5, 10, and 15 days) RMSE values resulted to be 0.028, 0.038 and 0.050, respectively. Additionally, ablation tests proved that the most important variable for enhancing the model's accuracy is the NDWI, and the most useful timesteps are the four most recent ones. To preliminary investigate the capability of the ANN to operate over a wider and different area it was applied over the entire Europe, using the LUCAS dataset as reference map to locate corn fields. Results show RMSE of 0.062, 0.083 and 0.105 for the 5, 10 and 15 days forecasting horizons, respectively. The methodology proposed in this paper can be a possible alternative to more ordinary approaches for NDVI forecasting that nowadays appears to be a fundamental step for a proactive precision agriculture where crop management can be significantly improved. Future developments should explore the use of sequence-to-sequence ANNs to predict the development of multiple spectral indices over multiple crop types simultaneously.

1. Introduction

Precision Agriculture (PA) is a pivotal approach for enhancing crop management and resources optimization (Sharma et al., 2021). It involves technologies like geographical information systems, satellite positioning and remote sensing to improve crop yield, economize input utilization, and reduce environmental impacts. Normalized Difference Vegetation Index (NDVI) is a widely used remotely sensed tool for estimating key crop metrics like biomass and vegetation health status, which are essential for precise agricultural planning and budgeting, significantly influencing crop yield projections and environmental management (Houborg and McCabe, 2016).

Despite the widespread adoption of NDVI, the intermittent nature of satellite acquisitions negatively affects crop monitoring and related agricultural deductions. Satellite data collection often encounters interruptions due to factors like cloud cover during imaging and the limited frequency of satellite passes over a given area (Li et al., 2021). These interruptions can hinder the timely assessment of crop conditions, leading to challenges in optimizing resource allocation and decision-making for agronomic activities.

* Corresponding author.
E-mail address: alessandro.farbo@unito.it (A. Farbo).

https://doi.org/10.1016/j.isprsjprs.2024.04.011
Received 4 January 2024; Received in revised form 15 March 2024; Accepted 12 April 2024
Available online 17 April 2024

In this regard, NDVI forecasting offers a game-changing solution. It offers a pivotal advantage by enabling the anticipation of agronomic activities. Rather than relying solely on historical or real-time NDVI data from satellite imagery, forecasting allows for a proactive decision-making process. With accurate predictions, farmers and land managers can preemptively plan and implement essential agronomic practices, such as irrigation, fertilization, and pest control, based on anticipated future conditions (Blaes et al., 2016, Radočaj et al., 2023, Soccolini and Vizzari, 2023). This proactive approach can significantly enhance the efficiency of resource allocation and crop management, ultimately contributing to improved agricultural productivity and sustainability. Moreover, in the droughts mitigation context, NDVI forecasting is a particularly useful tool for macro-scale regional water management. In fact, by offering predictive insights into vegetation health and land cover changes, NDVI forecasting empowers regional authorities and policy-makers to make informed decisions about water allocation, conservation measures, and the overall management of critical water resources (Maselli et al., 2020, Nouri et al., 2014, Poudel et al., 2021). During periods of severe drought, where water availability becomes a pressing concern, the capability to anticipate vegetation stress and water demand through NDVI forecasts plays a crucial role for optimizing water distribution, mitigating the impacts of drought, and safeguarding the ecological balance of ecosystems. NDVI forecasting at macro-scale can substantially improve regional resilience against water scarcity. Consequently, accurate NDVI forecasting has emerged as a pivotal element for enhancing the reliability and utility of NDVI as an agricultural indicator.

Possible approaches in forecasting tasks are based on Machine Learning (ML) and Deep Learning (DL). These methodologies utilize historical data to predict future values. Common ML algorithms are Random Forest (RF) and Support Vector Machine (SVM). They analyze patterns in data to make predictions, focusing on relationships within the features (Parmar et al., 2019, Mountrakis et al., 2011). RF and SVM have been widely used in time series analysis with notable success (de Castro et al., 2020, Son et al., 2020). However, it's important to note that these algorithms were not specifically designed to capture time dependencies such as autoregressive, seasonal, and cyclic phenomena that define time series. Indeed, each observation is considered an independent variable, not as part of a sequence. Additionally, RF and SVM are one-output models, which make them natively not suited for multi-step forecast problems (de Castro et al., 2020; Huang et al., 2017; Reuß et al., 2021; Roy et al., 2022). Conversely, DL architectures like Recurrent Neural Networks (RNNs) can capture more complex patterns, including short and long temporal dependencies, making them particularly suited for forecasting tasks that involve time-dependent data and provide multiple outputs (Torres et al., 2021, Yamak et al., 2020). However, it should be mentioned that training and fine tuning Artificial Neural Networks (ANNs) require large datasets to avoid overfitting and can be computationally intensive due to the amount of trainable parameters and dataset dimension. Therefore, depending on the task complexity and data availability, considering ML algorithms instead of DL ones could be the preferred way.

NDVI forecasting using DL typically operates at two distinct scales: large-area forecasts and field-level forecasts. Large-area forecasts focus on predicting the average NDVI values over extensive, relatively homogenous regions, such as irrigation districts or forest covers (Fathollahi et al., 2023). In contrast, field-level forecasts aim to predict NDVI values for individual fields covering smaller, heterogeneous areas that are more representative of actual field conditions (Cavalli et al., 2023). Field-level forecasts offer valuable insights for farmers, especially in a precision agriculture framework. However, predicting NDVI at this level presents unique challenges, given the inherent noise and variability in field-level NDVI data. Additionally, the forecasting horizon significantly impacts the complexity of the problem. Some methods focus on predicting only the immediate next NDVI value, while others strive to provide multi-step forecasts over extended periods (Ahmad et al., 2023, Cavalli et al., 2023). In this framework, state-of-the-art methodologies are significantly limited by three main problems: a) limited forecasting horizons, b) moderate prediction accuracies and c) ignoring complex patterns, including short and long temporal dependencies.

In this work, we address these research gaps by developing an ANN which incorporates a Long Short-Term Memory (LSTM)-based forecasting model. The research focuses on field-level NDVI prediction for short and medium-terms, specifically targeting forecasts at 5, 10, and 15 days ahead. Previous studies have effectively utilized historical NDVI, temperatures, and precipitation data for predicting future NDVI values, underscoring the importance of these variables in capturing the essential climatic and phenological influences on vegetation health (Cavalli et al., 2023). However, it is worth mentioning that Normalized Difference Water Index (NDWI) is able to reflect crop moisture availability, we included for the first time the water-related spectral index in our model to provide a more comprehensive assessment of vegetation health and water stress levels (Gao, 1996, Szabo et al., 2016). For this scope, Sentinel-2 (S2) data were chosen as the best suited for the aims of the study thanks to the high spatial, temporal, and spectral resolution. The study area selected for this work is a complex landscape of cultivated corn fields within the Piemonte Region (NW Italy), covering more than 160,000 total ha and extending across a study area of over 13000 km^2. Finally, the ANN is tested on a dataset covering the entire Europe in order to evaluate the ANN transferability to other ecozones.

2. Materials

2.1. Study Area

The Piemonte Region (NW-Italy) is made of 8 provinces showing different geomorphological and climatic features, leading to a complex agricultural landscape (Ghilardi et al., 2023, Sarvia et al., 2021). A Walter and Lieth climate diagram is reported in Figure 1 to better describe the average climate in Piemonte region.

It is one of Italy's main corn-producing regions, accounting for 23% of the total cultivated corn area and 26% of the total production, underscoring its significance in national agriculture (ISTAT - Coltivazioni: Cereali, legumi, radici bulbi e tuberi, last access 7 February 24) . The study area (AOI) develops over the Torino, Vercelli, Cuneo, Alessandria, and Asti provinces (Figure 2). Within AOI, corn fields were surveyed in the period 2018-2022. Prior to analysis, fields under 0.1 ha and with a shape index (SI, equation (1) greater than 3 were excluded to reduce mixed pixel effects resulting in a total of 130,401 corn fields (more than 160,000 ha) (Sarvia et al., 2021). This extensive exploration over multiple years across a large area is aimed at capturing a high spatial-temporal variability.

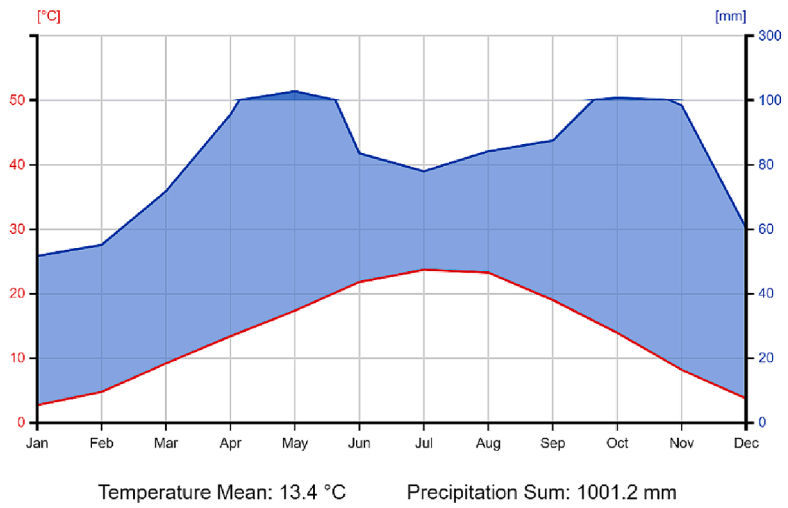
$$SI = \frac{P}{2\sqrt{\pi A}} \tag{1}$$

where P and A are the polygon perimeter and area, respectively.

In Northern Italy, corn cultivation includes two types of corn with distinct growing cycles: grains corn and silage corn. The first type, characterized by a longer growing cycle, is planted in mid-April to early May. Germination and emergence occur in May (BBCH 00-09), followed by vegetative growth until June (BBCH 10-39), then heading and flowering from mid to late July (BBCH 51-69), and ear development and ripening in August and September (BBCH 71-89), concluding in harvest from late September to October. The second type, with a shorter growing cycle, is sown in June to July post the winter wheat harvest. The silage corn rapidly progresses from germination to vegetative growth in July, with heading, flowering, and ear development, aiming for harvest by late September to October (Berti et al., 2019, Sarvia et al., 2021).

2.2. Rainfall

Water scarcity significantly constraints plant growth, development,



Month	Temp	Precip
Jan	2.7	51.7
Feb	4.8	55.1
Mar	9.2	71.8
Apr	13.4	95.5
May	17.4	127.9
Jun	21.8	83.5
Jul	23.7	78.0
Aug	23.3	84.2
Sep	19.0	87.5
Oct	13.9	107.5
Nov	8.2	98.4
Dec	3.7	60.3

Figure 1. Walter and Lieth climate diagram for the Piemonte region from 2010 to 2022.

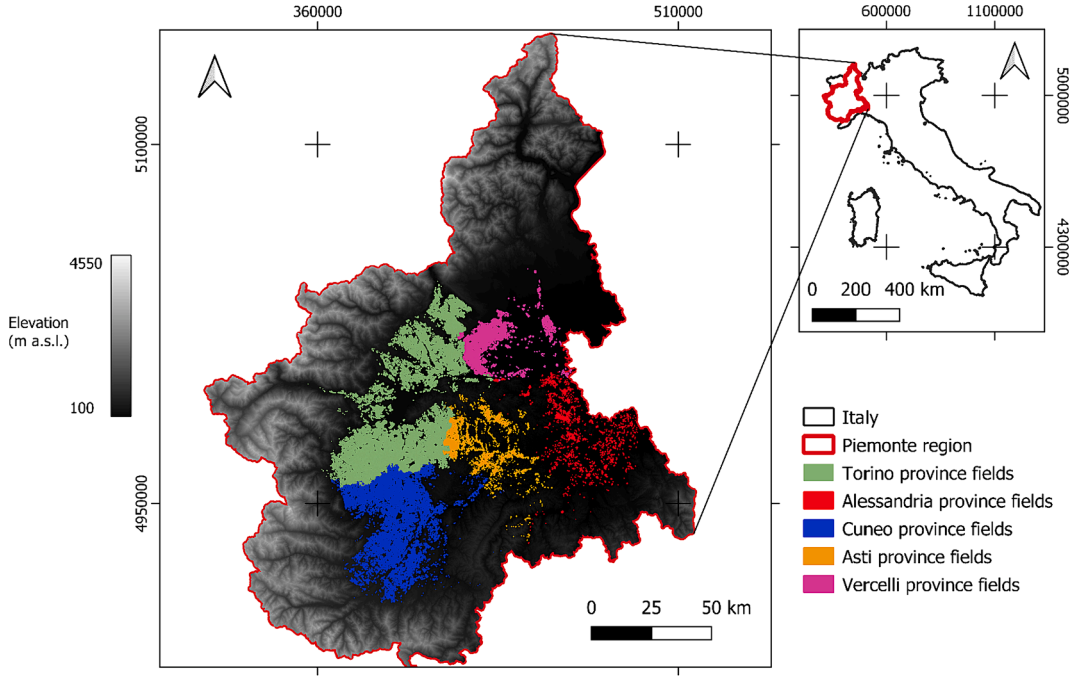


Figure 2. Corn fields collected in the Piemonte region (red outline) over different provinces (Reference system: WGS 84/UTM zone 32N, EPSG: 32632).

and yield (Rockström et al., 2010), particularly during periods of water deficit stress, commonly referred to as drought stress (Efeoğlu et al., 2009). Thus, precipitation plays a key role in biomass growth and phenology models. The Global Precipitation Measurement mission provides hourly calibrated precipitation data (Pcal) with a spatial resolution of 10 km. From these records, daily accumulated precipitation was computed and the yearly TS was derived using Google Earth Engine. To align Pcal data to the S2 spatial resolution (10 m), a nearest-neighbor resampling was applied, ensuring spatial coherence.

2.3. Air Temperature

Temperatures play a second key role in modelling the crop's phenological cycle and biomass production (Lizaso et al., 2018). As a result, maximum and minimum thermal data were sourced from a network of 116 meteorological stations (Figure 3) affiliated with the regional meteorological network (accessible at www.arpa.piemonte.it). These data were collected over the period spanning from January 1, 2018, to December 31, 2022, with daily frequency.

Growing Degree Days (GDD) is a recognized metric for characterizing biomass growth during the phenological season, taking into account the temperature's impact on crops, which is closely related to their thermal efficiency. In this study, daily GDD values were calculated using the formula from (Mcmaster, 1997). Daily temperature records were obtained from the meteorological station closest to the considered field and used to compute daily GDD according to eq. (2):

$$GDD_i = \left\{ max\left(\frac{T^i_{max} + T^i_{min}}{2} - T_{BASE}, 0 \right) \right\} \quad (2)$$

where T^i_{max} and T^i_{min} are the daily (i-th) maximum and minimum temperatures and T_{BASE} stands for the nominal base temperature, which varies depending on the crop and defines the minimum temperature threshold necessary to trigger the phenological development of

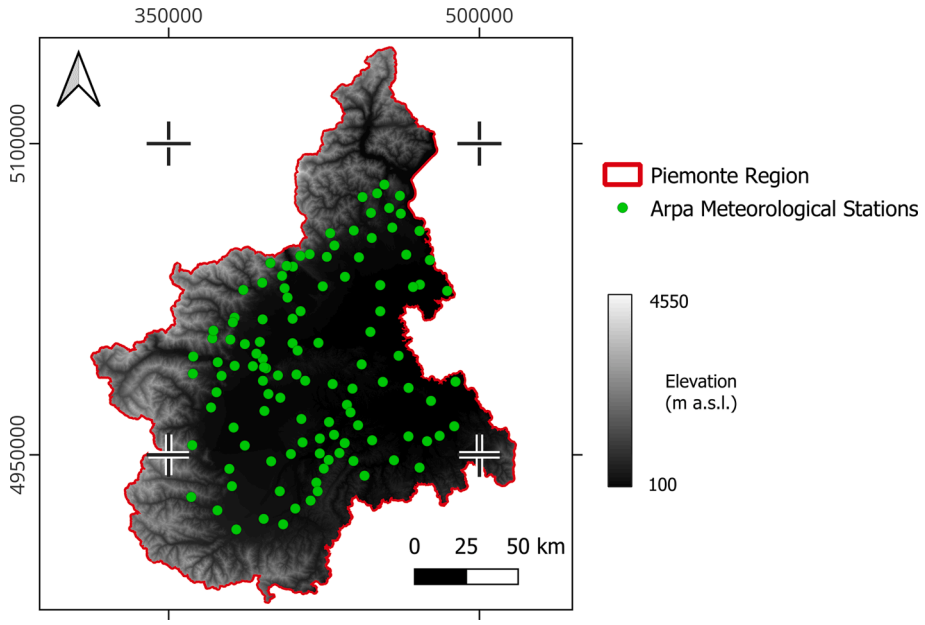


Figure 3. Arpa Meteorological Stations taken into account for the study. (Reference System: WGS84/UTM 32 N, EPSG: 32632).

vegetation (Salazar-Gutierrez et al., 2013). T_{BASE} selection is a largely discussed topic in literature (Gill et al., 2014; Giolo et al., 2021; Miller et al., 2001; Salazar-Gutierrez et al., 2013). T_{BASE} is a very complex factor depending on both the plant species and the cultivar (Salazar-Gutierrez et al., 2013). In this work, the T_{BASE} was set equal to 10 °C according to (Hou et al., 2014).

Finally, GDDs pertaining to each meteorological station were spatialized using the Voronoi Polygons algorithm (Gold et al., 1997). The resulting polygons were then rasterized with a 10 m spatial resolution in order to match the S2 one.

2.4. Satellite data

Despite the current availability of Earth observation satellite imagery, not all acquisitions are suitable for PA and agronomic applications. In the context of meeting the specific requirements for precision farming, several fundamental operational prerequisites must be addressed: (a) the imagery should possess adequate geometric resolution relative to the size of agricultural fields, ensuring that the details of individual fields can be effectively captured; (b) a high temporal resolution is essential to monitor the dynamic phenological phases of crops over time; (c) the spectral bands must exhibit sensitivity to crucial crop parameters, including biomass, photosynthetic activity, and leaf water content; and (d) the cost associated with acquiring this data should be aligned with the budget constraints of the agronomic sector, ideally being provided free of charge. These stringent criteria are crucial for precision farming applications, ensuring that the data used is not only accessible, but also highly relevant to the specific needs of agronomic and PA practices. In this context, the S2 mission stands out in meeting these criteria, as it offers a nominal time resolution of 5 days, although this is somewhat dependent on cloud cover. Moreover, the images are distributed free of cost and come pre-calibrated to at-the-ground reflectance and at a maximum geometric resolution of 10 meters. These attributes render them exceptionally compatible with the objectives of this study. The European Union's S2 mission is equipped with multispectral optical sensors that span the range from 400 to 2500 nm, including the visible to medium infrared spectrum.

Considering the extensive geographical coverage and the need for multi-year analysis in this study, Google Earth Engine was chosen as the platform for processing already geometrically and radiometrically corrected reflectance values from the S2 Harmonized level 2-A data. This product level has a nominal positional accuracy of 3m and a radiometric resolution of 12 bits (Gascon et al., 2017). The combination of spectral bands allows for the generation of spectral indices that can effectively describe distinct crop behaviours. Specifically, for this study, the NDVI and the NDWI have been selected. The focus on NDVI and NDWI aligns with the objective to manage computational resources efficiently while ensuring the predictive model remains focused on key parameters (Johnson and Kjell, 2019, Bargagli Stoffi et al., 2022). This approach minimizes data redundancy and enhances operational efficiency, enabling accurate analysis of crop status using the available spectral data. In particular, four spectral bands were selected for NDVI and NDWI computation: green (560 nm ± 35 nm) having a geometric resolution of 10 m; red (665 nm ± 30 nm) having a geometric resolution of 10 m; NIR (842 nm ± 115 nm) having a geometric resolution of 10 m and SWIR1 (1610 nm ± 90 nm) having a geometric resolution of 20 m. The latter was resampled using nearest neighbour method to the 10 m geometric resolution. This decision aligns with the study's requirements for efficient and scalable analysis over large areas and extended time periods, facilitating the comprehensive assessment of the datasets.

NDVI is a well-established spectral index in the scientific literature, renowned for its ability to provide insights into vegetation characteristics (Borgogno-Mondino et al., 2022). It is particularly valuable for phenological analysis (Farbo et al., 2022), ecosystem characterization (Orusa et al., 2023), crop yield prediction (Mkhabela et al., 2011), assessment of crop biomass (Meng et al., 2013), monitoring of urban green areas and heat islands (Grover and Singh, 2015), and the development of insurance strategies in agriculture (Sarvia et al., 2020). In this research, NDVI was utilized as a predictor for phenology and biomass estimation according to eq. (3):

$$NDVI = \frac{b8 - b4}{b8 + b4} \tag{3}$$

where $b8$ and $b4$ correspond to near infrared and red S2 spectral bands, respectively. Several other spectral indices can be effectively used in the PA context like the Normalized Difference RedEdge (NDRE) and the Enhanced Vegetation Index (EVI). However, both have limitations due to some intrinsic aspects of the spectral bands used for their computation. Specifically, the NDRE requires the use of the Red-Edge band, which for S2 has a coarser geometric resolution (20 m) compared to the Red and NIR spectral bands (10 m). This resolution is less suitable for the

small and fragmented fields typical of Italian agriculture. Additionally, depending on the red-edge band used to compute NDRE different results can be obtained, leaving a high degree of discretionality to the user and, therefore, a lower robustness about repeatability of the method. Concerning the EVI, this index incorporates the Blue spectral band, which is notoriously affected by atmospheric noise (Okin and Gu, 2015). It has been found that the noise of the EVI index increases with increasing EVI values (i.e., high Leaf Area Index - LAI) (Miura et al., 2000). Conversely, NDVI is more affected by noise at very low values (indicative of almost no vegetation) (Miura et al., 2000). Additionally, EVI was first introduced during the MODIS mission, and a set of empirical coefficients were developed specifically for this mission (Huete et al., 1994). Using EVI in a different mission without adjusting the parameters could introduce errors that are difficult to track, as reported for other indices (Zhen et al., 2023). Moreover, EVI is considered a viable solution to the well-known NDVI saturation effect only at high LAI values, such as those found in tropical forests. However, it should also be noted that the NDVI saturation effect becomes problematic at very high LAI values (>4), as reported by Potithep et al. (Potithep et al., 2013). Conversely, studies have found that such high LAI levels rarely occur in corn fields (Hosseini et al., 2015; Pacheco et al., 2001; Fei et al., 2012). Considering all these aspects, NDVI emerges as the most suitable spectral index for analysis in this study.

Within the agricultural landscape of northern Italy, a substantial portion of the corn cultivation, exceeding 52%, relies on irrigation practices. This significant reliance on irrigation implies that precipitation data alone may not be sufficient of estimating soil water availability. A noteworthy case in point is the Piemonte region, where approximately 70% of corn cultivation is under irrigation (Regione Piemonte: Censimenti generali dell'agricoltura - dati di sintesi, last access 13 October 23). Therefore, an additional spectral index able to relate to water content has been considered (i.e. NDWI).

NDWI plays a pivotal role in assessing agricultural landscapes. NDWI primarily focuses on the presence of water and its variations, making it valuable for monitoring and managing water resources in agricultural regions. This index is especially sensitive to changes in water content, which can provide critical information for irrigation management, drought detection, and soil moisture assessment in agricultural settings (Gao, 1996, Szabo et al., 2016). Consequently, NDWI was also selected in this study to enhance the understanding of water-related dynamics in corn fields. NDWI was computed according to eq. (4):

$$NDWI = \frac{b3 - b11}{b3 + b11} \tag{4}$$

where $b3$ and $b11$ correspond to green and short-wave infrared-1 S2 spectral bands, respectively.

NDVI and NDWI time series (TS) were generated, recognizing that Sentinel-2 data are prone to interference from clouds and shadows, which can alter the accuracy of the derived indices. To address this issue, the Scene Classification Layer (SCL) was employed to identify and exclude pixels compromised by clouds, snow, and shadows. This preprocessing step is critical for minimizing the influence of atmospheric conditions on the TS. After this correction, TS were regularized to a 5-day frequency, with linear interpolation used to fill gaps in the time domain. Satellite-based TS are notoriously noisy, as acknowledged in numerous studies (Hird and McDermid, 2009). To isolate the noise-free trend within TS, a first-order Savitzky-Golay smoothing algorithm with a five-observation window was used (Michishita et al., 2014). This smoothing was exclusively applied to the NDVI (NDVIs), as it is the target variable for the ANN forecasting. Conversely, the unsmoothed NDVI and NDWI were utilized as input variables because the Savitzky-Golay smoothing method results in the loss of the initial and final observations, making the use of smoothed TS as ANN inputs unsuitable in practical and real-time applications. This misalignment between the input and target data processing is necessary to train the ANN to process previous noisy S2 TS to forecast NDVI crop trends.

3. Methods

3.1. Data processing

This work was aimed at improving NDVI value prediction for PA applications and land planning management. The workflow, illustrated in Fig. 4, integrates meteorological and satellite data through a structured process comprising three key steps: (i) data preparation was intended for generating the appropriate multivariate TS serving as inputs for the ANN; (ii) ANN design, training, tuning and testing was aimed at verifying actual potentialities of the explored AI-based approach to NDVI prediction. This step involves different phases outlined in sections 3.3, 3.3.1, and 3.3.2; (iii) ANN generalization test was intended to verify the capability of a locally trained ANN to predict NDVI value out of its comfort zone. This was achieved with reference to the entire Europe exploiting corn crop maps from the LUCAS dataset. The description of data processing details is provided in the subsequent sections, offering comprehensive insights into each step of the proposed methodology.

3.1.1. Time series harmonization

Harmonizing spectral indices, precipitation and GDD time series is a necessary step for comparing and relating these factors. Therefore, spectral, precipitation and thermal time series were computed for each field. Furthermore, since corn growing season may vary significantly between fields, a common period (from emergence to senescence) must be locally defined to make comparable different corn fields. For this task, the Start of Season (SOS) and the End of Season (EOS) for each field were extracted from NDVIs TS according to the following procedure:

- NDVImax (day of maximum NDVI value) identification: corn usually express the maximum NDVI values between mid-June and mid-August, therefore this period has been used to identify the day in which the maximum NDVI value occurs;
- SOS identification: the first lowest value below a defined threshold was found in a backward-looking process from NDVImax to the start of NDVIs TS. Initially the threshold was set to 0.3, indicating a non-vegetated moment in the field. However, some fields might be managed with no-tillage practices and the bare soil might not be present. Therefore, to identify the SOS in unconventional managed fields, the process was repeated with an increased threshold of 0.05 until 0.5. Once SOS was identified, all the previous data including SOS in NDVIs TS were removed;
- EOS identification: a similar process used for SOS identification was used but with a forward-looking instead of a backward-looking process. Similarly, all the subsequent data in NDVIs TS after EOS were removed.

Once the NDVIs TS for each field was processed and the non-corn related data removed, all the remaining TS (NDVI, NDWI, Pcal and GDD) were filtered based on NDVIs TS and all data outside the SOS-EOS interval were removed for each field individually.

Subsequently, Pcal and GDD data were accumulated for each corn field using equations (5) and (6), respectively, resulting in the cumulative Global Precipitation Measurement (GPM) and Accumulated Growing Degree Days (AGDD) for each day (i-th) up to the last available day (n):

$$GPM = \sum_{i=1}^{n} Pcal_i \tag{5}$$

$$\mathrm{AGDD} = \sum_{i=1}^{n} GDD_i \tag{6}$$

To ensure consistency in the time series data, they were harmonized to match the 5-day timestep of the S2 TS. Finally, for each field, the Day of

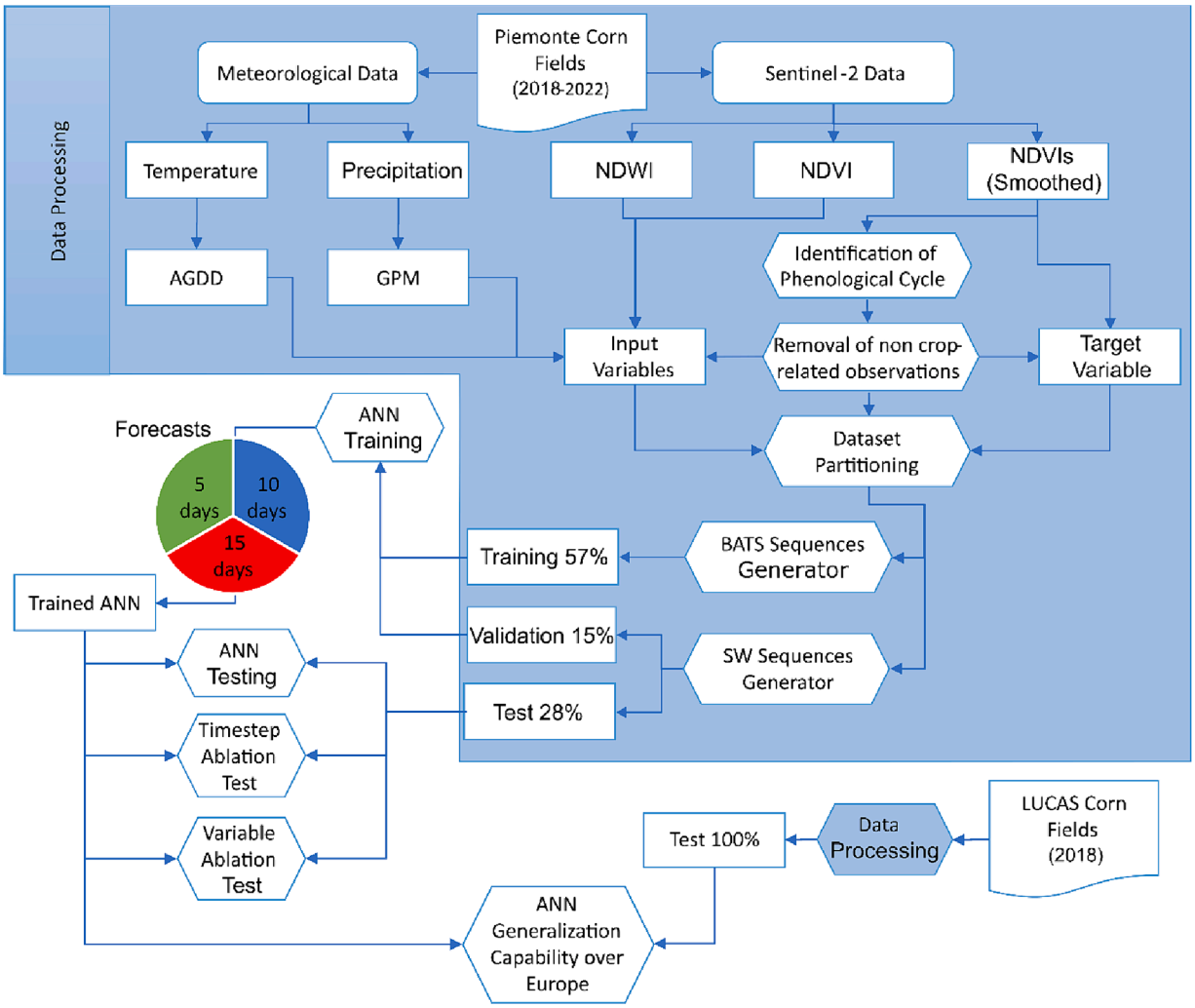


Figure 4. Schematic workflow of data processing and deep learning processes.

the Year (DOY) TS was retrieved, indicating the specific day on which each observation occurred.

At the conclusion of the time series harmonization process, the input variables exhibit varying scales, potentially posing a challenge during the ANN training phase. It is a well-established practice that ANNs tend to perform more effectively when all input and target variables are standardized or normalized (Djordjević et al., 2022). Finally, the six variables (NDVIs, NDVI, NDWI, AGDD, GPM, and DOY) were individually subjected to normalization using a min-max scaler. This preprocessing step ensures that the variables are brought to a consistent scale, enhancing the ANN's ability to learn and make accurate predictions.

3.1.2. *Dataset partitioning*

In this study, the dataset was partitioned into three subsets: 57% for training, 15% for validation, and 28% for testing. To ensure a robust estimation of model performance, Monte Carlo cross-validation was employed, using five distinct seeds to randomize the data division process (Xu and Liang, 2001).

3.1.3. *Minibatch generation*

The proposed model was developed to generate NDVI predictions for corn at any point during the phenological cycle, necessitating the adoption of a mini-batch approach. To begin with, the input and target variables were defined: in this study, the input data consisted of the previous 10 observations (equivalent to 50 days) of NDVI, NDWI, AGDD, GPM, and DOY. Meanwhile, the target data consisted of the subsequent three observations (i.e. 5, 10, and 15 days) of NDVIs. NDVIs was chosen as target variable to train the ANN to predict de-noised NDVI values.

Considering ten past observations (50 days) as input data would prevent the ANN to be applied at the beginning of the corn growing period. To avoid this problem, two different approaches were used: a Sliding Window (SW) was used to cut the TS for the test and validation dataset and a Boosting-Adaptive Timeseries Slicer (BATS) was used to cut the TS for the train dataset. Both SW and BATS ultimately produce mini-batches of ten timesteps for the input and three for the target, however the content of the mini-batches differs significantly. In Figure 5 a graphical representation of the two mechanisms is reported.

SW was designed to slide through TS generating mini-batches. Initially SW considers the first three observations and zero-pads the seven empty spaces. As it progresses, the number of considered observations increases while the number of padded empty spaces decreases. This continues until the mini-batches consist solely of actual observations. It was found out that pre-padding led to better results compared to post-padding in Recurrent Neural Network (Dwarampudi and Reddy, 2019). Therefore, pre-padding was used to make all minibatches the same length (i.e. ten timesteps).

BATS was designed to have 7 sliding windows that use different sequence lengths to generate the ten timesteps mini-batches. Specifically, the sliding windows consider a minimum of three observations up to a maximum of ten. All the mini-batches were pre-padded similarly to SW.

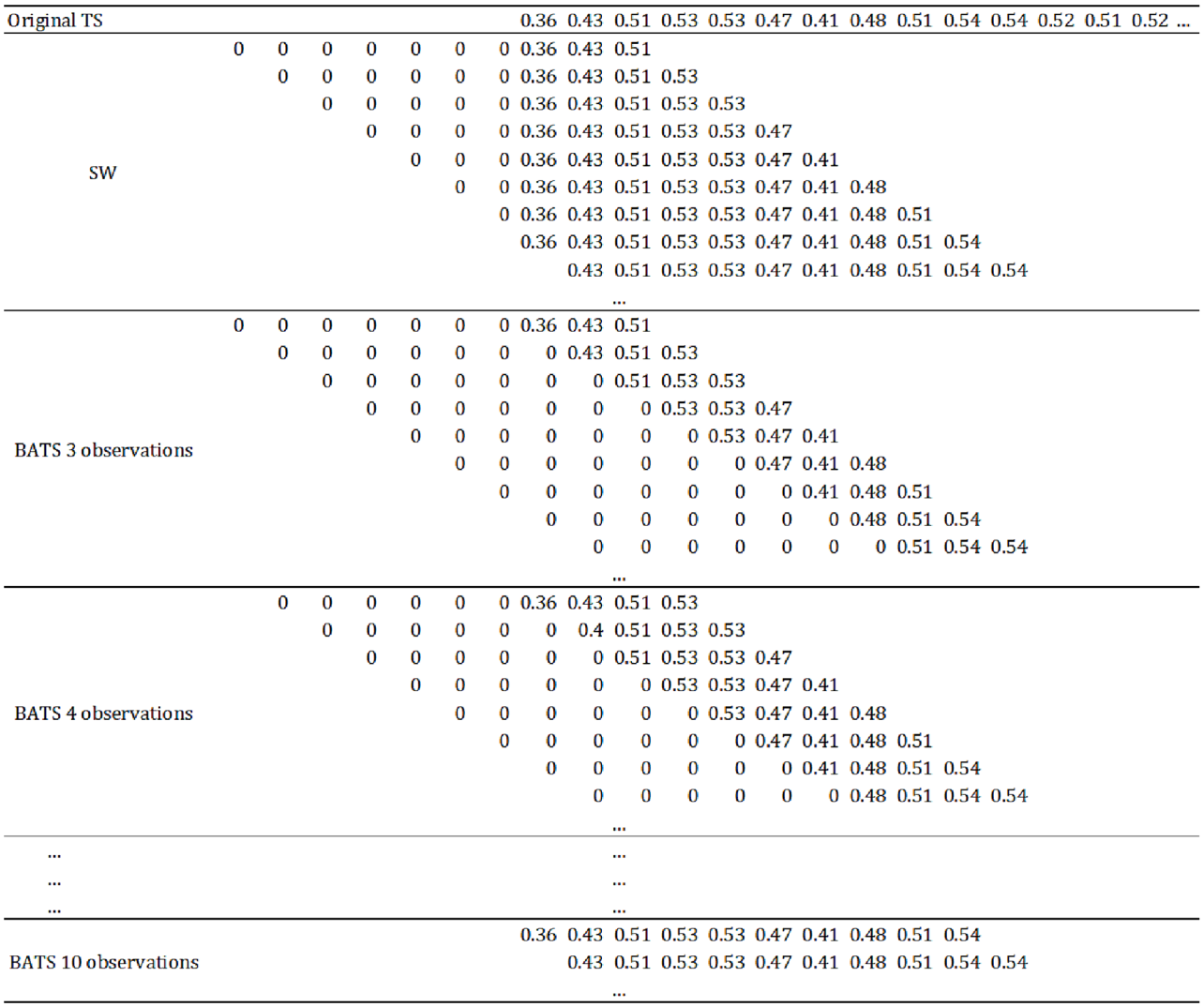


Figure 5. Graphical representation of SW and BATS approaches in mini-batch generation from a generic TS.

ANN was trained utilizing BATS data as input to ensure robust training even under circumstances where data availability is constrained, such as during the initial stages of the growth cycle. For the ANN validation and testing phase, data were processed with the SW approach. This strategic choice was made to enable the ANN to harness all available information for making accurate predictions when a wealth of data is accessible. Conversely, it ensures that the ANN remains proficient even in scenarios characterized by limited data availability during the early phases of the cycle. In essence, the ANN was trained to adapt and perform effectively under diverse data conditions, ranging from data-rich to data-scarce situations.

3.2. ANN settings

The forecasting method is carried out at field-level for each provided sequence. The ANN was designed with 2-Bidirectional Long Short-Term Memory (BiLSTM) layers. The employed ANN is a type of RNN that integrates two LSTM units at each time step. LSTM is a specialized recurrent hidden unit designed to capture and learn long and short-term dependencies, as originally proposed by Hochreiter and Schmidhuber (Hochreiter and Schmidhuber, 1997). The core of LSTM units revolves around the concept of memory cells, which maintain temporal dependencies. Within an LSTM unit, there exist three critical gates: the input gate denoted as i_t, the forget gate as f_t, and the output gate as o_t. At any given time step t, the LSTM unit processes the input x_t along with the previous hidden state h_{t-1}, combining them under the regulation of a hyperbolic tangent function (eq. (7)) as follows:

$$g_t = tanh\left(W_{xg}x_t + W_{hg}h_{t-1} + b_g\right) \tag{7}$$

where W_{xg} represents the input weights, W_{hg} the recurrent weights, and b_g the bias term. The input gate i_t plays a pivotal role in determining what information should be stored in the memory cell, and it operates through a sigmoid function (eq. (8)):

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{8}$$

where W_{xi} represents the input weights, W_{hi} the recurrent weights, and b_i the bias term. The forget gate f_t serves to decide what content in the existing memory cell should be retained or forgotten (eq. (9)):

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + b_f\right) \tag{9}$$

where W_{xf} represents the input weights, W_{hf} the recurrent weights, and b_f the bias term. Information is then updated in the memory cell C_t by incorporating inputs from both the input gate and the forget gate, allowing for the integration of new information g_t while discarding part of the current memory content (eq. (10)):

$$C_t = f_t \odot C_{t-1} + i_t \odot g_t \tag{10}$$

Finally, the output (hidden) state h_t is computed by the output gate o_t and the updated memory cell C_t (eq. (11)):

$$h_t = o_t \odot tanh(C_t) \tag{11}$$

Where o_t is determined by yet another sigmoid function:

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \tag{12}$$

By combining two LSTM units in a bidirectional manner, a BiLSTM network is derived, enabling the analysis of past states from two different point of view: forward looking and backward looking. BiLSTM networks provide a comprehensive perspective on time series data, as they learn from sequences at each time step, thereby offering a holistic view of temporal dependencies (Bin et al., 2019, Graves et al., 2005).

ANN hyperparameters were fine-tuned with a systematic exploration of the hyperparameter space, guided by a range of potential configurations for layers and neurons as suggested by existing literature (Stepchenko and Chizhov, 2015, Reddy and Prasad, 2018, Ahmad et al., 2023, Cavalli et al., 2023). Ultimately, the best model was chosen as the one which reached the lowest overall RMSE from the cross validation procedure. Additionally, the ANNs were initialized with different seeds in order to minimize variability. The resulting network's structure includes an input layer that houses the predictors. In this case, these are mini batches of ten timesteps that include five variables (NDVI, NDWI, AGDD, GPM, and DOY). This is followed by two BiLSTM layers, each with 64 hidden units and a subsequent 20% dropout layer to prevent overfitting. A linear layer then calculates the predictions for the next three timesteps (i.e. 5, 10, 15 days) in the network output. The final layer, the output layer, contains the target forecasts. The ANN structure is reported in figure 6.

The ANN was trained with a batch size of 64. The Adam optimizer was used with a learning rate set to 1e-3 and Mean Squared Error (MSE) was used as the loss function. The total number of training epochs was set to 100. However, if the validation loss did not decrease for ten consecutive epochs, the training would have been halted, and the weights from the last epoch with decreased loss would have been restored.

3.3. Model performance assessment

The assessment of ANN performance was conducted by examining one timestep forecast at a time on the test datasets derived from the Monte Carlo cross validation. Specifically, the evaluation involved calculating the Root Mean Squared Error (RMSE) between the forecasts generated by the ANN and the corresponding target values for each timestep according to eq. (13):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{13}$$

where $\hat{y}_i$ and y_i are respectively the predicted and target i-th forecasts and n is the total number of forecasts. RMSE is used to measure the average magnitude of the errors, providing a sense of overall model performance.

Following this, to visually assess the output of the three linear neurons and investigate potential biases linked to NDVI values, density scatter plots and their corresponding linear models were generated.

For a more comprehensive evaluation of the model's performance throughout the phenological cycle, the RMSE and Mean Error (ME, eq. (14)) of each output linear neuron were calculated for each DOY. This approach also helped to identify periods where the model excelled or faced challenges compared to others.

$$ME = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \tag{14}$$

where $\hat{y}_i$ and y_i are respectively the predicted and target i-th forecasts and n is the total number of forecasts. ME is crucial for identifying any biases in the model, as it indicates whether the errors are systematically over or under predicting the target values.

From literature it was observed that the S2-retrieved NDVI uncertainty ranges from 0.01 to 0.07 (Borgogno-Mondino et al., 2016, De Petris et al., 2023). To enhance the understanding of the errors made by the ANN throughout the phenological cycle, the percentage of errors exceeding a cautious NDVI uncertainty threshold of 0.04 (i.e. significant errors) was computed for each forecasting horizon. It's important to note that this uncertainty pertains to the original NDVI, not the smoothed NDVI (NDVIs), which is the output of the ANN. However, since NDVIs is derived from the original NDVI, the uncertainty affecting the original NDVI also impacts the NDVIs to some extent.

3.3.1. Timesteps Ablation Test

The model's training involved the use of ten fixed timesteps, which were occasionally zero-padded through the BATS algorithm. This strategy aimed to enable the model to acquire the ability to make accurate NDVI forecasts during periods with abundant data (such as the mid-late growing season) as well as during periods characterized by limited data availability (such as the initial phase of the growing season where few past observations are present). However, it's worth noting that this approach could be further refined by determining the minimum number of timesteps necessary to achieve satisfactory results.

To address this optimization challenge, the test dataset was subjected to perturbations by incrementally zero-padding one timestep at a time. This process continued until the inputs provided to the ANN consisted solely of sequences of zeros, essentially offering no meaningful information. For each set of removed timesteps, the RMSE for each forecast was calculated and then compared to the RMSE obtained without perturbations. This systematic examination helped pinpoint the minimum number of timesteps required to maintain acceptable forecasting performance, thereby fine-tuning the model's efficiency.

3.3.2. Variables Ablation Test

Five input variables were used to train the ANN, however some doubts arise regarding the importance of all of them. Ordinarily, multivariate models use only highly significant variables and exclude the less important ones (Sauerbrei et al., 2020). It is good practice to check which variable is more important for the model performances in order to understand the underlying processes.

To determine less important variables, a similar approach to that outlined in section 3.3.1 was utilized. Specifically, each variable was

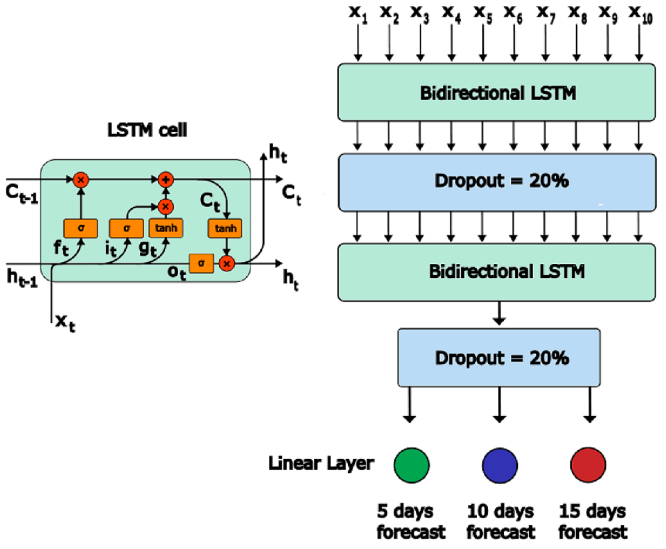


Figure 6. ANN structure representation. X_1-X_{10} are the inputs at each timestep.

individually subjected to zero-padding, and the corresponding perturbed RMSE was calculated. $\Delta RMSE_i$ was then computed according to eq. (15) to assess the impact of omitting each variable i:

$$\Delta RMSE_i = RMSE_i^p - RMSE_o \quad (15)$$

where $RMSE_i^p$ represents the RMSE observed when each variable i is zero-padded, effectively omitting its information from the ANN while $RMSE_o$ represent the overall RMSE. This analysis provides insights into the extent of errors introduced by the absence of individual variables, aiding in the identification of their relative importance in the modeling process.

3.3.3. ANN generalization capability

In literature, many works use ANNs and remote sensing or geographical data over relatively small study areas and are not tested over bigger and more diverse dataset (Small, 2021). Such limitation can result in well-trained ANNs that struggle to generalize effectively across more variable AOIs. In our work, the AOI is confined to the Piemonte region, as described in section 2.1. To mitigate this constraint, we collected data spanning five years, from 2018 to 2022, in order to capture a wide range of temporal and climatic variability. However, even with this extended dataset, there are inherent limitations, and it remains crucial to test the trained ANN against a completely different and more diverse dataset.

To address this need, we gathered 1002 corn fields from the year 2018 using the LUCAS dataset (Figure 7). The LUCAS (Land Use/Cover Area frame statistical Survey) dataset is an initiative organized by the European Statistical Office (EUROSTAT), designed to collect information on land cover and land use across all Member States of the European Union. This layer was used to detect the position of corn fields outside the study area (EU level). For each field in this dataset, we collected and computed NDVI, NDVIs, GPM, and DOY TS data, following the methodology outlined in sections 2.2, 2.3, 2.4 and 3.1.1.

However, obtaining data for Growing Degree Days (AGDD) presented a distinct challenge, as it was impractical to collect data from meteorological stations across the entire European region. To address this, we sourced daily maximum and minimum air temperature data at a 2-meter height from the ERA5 reanalysis dataset provided by ECMWF (European Center for Medium-Range Weather Forecasts). Previous studies have indicated that ERA5 air temperature data generally exhibit satisfactory levels of accuracy and precision. However, it's important to note that these levels can vary significantly depending on the specific study area (McNicholl et al., 2022, Tetzner et al., 2019). Consequently, ERA5 is considered a valuable tool, though it is important to anticipate the possibility of encountering increased errors during the inference phase in certain scenarios.

Finally, temperature data were converted from °Kelvin to °Celsius to compute AGDD as outlined in section 3.1.1.

All time series were normalized using previously derived min-max scalers for each variable. Subsequently, SW approach was employed to create mini batches for input and target data in the ANN.

Since GPM data over European dataset were approximately 2.5 to 3 times higher than the GPM values computed specifically for the Piemonte region, concerns arise regarding the accuracy of ANN in processing such heterogeneous data. In fact, feeding ANN with values outside the ones of calibration range could potentially lead to misinterpretations and erroneous forecasts. Therefore, the variable ablation test was performed to assess the importance of each variable and specifically evaluate the role of each variable. Based on the $\Delta RMSE_i$ analysis, it was decided to include only the most crucial variables in the subsequent transferability analysis.

To evaluate whether the ANN could achieve satisfactory results with varying amounts of timestep information, the timestep importance assessment as described in section 3.3.1 was conducted. Simultaneously, the transferability performance of the ANN was evaluated, following the methodology outlined in section 3.3.

Europe host a variety of climatic regions (Beck et al., 2018), which can significantly influence crop growth (Hatfield et al., 2018). Simply comparing the ANN's transferability results to those computed for the Piemonte region might obscure certain spatial patterns. Therefore, the LUCAS dataset was divided based on the present Köppen bioclimatic regions generated by Beck et al. (Beck et al., 2018), resulting in a total of seven groups (Figure 7):

- BSk - arid, steppe, cold: 34 fields;
- Csa - temperate, dry summer, hot summer: 19 fields;
- Csb - temperate, dry summer, warm summer: 19 fields;
- Cfa - temperate, no dry season, hot summer: 79 fields;
- Cfb - temperate, no dry season, warm summer: 396 fields;
- Dfa - cold, no dry season, hot summer: 58 fields;
- Dfb - cold, no dry season, warm summer: 398 fields.

Subsequently, the RMSE for the three-timestep forecasts was attributed to each bioclimatic region, enabling visualization of the model's performance across the entire European region.

4. Results

4.1. Model performance assessment

ANN was evaluated on the test datasets derived from the Monte Carlo cross validation (28% of total corn fields). The average RMSE for each of the three output neurons, along with the respective standard deviations

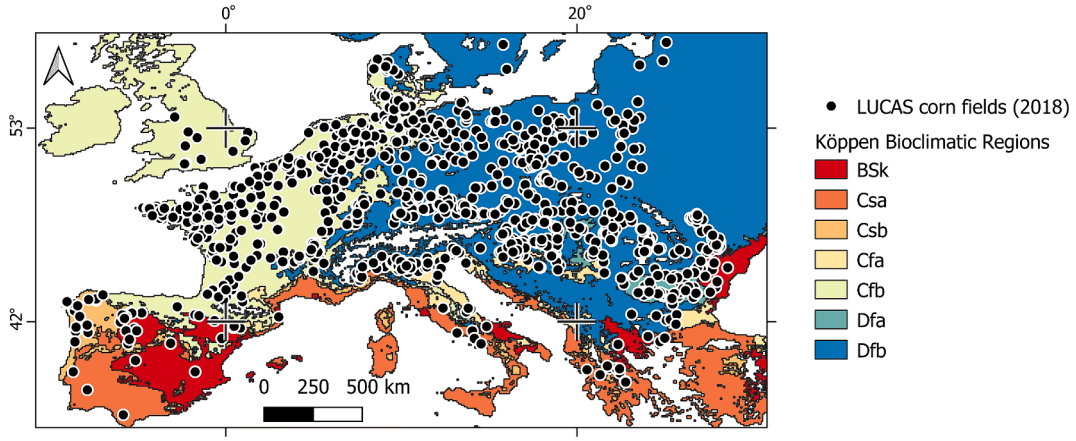


Figure 7. Spatial distribution of corn fields from LUCAS dataset in 2018 over different Köppen bioclimatic regions. (Reference System: WGS84, EPSG: 4326).

(STD), is reported in Table 1. Results from the Monte Carlo cross-validation, particularly the low STD, indicate the model's stability. Consequently, for simplicity, all subsequent results are based on a single ANN.

As expected, the ANN exhibits higher errors at increasing forecasts timestep. From literature it was observed that the S2-retrieved NDVI uncertainty ranges from 0.01 to 0.07 (Borgogno-Mondino et al., 2016). Notably, the ANN demonstrates the capability to generate forecasts up to 15 days in advance with errors comparable to NDVI theoretical uncertainty.

To gain further insights into potential biases between the predicted and target values, a first order linear model was constructed for each prediction horizon and is illustrated in Figure 8.

As previously noted in Table 1, the errors exhibit an increasing trend as the prediction horizon extends. Specifically, R^2 decreases from 0.96 at the 5-day horizon to 0.91 at the 15-day horizon. A closer examination of the linear model parameters provides further insights into the forecasting process: the intercept consistently exceeds 0, while the slope consistently falls below 1. Notably, at low NDVI values, the model tends to overestimate the actual values. This is particularly pronounced at the 15-day horizon, where values below 0.5 NDVI are significantly overestimated. This outcome aligns with expectations since the ANN was trained to forecast NDVI over vegetated areas. Low NDVI values are often associated with fields featuring sparse or small vegetation or non-photosynthetically active plants in senescent phases (Gao et al., 2020). Consequently, predicting NDVI under these conditions appears to be more challenging compared to when crops are in their full development phase.

Furthermore, scatterplot's points density highlights that only a small fraction of the forecasts deviates significantly from the bisector, while the majority of forecasts are concentrated around high NDVI values.

To explore the model's performance at different dates, the average NDVIs profile with its standard deviation was plotted, along the NDVI RMSE for the three temporal horizons for the entire forecasting period (Figure 9).

NDVIs first observations start with values about 0.5 despite the filtering of NDVIs TS to retain values above 0.3. The time series depicted in Figure 9 pertains to the forecasting period, which initiates from the fourth observation (i.e., 20 days) after the first recorded NDVIs value exceeding 0.3. It's important to note that during this period, corn undergoes a rapid growth phase, which readily results in NDVI values surpassing the 0.3 threshold.

Remarkably, the most elevated RMSE values for all the forecasting horizons are notably concentrated after DOY 230. This particular period corresponds to the late season, typically at the end of August, when corn plants enter the senescence phase and harvesting commences (Pan et al., 2015). Consequently, during both the training and test phases, the ANN demonstrated challenges in accurately capturing the extreme NDVI variability induced by these two phenomena. In contrast, throughout the phases from crop growth to full development, the RMSE consistently remains below the NDVI uncertainty identified by Borgogno-Mondino et al. (Borgogno-Mondino et al., 2016). This highlights how the proposed model achieves good predictions during the critical stages of the phenological cycle.

To conduct a more in-depth examination of whether the ANN demonstrates any consistent bias during the test phase throughout the phenological cycle, ME of the three forecasting horizons is presented in Figure 10.

Table 1
NDVI forecasts average RMSE and respective STD for each output neuron.

Forecast Timesteps(Output Neurons)	NDVI RMSE	
	Mean	STD
5 days	0.028	2.33E-04
10 days	0.038	3.33E-04
15 days	0.050	4.27E-04

Figure 10 reveals a notable period from DOY 190 to 230, corresponding to the timeframe from mid-July to mid-August, during which the mean error consistently exhibits a positive trend, albeit remaining below 0.005, for all the forecasting horizons. This observation might suggest a slight overestimation by the ANN during the corn flowering phase.

As previously stated, NDVI values are affected by variable levels of uncertainty (from 0.01 up to 0.07) depending on the observation period and the NDVI value itself, therefore it is essential to set a cautious uncertainty threshold of 0.04 in order to separate the significant errors to the non significant ones. An enhanced comprehension of instances where the ANN registers significant errors surpassing the NDVI threshold can be obtained by observing the progression of the percentage of it. This progression, which spans the entire corn phenological season, is depicted in Figure 11.

Figure 11 provides valuable insights into the forecasting outcomes. Notably, in the most challenging scenario, involving the furthest forecasting horizon of 15 days, the ANN is unable to make accurate predictions for approximately 40-46% of the total fields. Notably, the most pronounced errors are observed in the initial forecast of the 15-day horizon. This outcome is reasonably expected, considering the limited information available to the ANN during the early stages of the phenological cycle, where only 3 timesteps (equivalent to 15 days) contain useful data, as reported in section 3.1.3. The addition of a single valuable observation to the input data was effective in reducing the errors to 38% of the total observations.

On average, the ANN falls short in predicting outcomes in 9.7% of cases for the 5-day horizon, 18.7% of cases for the 10-day horizon, and 27.9% of cases for the 15-day horizon. During the period spanning from DOY 140 to 220, a timeframe characterized by significant agricultural activities, the ANN exhibits inaccuracies in forecasts for 8.2%, 15.5%, and 24% of cases for the 5-day, 10-day, and 15-day forecasting horizons, respectively. These results highlight the efficacy of ANN making it a reliable tool with low error thresholds, specifically for the 5 and 10 days forecasting horizon.

4.1.1. *Timesteps Ablation Test*

The model's training approach, which involved the use of ten fixed timesteps with variable zero-padding sequences, was undertaken to impart the capacity to make accurate NDVI forecasts across periods with varying data availability. This approach sought to accommodate both data-rich phases, such as the mid-late growing season, and periods with limited data, as seen at the start of the growing season. However, in an effort to optimize this process by identifying the minimum number of timesteps required for satisfactory results, a further investigation was conducted. Figure 12 illustrates the impact of incrementally removing timesteps on the RMSE for the three forecasting horizons. The purpose of this analysis was to determine the critical threshold beyond which the forecasting performance of the model becomes significantly compromised.

Carefully inspecting figure 12 it is possible to identify the moment the ablation test causes a dramatic increase of RMSE. Specifically, for all the forecasting horizons removing 8 timesteps of information leads to significantly higher errors. That was quite expected since the entire training process involved no less than 3 useful timesteps. Removing 7 timesteps leads to an appreciable, although smaller, error increase; on the other hand, removing 1 to 6 timesteps causes no appreciable effects on the ANN results. This means that the minimum amount of useful timestep data to get accurate forecasts is 3 (i.e. 15 days) and the optimal is 4 (i.e. 20 days).

4.1.2. *Variables Ablation Test*

A variable importance assessment was conducted to identify the variables with the most impact on the ANN model. This was achieved by zero-padding each input variable individually and measuring the

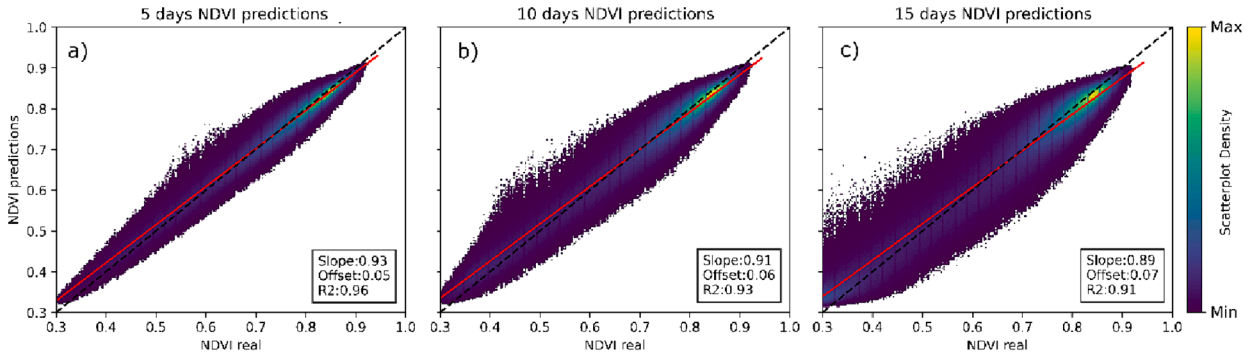


Figure 8. Target NDVI (X axis) vs Predicted (Y axis) NDVI at 5 days prediction horizon (a), at 10 days prediction horizon (b), at 15 days prediction horizon (c). The color scale of each pixel in the scatter plot corresponds to the density of data points within that pixel. Red line represents the fitted linear model.

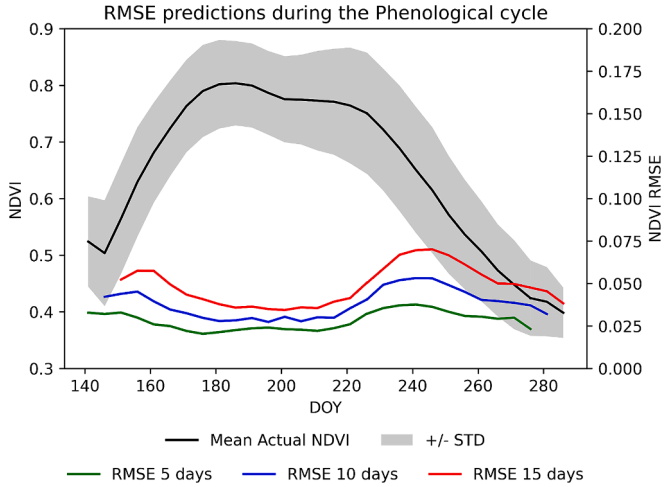


Figure 9. Average NDVIs profile (black line) with its standard deviation. NDVI forecasts RMSE at 5, 10 and 15 days horizons during the entire phenological cycle are represented by the green, blue and red lines, respectively.

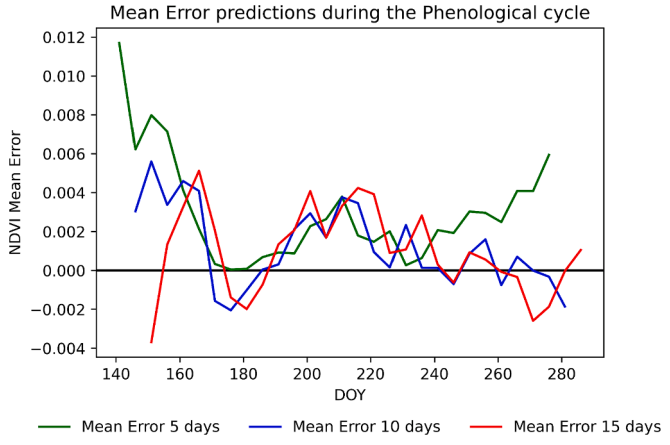


Figure 10. NDVI forecasts ME at 5, 10 and 15 days horizons during the entire phenological cycle are represented by the green, blue and red lines, respectively.

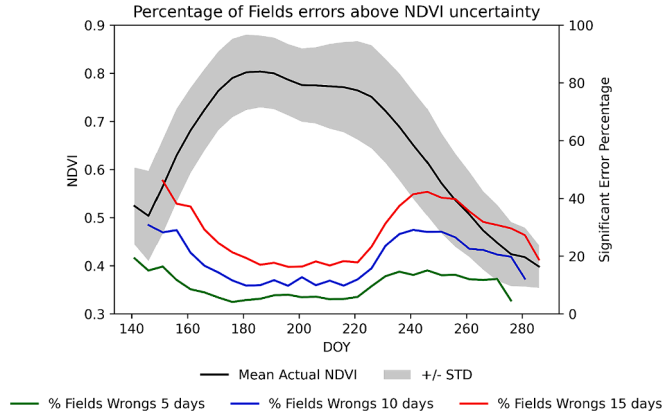


Figure 11. Average NDVIs profile (black line) with its standard deviation. Percentage of significant forecasting errors above NDVI threshold (0.04) at 5, 10 and 15 days horizon during the entire phenological cycle are represented by the green, blue and red lines, respectively.

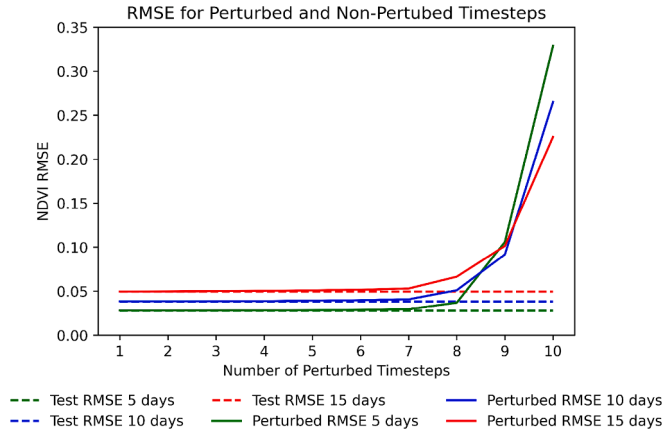


Figure 12. Timesteps Ablation Test. On the X-axis is reported the number of ablated input timesteps. Dotted lines correspond to the ANN average RMSE as reported in Table 1 (Section 4.1) and solid lines correspond to the RMSE derived from the incremental ablation of the inputs timesteps. Green, blue and red lines indicate the RMSE at the 5, 10 and 15 days forecasting horizons, respectively.

subsequent effect on RMSE. This method provided data on the relative importance of these variables in the forecasting process. Figure 13 illustrates the results, highlighting the variables that have a substantial effect on the predictive performance of the ANN.

Interestingly, the GPM ΔRMSE is the lowest among all the other variables, meaning that the variable itself is not that important for the model. However, it is known of the precipitation importance on corn development (Yamoah et al., 2000). The answer to this apparent contradiction can be found in the GPM characteristics, in the corn water management and in the presence of the NDWI as input variable.

Specifically, the Global Precipitation Measurement data has

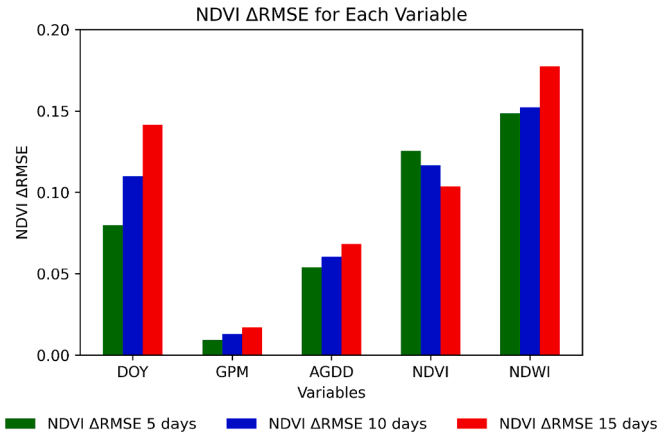


Figure 13. Variables Ablation Test. On the X-axis are reported the input ablated variables. Green, blue and red bars indicate the ΔRMSE at the 5, 10 and 15 days forecasting horizons, respectively.

exhibited a propensity to either slightly under- or over-estimate ground reference data, and the degree of this variation appears to be contingent upon the specific study area (Ramsauer et al., 2018). While the GPM data has emerged as a valuable tool for quantifying daily precipitation on a global scale, concerns have been raised regarding its reliability for applications in smaller geographic areas due to its relatively coarse geometric resolution (10 km).

Additionally, NDWI proves to be a valuable predictor for estimating vegetation water content (VWC), which is intricately linked to soil water content (Cosh et al., 2019). Consequently, providing VWC information to the ANN, as opposed to relying solely on precipitation data, appears to enhance the accuracy of NDVI forecasts. Figure 13 underscores the significance of NDWI as the most influential variable in the forecasting process thanks to the ablation test. It's noteworthy that NDWI's ΔRMSE is 0.15 for the 5 and 10-day horizons and increasing to 0.17 for the 15-day horizon. This trend is consistent across various variables, where the importance of a variable intensifies with the forecasting horizon's length. An exception is NDVI, which gradually diminishes in importance with extended forecasting horizons. This observation suggests that past NDVI values are vital for precise predictions in the nearest horizons, while for longer-term predictions, past NDWI assumes the pivotal role. This implies that historical VWC significantly influences future crop biomass, with its impact growing as the forecasting horizon extends.

The relationship between AGDD and NDVI is a well-established connection in the existing literature (Ghamghami et al., 2019). Consequently, it was anticipated that this variable would exert a significant impact on the ANN's predictions, although not to the same degree as NDVI values. This discrepancy can be attributed to the method employed in deriving air temperature data from the nearest meteorological station. Since air temperature measurements might originate from several kilometres away from the specific field, the influence of AGDD on predictions is somewhat mitigated in comparison to the more localized and direct impact of NDVI values.

Unexpectedly, DOY plays an important role in the ANN, specifically at the 10 and 15 days horizons. Knowing in which period of the year the past NDVI, NDWI, AGDD and GPM observations occur allows the ANN to correctly identify the growing period. This is a critical insight since it's well-established that the timing of corn seeding significantly impacts the growth (Djaman et al., 2022).

4.1.3. ANN generalization capability

The ANN's ability to generalize over a different AOI was tested using a dataset comprising 1002 corn fields from the LUCAS dataset, collected across Europe in 2018. However, before delving into the ANN's performance, it's crucial to evaluate the relevance of each input variable. It's worth noting that variations in bioclimatic areas can result in significant differences in monthly average temperatures and annual total precipitation levels (Cortesi et al., 2012; Metzger et al., 2005). Of particular concern were the GPM values, which frequently exceeded those observed in the Piemonte region by a factor of 2 to 3. This raised questions about the ANN's capacity to produce accurate forecasts when presented with data points outside the range encountered during the training phase.

For this reason, the first conducted analysis was the variables ablation test on the LUCAS dataset (Figure 14).

In comparison to Figure 13, it is noteworthy that all variables appear to have diminished in importance except the NDVI. However, it is crucial to consider that the overall RMSE values have also shifted to higher values, specifically to 0.098, 0.124, and 0.154 for the 5, 10, and 15 days forecasting horizons (Table 2). ΔRMSE serves as a relative measure of the variables' importance within the model.

Examining individual variables, GPM ΔRMSE consistently displays negative values, indicating a general decline in the ANN's performance when the GPM variable is included. Interestingly, NDVI has emerged as the most crucial variable, and the diminishing importance observed in Figure 13 has been further accentuated. NDWI, while still valuable, has lost some of its utility to the ANN. DOY and AGDD continue to provide information to the ANN, although their significance has become questionable. This was expected due to the specific training area characterized by distinct agricultural practices and temperatures which may vary compared to the entire European agricultural context. Additionally, the training dataset represents only a subset of the diverse corn cultivars in Europe (more than 5000 according to European Commission, 2024), each with varying phenological cycles and thermal efficiencies (Hou et al., 2014). Consequently, DOY and AGDD variables in the LUCAS dataset might not align with the same temporal patterns and relationships with other variables as observed in the Piemonte dataset.

Following the variable ablation test, it was evident that the GPM variable had an adverse impact on the ANN's performance when tested on the LUCAS dataset. Consequently, to optimize results, GPM was excluded from the ANN inputs. With this adjustment, the ANN was evaluated using the LUCAS dataset, and the overall RMSE was calculated for each forecasting horizon (Table 2). However, this approach causes a misalignment between the results obtained in the Piemonte region (with GPM variable included) and the results in the LUCAS dataset (without GPM). Therefore, to overcome this inconsistency, the GPM variable was also removed from the Piemonte test inputs and the ANN was tested again. From this point onward, when comparing LUCAS and Piemonte region results', we refer to the ones where the GPM variable is not included in the inputs.

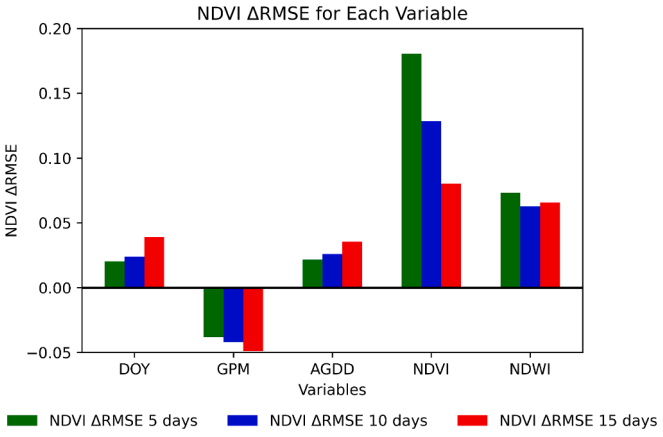


Figure 14. LUCAS Variables Ablation Test. On the X-axis are reported the input ablated variables. Green, blue and red bars indicate the ΔRMSE at the 5, 10 and 15 days forecasting horizons, respectively.

Table 2
NDVI forecasts RMSE for each output neuron on Piemonte (AOI) and LUCAS dataset with and without the GPM as input variable.

Forecast Timesteps(Output Neurons)	NDVI RMSE (AOI)	NDVI RMSE – NO GPM (AOI)	NDVI RMSE (LUCAS)	NDVI RMSE – NO GPM (LUCAS)
5 days	0.028	0.037	0.098	0.062
10 days	0.038	0.051	0.124	0.083
15 days	0.050	0.067	0.154	0.105

When comparing ANN RMSE – NO GPM for AOI and LUCAS, a noticeable increase in the RMSE for all three output neurons can be detected. Specifically, the RMSE for each of the forecasting horizons has more than doubled, rising from 0.037 to 0.062, from 0.051 to 0.083, and from 0.067 to 0.105 for the 5, 10, and 15-day horizons respectively. It is important to note that all RMSE values have exceeded the average NDVI threshold, indicating a higher number of significant errors. However, this increase in error should be viewed in context and refers to the entire Europe with all its climatic variability as depicted in Figure 7.

In order to check the timestep importance for the ANN on the LUCAS dataset, the ablation test conducted on the timesteps is reported in figure 15.

Carefully inspecting figure 15 it is possible to identify the moment the ablation test causes a dramatic increase of RMSE. Specifically, for all the forecasting horizons removing nine information timesteps leads to significantly higher errors. Compared to figure 12, it is evident that the ANN only uses the last three timesteps instead of the last four. Probably, the input and forget gates (i_t and f_t, respectively) do not expect so much data diversity and do not allow the usage of all the information but only the most recent one when a so variable dataset is tested.

To better understand where the biggest errors lie and highlight potential biases between the predicted and target values, a first order linear model was used for each prediction horizon and is illustrated in Figure 16.

The first notable distinction, when compared to Figure 8(a, b, c), is the reduced density of point clouds, primarily attributed to the differences in dataset dimensions between the Piemonte and LUCAS datasets. All three figures 16 - a, b, c exhibit a trend similar to Figure 8 (a, b, c), wherein R^2 decreases as the forecasting horizon increases. Furthermore, in this case, the offset and slope consistently maintain values greater than 0 and less than 1, respectively, although the extent of overestimation at low NDVI values is more pronounced.

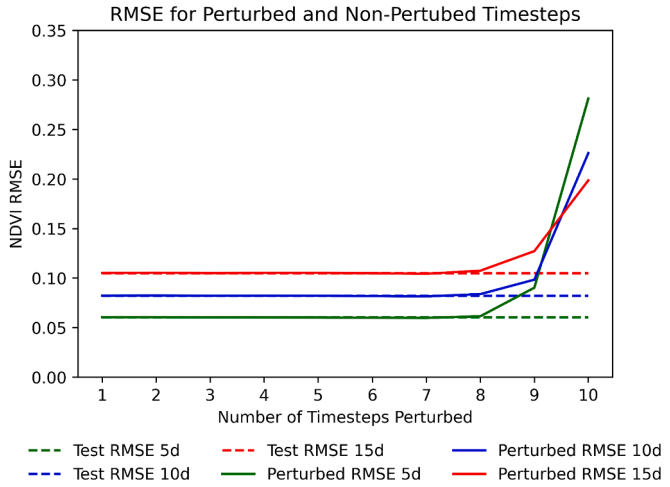


Figure 15. Timesteps Ablation Test. On the X-axis is reported the number of ablated input timesteps. Dotted lines correspond to the ANN overall LUCAS RMSE – NO GPM as reported in Table 2 and solid lines correspond to the RMSE derived from the incremental ablation of the inputs timesteps. Green, blue and red lines indicate the RMSE – NO GPM at the 5, 10 and 15 days forecasting horizons, respectively.

Just like the scatterplots shown in Figures 8, the concentration of points indicates that only a small portion of the forecasts significantly diverge from the bisector of the scatterplot. Meanwhile, most of the forecasts are clustered in the area representing high NDVI values. To have a better understanding of the ANN performances during the phenological cycle, the RMSE for the three forecasting horizons and the average NDVI profile are reported in Figure 17.

There are several crucial observations to make. First, Figure 17.a shows similar RMSE temporal profiles in comparison to the ones reported in Figure 11 although slightly higher as expected by the analysis conducted in section 4.1.2. Specifically, it seems that GPM variable mostly affects the period from DOY 230 to 260. Second, it is essential to note the remarkable extension of the standard deviation in the NDVI average profile when comparing Figure 17.a and 17.b. It is evident that NDVI values throughout the year exhibit considerably more variability in LUCAS dataset, with standard deviations consistently exceeding 0.2. In contrast, the Piemonte region, despite data originating from different years, typically maintains an average standard deviation value below 0.1. This heightened variability is a clear explanation for the increase in ANN RMSE on the LUCAS dataset. Moreover, the patterns in the ANN RMSEs for the three forecasting horizons are similar, where the lowest errors are predominantly concentrated in the middle parts of the phenological cycle.

Finally, LUCAS dataset was divided into the 7 european Köppen climatic zones and the average RMSE was reported in figure 18 in order to better visualize its spatial distribution across different climatic zones.

Figure 18 highlights the spatial distribution of the errors the ANN makes when inference is performed on a completely different and more variable dataset in respect to the one which was trained on. It is evident that depending on the climatic zone, the errors dramatically change. Specifically, BSk, Csa and Cfa (arid, steppe, cold; temperate, dry summer, hot summer; temperate, no dry season, hot summer- respectively) show the smaller errors when compared to the “continental” climates like Cfb and Dfb. However, it is worth noting that Cfb and Dfb have a much larger number of corn fields, which could potentially affect the reliability of this comparison.

5. Discussions

NDVI is one of the most popular vegetation indices used in remote sensing applications and precision farming contexts. However, cloud coverage and processing temporal lags pose significant challenges for real-time applications, particularly in the agricultural sector. Furthermore, NDVI forecasting offers a pivotal advantage by enabling the anticipation of vital agronomic activities and proactive decision-making. With accurate predictions, farmers and land managers can preemptively plan and implement essential agronomic practices, enhancing the efficiency of resource allocation and contributing to improved agricultural productivity and sustainability. In addition to its agricultural applications, NDVI forecasting holds immense potential for micro and macro-scale regional water resources management, particularly in the context of severe droughts. By offering predictive insights into vegetation health and land cover changes, NDVI forecasting empowers regional authorities and policymakers to make informed decisions about water allocation, conservation measures, and the overall management of critical water resources. During periods of severe drought, where water availability becomes a pressing concern, the ability to anticipate vegetation stress and water demand through NDVI forecasts plays a crucial role in optimizing water distribution, mitigating the impacts of drought, and safeguarding the ecological balance of ecosystems. This macro-scale utilization of NDVI forecasting can substantially improve regional resilience in the face of water scarcity, ultimately contributing to the sustainable management of water resources.

Previous research in NDVI estimation for the current growing season has largely relied on conventional time-series forecasting like

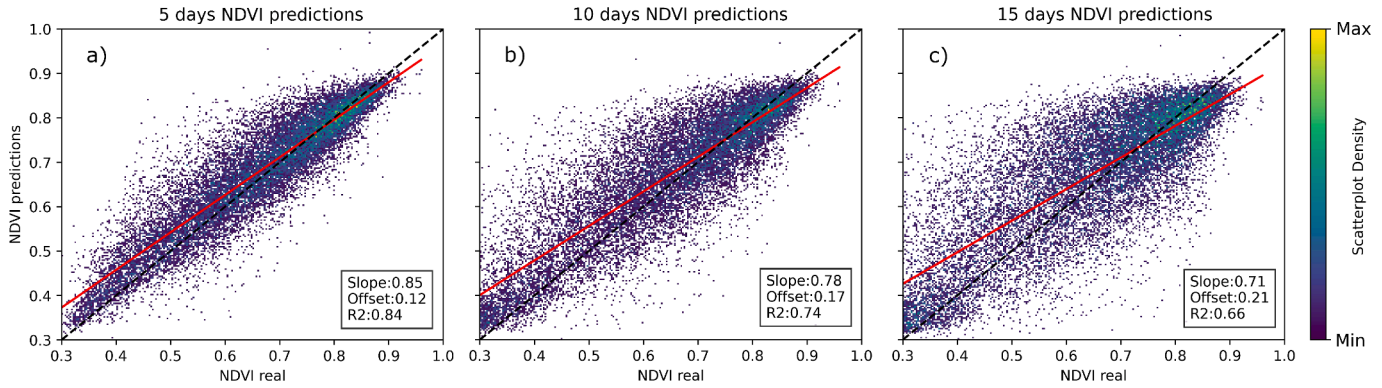


Figure 16. Target NDVI (X axis) vs Predicted (Y axis) NDVI at 5 days prediction horizon (a), at 10 days prediction horizon (b), at 15 days prediction horizon (c) for the LUCAS dataset. The color scale of each pixel in the scatter plot corresponds to the density of data points within that pixel. Red line represents the fitted linear model.

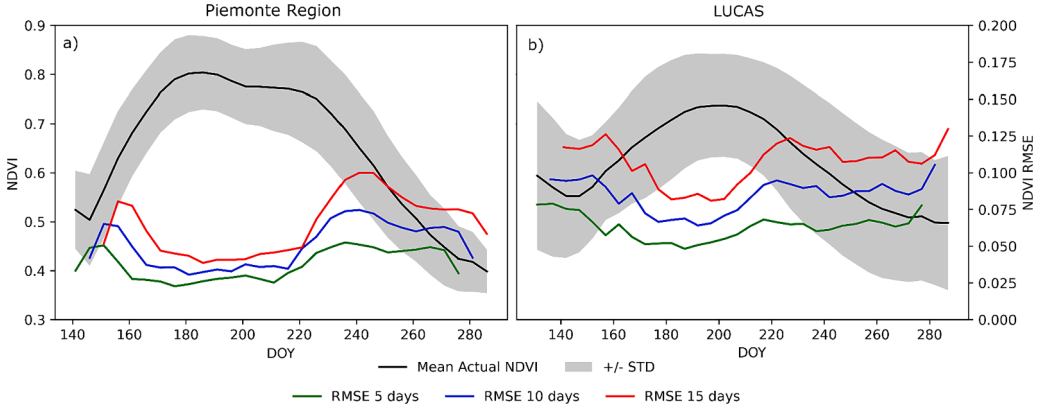


Figure 17. Plots a and b report the average NDVIs profiles (indicated by the black lines) along with their associated standard deviations. The RMSEs – NO GPM of NDVI forecasts at 5, 10, and 15-day horizons throughout the complete phenological cycle are presented by the green, blue, and red lines, respectively. Plot *a* pertains to the Piemonte test dataset, while plot *b* represents the LUCAS dataset.

autoregressive and regression techniques (Atkinson et al., 2012, Cao et al., 2015, Vorobiova and Chernov, 2017). However, these approaches have consistently been surpassed by deep learning methods, particularly those integrating time series analysis. Gomez-Lagos et al. (Gómez-Lagos et al., 2023) employed a standard multilayer perceptron but was limited by its testing solely on NDVI pixels for a single date. Analyses on a single date are not viable for a model that must be tested over an entire growing season, therefore ANN which specifically deal with time series are needed.

When it comes to forecasting tasks, a time series approach is needed and as such, advanced methods have begun to incorporate RNNs or variants like LSTM. For instance, Stepchenko and Chizhov (Stepchenko and Chizhov, 2015) accomplished MODIS pixel-level forecasts with an RMSE of 0.035 using LSTM architectures. However, questions arose regarding their training procedure and the absence of a validation set. Similarly, Reddy and Prasad (Reddy and Prasad, 2018) employed LSTM for large-area NDVI forecasts, achieving a remarkably low RMSE of less than 0.03. However, this study focused on forecasting averages over extensive regions using 250m MODIS pixels, a task that is less intricate than field-level predictions which are influenced by agronomic practices.

More recently, Ahmad et al. (Ahmad et al., 2023) introduced the ConvLSTM methodology for field-level predictions, achieving an RMSE of 0.0782. However, it's important to note that Ahmad et al. used relatively large 250m pixels suitable for the vast fields common in Southern American agriculture, which differs from our goal of predicting NDVI in smaller and more fragmented fields typically found in Italian agriculture.

Finally, Cavalli et al. (Cavalli et al., 2023) proposed several LSTM networks for forecasting NDVI values at 1, 2, and 3 days with RMSE that surpassed every other approach in the field-level forecasting task. Their best results achieved RMSE values of 0.032, 0.062, and 0.085 for the respective time steps. While comparing these outcomes to our results, some differences must be considered: Cavalli et al. trained several LSTM ANN to forecast at 1, 2, and 3 days, while our methodology aimed to create a single model for forecasting at 5, 10, and 15 days. Our ANN achieved an RMSE of 0.028 at 5 days, outperforming the 1-day forecast obtained by Cavalli et al. At 15 days in the future, our ANN's RMSE of 0.051 surpassed the 2-day forecast of 0.06201 by Cavalli et al.

It is worth noting that our ANN was trained on a considerably larger area compared to other studies (>160000 ha in our study compared to 120000 ha in Cavalli et al. and 1945 ha in Ahmad et al.). Moreover, AOI explores an extensive area (13000 km^2) which guarantees a high pedological, topographical and climatic variability. The increased variability in the training dataset posed an extremely difficult task to the ANN which ultimately aided in the generalization phase. We tested the ANN's ability to generalize across the entire European continent, a task not previously undertaken to our knowledge. The performance of the ANN decreased as expected, but it still outperformed the models proposed by Cavalli et al.

Our methodology introduced several novel elements: the use of Bi-LSTM, exploration of significant variability, introduction of NDWI and DOY as input variables, and implementation of BATS - a minibatch procedure that allowed training the ANN to perform effectively in

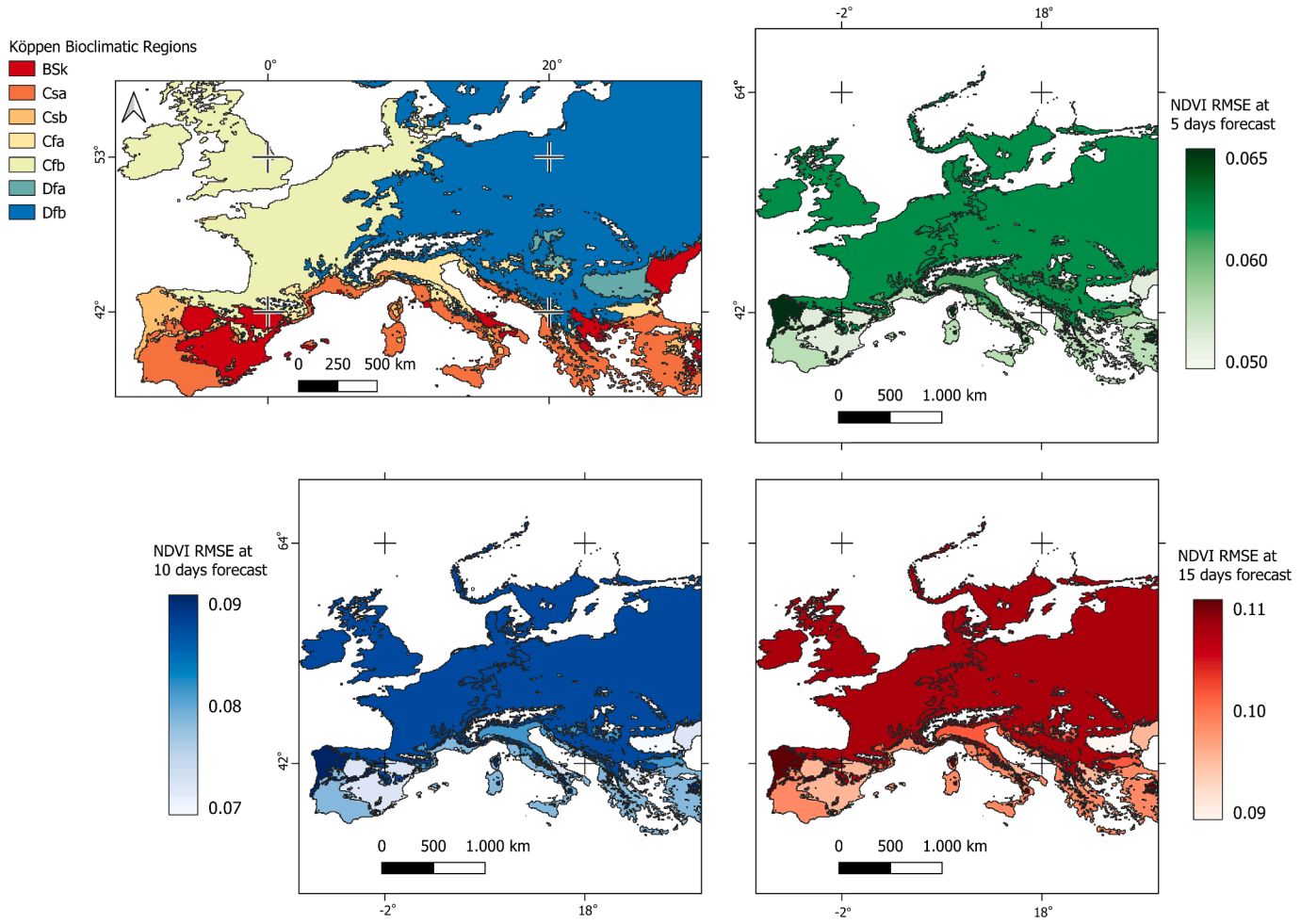


Figure 18. a) Spatial distribution of Köppen climatic zones, b) 5-day horizon RMSE (green) for Köppen climatic zones, c) 10-day horizon RMSE (blue) for Köppen climatic zones, d) 15-day horizon RMSE (red) for Köppen climatic zones.

periods with both small and large data availability. All the previous elements set the new state of the art when facing the NDVI forecasting task.

6. Conclusion

In this paper, a double-layer Bi-LSTM was employed to forecast NDVI at three different timesteps (5, 10 and 15 days) for corn crop. Specifically, the ANN was trained in Piemonte region (NW Italy) over a vast dataset that comprised a total of 130401 corn fields collected over five years (from 2018 to 2022). The ANN was intended to predict NDVI time series based on past meteorological (temperature and precipitation), spectral indices (NDVI and NDWI from Sentinel-2 mission) and the day of the year in which each observation occurs in periods with both large and limited data availability thanks to a peculiar training minibatch generation we call BATS. Additionally, conducted ablation tests confirmed that the key factor in improving the model's precision is the NDWI, whereas GPM contributed minimally to the ANN's performance, likely due to its coarse spatial resolution. Also, the last four timesteps were found to be the most informative, indicating that long sequences are not necessarily needed for accurate forecasts.

The proposed methodology, thanks to the described novelties, managed to outperform all the previous models. Furthermore, it is essential to emphasize that the ANN model was intentionally designed for real-time application, allowing its use at any point during the phenological season, whether data availability is limited or abundant. This adaptability enhances its practical utility in agricultural decision-making and monitoring, making it a valuable tool for farmers, authorities and researchers alike. Nevertheless, there are limitations to this approach. Training the ANN, particularly with very large datasets, requires considerable computational resources and time. Moreover, while the ANN inputs are sequential, the outputs are three discrete future values. Future enhancements should evolve the model into a sequence-to-sequence ANN, where outputs also form a sequential series while forecasting multiple spectral indices at the same time such as NDRE and EVI. Moreover, comparative studies with multiple machine learning and deep learning algorithms, including Random Forest, Support Vector Machine, sequence-to-sequence LSTM, and Transformer models, should be undertaken to further validate and improve the forecasting capability. Although the current workflow is optimized for a single crop type, expanding its applicability to multiple crops remains a key objective for future development.

Finally, we propose to define a common benchmark dataset on which crop forecasting models should be tested in order to better compare them. A large open dataset is not easy to find, however LUCAS dataset can be the reference for all the future works until larger datasets become available.

Ultimately, NDVI forecasts are essential in both the precision farming context and in the regional monitoring services. Knowing beforehand the crop behaviour is a powerful tool to be used to better plan agronomic practices and to manage water resources (at field, consortium and regional level).

CRediT authorship contribution statement

A. Farbo: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **F. Sarvia:** Data curation, Validation, Writing – review & editing. **S. De Petris:** Validation, Writing – review & editing. **V. Basile:** Methodology, Supervision, Writing – review & editing. **E. Borgogno-Mondino:** Project administration, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research was supported by the project PSR 2014-2020 Regione Piemonte - Misura 16 - Operazione 16.1.1 “Costituzione, gestione e operatività dei gruppi operativi dei PEI (AGRICOLTURA)”. Project “TELECER - Applicazione del telerilevamento per il miglioramento produttivo e qualitativo dei cereali per le filiere avanzate”.

References

Ahmad, R., Yang, B., Ettlin, G., Berger, A., Rodríguez-Bocca, P., 2023. A machine-learning based ConvLSTM architecture for NDVI forecasting. Int. Trans. Oper. Res. 30, 2025–2048. https://doi.org/10.1111/itor.12887.

Atkinson, P.M., Jeganathan, C., Dash, J., Atzberger, C., 2012. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. Remote Sens. Environ. 123, 400–417. https://doi.org/10.1016/j.rse.2012.04.001.

Bargagli Stoffi, F.J., Cevolani, G., Gnecco, G., 2022. Simple models in complex worlds: Occam’s razor and statistical learning theory. Minds Mach. 32, 13–42. https://doi.org/10.1007/s11023-022-09592-z.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci. Data 5, 180214. https://doi.org/10.1038/sdata.2018.214.

Berti, A., Maucieri, C., Bonamano, A., Borin, M., 2019. Short-term climate change effects on maize phenological phases in northeast Italy. Ital. J. Agron. 14, 222–229. https://doi.org/10.4081/ija.2019.1362.

Bin, Y., Yang, Y., Shen, F., Xie, N., Shen, H.T., Li, X., 2019. Describing video with attention-based bidirectional LSTM. IEEE Trans. Cybern. 49, 2631–2641. https://doi.org/10.1109/TCYB.2018.2831447.

Blaes, X., Chomé, G., Lambert, M.-J., Traoré, P.S., Schut, A.G.T., Defourny, P., 2016. Quantifying fertilizer application response variability with VHR satellite NDVI time series in a rainfed smallholder cropping system of Mali. Remote Sens. 8, 531. https://doi.org/10.3390/rs8060531.

Borgogno-Mondino, E., Lessio, A., Gomarasca, M.A., 2016. A fast operative method for NDVI uncertainty estimation and its role in vegetation analysis. Eur. J. Remote Sens. 49, 137–156. https://doi.org/10.5721/EuJRS20164908.

Borgogno-Mondino, E., Farbo, A., Novello, V., de Palma, L., 2022. A fast regression-based approach to map water status of pomegranate orchards with sentinel 2 data. Horticulturae 8, 759. https://doi.org/10.3390/horticulturae8090759.

Cao, R., Chen, J., Shen, M., Tang, Y., 2015. An improved logistic method for detecting spring vegetation phenology in grasslands from MODIS EVI time-series data. Agric. For. Meteorol. 200, 9–20. https://doi.org/10.1016/j.agrformet.2014.09.009.

Cavalli, S., Penzotti, G., Amoretti, M., Caselli, S., 2023. A Machine Learning Approach for NDVI Forecasting based on Sentinel-2 Data. In: Presented at the 16th International Conference on Software Technologies, pp. 473–480.

Cortesi, N., Gonzalez-Hidalgo, J.C., Brunetti, M., Martin-Vide, J., 2012. Daily precipitation concentration across Europe 1971–2010. Nat. Hazards Earth Syst. Sci. 12, 2799–2810. https://doi.org/10.5194/nhess-12-2799-2012.

Cosh, M.H., White, W.A., Colliander, A., Jackson, T.J., Prueger, J.H., Hornbuckle, B.K., Hunt, E.R., McNairn, H., Powers, J., Walker, V.A., Bullock, P., 2019. **Estimating vegetation water content during the Soil Moisture Active Passive Validation Experiment 2016**. J. Appl. Remote Sens. 13, 014516 https://doi.org/10.1117/1.JRS.13.014516.

de Castro, C., Filho, H., de Carvalho, A., Júnior, O., Ferreira de Carvalho, O.L., Pozzobon de Bem, P., dos Santos de Moura, R., Olino de Albuquerque, A., Rosa Silva, C., Guimarães Ferreira, P.H., Fontes Guimarães, R., Trancoso Gomes, R.A., 2020. Rice crop detection using LSTM, Bi-LSTM, and machine learning models from sentinel-1 time series. Remote Sens. 12, 2655. https://doi.org/10.3390/rs12162655.

De Petris, S., Sarvia, F., Borgogno-Mondino, E., 2023. Uncertainty assessment of Sentinel-2-retrieved vegetation spectral indices over Europe. Eur. J. Remote Sens. 2267169. https://doi.org/10.1080/22797254.2023.2267169.

Djaman, K., Allen, S., Djaman, D.S., Koudahe, K., Irmak, S., Puppala, N., Darapuneni, M. K., Angadi, S.V., 2022. Planting date and plant density effects on maize growth, yield and water use efficiency. Environ. Chall. 6, 100417 https://doi.org/10.1016/j.envc.2021.100417.

Djordjević, K.L., Jordović-Pavlović, M.I., Ćojbašić, Ž.M., Galović, S.P., Popović, M.N., Nešić, M.V., Markushev, D.D., 2022. Influence of data scaling and normalization on overall neural network performances in photoacoustics. Opt. Quantum Electron. 54, 501. https://doi.org/10.1007/s11082-022-03799-1.

Dwarampudi, M., Reddy, N.V.S., 2019. Effects of padding on LSTMs and CNNs. https://doi.org/10.48550/arXiv.1903.07288.

Efeoğlu, B., Ekmekçi, Y., Çiçek, N., 2009. Physiological responses of three maize cultivars to drought stress and recovery. South Afr. J. Bot. 75, 34–42. https://doi.org/10.1016/j.sajb.2008.06.005.

Coltivazioni: Cereali, legumi, radici bulbi e tuberi [WWW Document], n.d. URL http://dati.istat.it/Index.aspx?QueryId=33702# (accessed 2.7.24).

Plant variety catalogues, databases & information systems - European Commission [WWW Document], URL https://food.ec.europa.eu/plants/plant-reproductive-material/plant-variety-catalogues-databases-information-systems_en (accessed 2.9.24).

Farbo, A., Sarvia, F., De Petris, S., Borgogno-Mondino, E., 2022. Preliminary Concerns About Agronomic Interpretation Of Ndvi Time Series From Sentinel-2 Data: Phenology And Thermal Efficiency Of Winter Wheat In Piemonte (NW ITALY). Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLIII-B3-2022, 863–870. Doi: 10.5194/isprs-archives-XLIII-B3-2022-863-2022.

Fathollahi, L., Wu, F., Melaki, R., Jamshidi, P., 2023. Global Ndvi Forecasting from Air Temperature, Soil Moisture and Precipitation Using a Deep Neural Network. Doi: 10.2139/ssrn.4598952.

Fei, Y., Jiulin, S., Hongliang, F., Zuofang, Y., Jiahua, Z., Yunqiang, Z., Kaishan, S., Zongming, W., Maogui, H., 2012. Comparison of different methods for corn LAI estimation over northeastern China. Int. J. Appl. Earth Obs. Geoinformation 18, 462–471. https://doi.org/10.1016/j.jag.2011.09.004.

Gao, B.-C., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266.

Gao, L., Wang, X., Johnson, B.A., Tian, Q., Wang, Y., Verrelst, J., Mu, X., Gu, X., 2020. Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: A review. ISPRS J. Photogramm. Remote Sens. 159, 364–377. https://doi.org/10.1016/j.isprsjprs.2019.11.018.

Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., Languille, F., Alhammoud, B., Viallefont, F., Pflug, B., Bieniarz, J., Clerc, S., Pessiot, L., Trémas, T., Cadau, E., De Bonis, R., Isola, C., Martimort, P., Fernandez, V., 2017. Copernicus sentinel-2A calibration and products validation status. Remote Sens. 9, 584. https://doi.org/10.3390/rs9060584.

Ghamghami, M., Ghahreman, N., Irannejad, P., Ghorbani, K., 2019. Comparison of data mining and GDD-based models in discrimination of maize phenology. Int. J. Plant Prod. 13, 11–22. https://doi.org/10.1007/s42106-018-0030-2.

Ghilardi, F., Virano, A., Prandi, M., Borgogno-Mondino, E., 2023. Zonation of a viticultural territorial context in piemonte (NW Italy) to support terroir identification: the role of pedological. Topographical and Climatic Factors. Land 12, 647. https://doi.org/10.3390/land12030647.

Gill, K., Babuta, R., Kaur, N., Sandhu, S., 2014. Thermal requirement of wheat crop in different agroclimatic regions of Punjab under climate change scenarios. Mausam 65, 417–424.

Gold, C.M., Remmele, P.R., Roos, T., 1997. Voronoi methods in GIS, in: van Kreveld, M., Nievergelt, J., Roos, T., Widmayer, P. (Eds.), Algorithmic Foundations of Geographic Information Systems, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 21–35. Doi: 10.1007/3-540-63818-0_2.

Giolo, M., Sallenave, R., Pornaro, C., Velasco-Cruz, C., Macolino, S., Leinauer, B., 2021. Base temperatures affect accuracy of growing degree day model to predict emergence of bermudagrasses. Agron. J. 113, 2960–2966. https://doi.org/10.1002/agj2.20660.

Gómez-Lagos, J.E., González-Araya, M.C., Blu, R.O., Espejo, L.G.A., 2023. Using Data Mining Techniques to Forecast the Normalized Difference Vegetation Index (NDVI). In: in Table Grape. Presented at the 8th International Conference on Operations Research and Enterprise Systems, pp. 189–194.

Graves, A., Fernández, S., Schmidhuber, J., 2005. Bidirectional LSTM Networks for Improved Phoneme Classification and Recognition, in: Duch, W., Kacprzyk, J., Oja, E., Zadrożny, S. (Eds.), Artificial Neural Networks: Formal Models and Their Applications – ICANN 2005, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 799–804. Doi: 10.1007/11550907_126.

Grover, A., Singh, R.B., 2015. Analysis of urban heat island (UHI) in relation to normalized difference vegetation index (NDVI): a comparative study of Delhi and Mumbai. Environments 2, 125–138. https://doi.org/10.3390/environments2020125.

Hatfield, J.L., Dold, C., Hatfield, J.L., Dold, C., 2018. Climate Change Impacts on Corn Phenology and Productivity, in: Corn - Production and Human Health in Changing Climate. IntechOpen. Doi: 10.5772/intechopen.76933.

Hird, J.N., McDermid, G.J., 2009. Noise reduction of NDVI time series: an empirical comparison of selected techniques. Remote Sens. Environ. 113, 248–258. https://doi.org/10.1016/j.rse.2008.09.003.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9, 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Hosseini, M., McNairn, H., Merzouki, A., Pacheco, A., 2015. Estimation of leaf area index (LAI) in corn and soybeans using multi-polarization C- and L-band radar data. Remote Sens. Environ. 170, 77–89. https://doi.org/10.1016/j.rse.2015.09.002.

Hou, P., Liu, Y., Xie, R., Ming, B., Ma, D., Li, S., Mei, X., 2014. Temporal and spatial variation in accumulated temperature requirements of maize. Field Crops Res. 158, 55–64. https://doi.org/10.1016/j.fcr.2013.12.021.
Houborg, R., McCabe, M.F., 2016. High-resolution NDVI from planet's constellation of earth observing nano-satellites: a new data source for precision agriculture. Remote Sens. 8, 768. https://doi.org/10.3390/rs8090768.
Huang, S., Ming, B., Huang, Q., Leng, G., Hou, B., 2017. A case study on a combination NDVI forecasting model based on the entropy weight method. Water Resour. Manag. 31, 3667–3681. https://doi.org/10.1007/s11269-017-1692-8.
Huete, A., Justice, C., Liu, H., 1994. Development of vegetation and soil indices for MODIS-EOS. Remote Sens. Environ. 49, 224–234. https://doi.org/10.1016/0034-4257(94)90018-3.
Johnson, M.K., Kjell, 2019. Feature engineering and selection: a practical approach for predictive models. Chapman and Hall/CRC, Boca Raton. https://doi.org/10.1201/9781315108230.
Li, S., Xu, L., Jing, Y., Yin, H., Li, X., Guan, X., 2021. High-quality vegetation index product generation: A review of NDVI time series reconstruction techniques. Int. J. Appl. Earth Obs. Geoinformation 105, 102640. https://doi.org/10.1016/j.jag.2021.102640.
Lizaso, J.I., Ruiz-Ramos, M., Rodríguez, L., Gabaldon-Leal, C., Oliveira, J.A., Lorite, I.J., Sánchez, D., García, E., Rodríguez, A., 2018. Impact of high temperatures in maize: Phenology and yield components. Field Crops Res 216, 129–140. https://doi.org/10.1016/j.fcr.2017.11.013.
Maselli, F., Chiesi, M., Angeli, L., Fibbi, L., Rapi, B., Romani, M., Sabatini, F., Battista, P., 2020. An improved NDVI-based method to predict actual evapotranspiration of irrigated grasses and crops. Agric. Water Manag. 233, 106077 https://doi.org/10.1016/j.agwat.2020.106077.
Mcmaster, G., 1997. Growing degree-days: one equation, two interpretations. Agric. For. Meteorol. 87, 291–300. https://doi.org/10.1016/S0168-1923(97)00027-0.
McNicholl, B., Lee, Y.H., Campbell, A.G., Dev, S., 2022. Evaluating the reliability of air temperature from ERA5 reanalysis data. IEEE Geosci. Remote Sens. Lett. 19, 1–5. https://doi.org/10.1109/LGRS.2021.3137643.
Meng, J., Du, X., Wu, B., 2013. Generation of high spatial and temporal resolution NDVI and its application in crop biomass estimation. Int. J. Digit. Earth 6, 203–218. https://doi.org/10.1080/17538947.2011.623189.
Metzger, M.J., Bunce, R.G.H., Jongman, R.H.G., Mücher, C.A., Watkins, J.W., 2005. A climatic stratification of the environment of Europe. Glob. Ecol. Biogeogr. 14, 549–563. https://doi.org/10.1111/j.1466-822X.2005.00190.x.
Michishita, R., Jin, Z., Chen, J., Xu, B., 2014. Empirical comparison of noise reduction techniques for NDVI time-series based on a new measure. ISPRS J. Photogramm. Remote Sens. 91, 17–28. https://doi.org/10.1016/j.isprsjprs.2014.01.003.
Miller, P., Lanier, W., Brandt, S., 2001. Using growing degree days to predict plant stages. AgExtension Commun. Coord. Commun. Serv. Mont. State Univ.-Bozeman Bozeman MO 59717, 994–2721.
Miura, T., Huete, A.R., Yoshioka, H., 2000. Evaluation of sensor calibration uncertainties on vegetation indices for MODIS. IEEE Trans. Geosci. Remote Sens. 38, 1399–1409. https://doi.org/10.1109/36.843034.
Mkhabela, M.S., Bullock, P., Raj, S., Wang, S., Yang, Y., 2011. Crop yield forecasting on the Canadian Prairies using MODIS NDVI data. Agric. For. Meteorol. 151, 385–393. https://doi.org/10.1016/j.agrformet.2010.11.012.
Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: a review. ISPRS J. Photogramm. Remote Sens. 66, 247–259. https://doi.org/10.1016/j.isprsjprs.2010.11.001.
Nouri, H., Beecham, S., Anderson, S., Nagler, P., 2014. High spatial resolution worldview-2 imagery for mapping NDVI and its relationship to temporal urban landscape evapotranspiration factors. Remote Sens. 6, 580–602. https://doi.org/10.3390/rs6010580.
Okin, G.S., Gu, J., 2015. The impact of atmospheric conditions and instrument noise on atmospheric correction and spectral mixture analysis of multispectral imagery. Remote Sens. Environ. 164, 130–141. https://doi.org/10.1016/j.rse.2015.03.032.
Orusa, T., Viani, A., Cammareri, D., Borgogno Mondino, E., 2023. A Google earth engine algorithm to map phenological metrics in mountain areas worldwide with landsat collection and sentinel-2. Geomatics 3, 221–238. https://doi.org/10.3390/geomatics3010012.
Pacheco, A., Bannari, A., Staenz, K., Mcnairn, H., 2001. LAI measurements in white beans and corn canopies with two optical instruments. Presented at the Mesures physiques et signatures en télédétection (Aussois, 8-12 January 2001), pp. 374–379.
Pan, Z., Huang, J., Zhou, Q., Wang, L., Cheng, Y., Zhang, H., Blackburn, G.A., Yan, J., Liu, J., 2015. Mapping crop phenology using NDVI time-series derived from HJ-1 A/B data. Int. J. Appl. Earth Obs. Geoinformation 34, 188–197. https://doi.org/10.1016/j.jag.2014.08.011.
Parmar, A., Katariya, R., Patel, V., 2019. A Review on Random Forest: An Ensemble Classifier, in: Hemanth, J., Fernando, X., Lafata, P., Baig, Z. (Eds.), International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018, Lecture Notes on Data Engineering and Communications Technologies. Springer International Publishing, Cham, pp. 758–763. Doi: 10.1007/978-3-030-03146-6_86.
Potithep, S., Nagai, S., Nasahara, K.N., Muraoka, H., Suzuki, R., 2013. Two separate periods of the LAI–VIs relationships using in situ measurements in a deciduous broadleaf forest. Agric. For. Meteorol. 169, 148–155. https://doi.org/10.1016/j.agrformet.2012.09.003.
Poudel, U., Stephen, H., Ahmad, S., 2021. Evaluating irrigation performance and water productivity using EEFlux ET and NDVI. Sustainability 13, 7967. https://doi.org/10.3390/su13147967.
Radočaj, D., Šiljeg, A., Marinović, R., Jurišić, M., 2023. State of major vegetation indices in precision agriculture studies indexed in web of science: a review. Agriculture 13, 707. https://doi.org/10.3390/agriculture13030707.
Ramsauer, T., Weiß, T., Marzahn, P., 2018. Comparison of the GPM IMERG final precipitation product to RADOLAN weather radar data over the topographically and climatically diverse Germany. Remote Sens. 10, 2029. https://doi.org/10.3390/rs10122029.
Reddy, D.S., Prasad, P.R.C., 2018. Prediction of vegetation dynamics using NDVI time series data and LSTM. Model. Earth Syst. Environ. 4, 409–419. https://doi.org/10.1007/s40808-018-0431-3.
Censimenti generali dell'agricoltura - dati di sintesi | Servizionline [WWW Document], n. d. URLhttps://servizi.regione.piemonte.it/catalogo/censimenti-generali-dellagricoltura-dati-sintesi (accessed 10.13.23).
Reuß, F., Greimeister-Pfeil, I., Vreugdenhil, M., Wagner, W., 2021. Comparison of long short-term memory networks and random forest for sentinel-1 time series based large scale crop classification. Remote Sens. 13, 5000. https://doi.org/10.3390/rs13245000.
Rockström, J., Karlberg, L., Wani, S.P., Barron, J., Hatibu, N., Oweis, T., Bruggeman, A., Farahani, J., Qiang, Z., 2010. Managing water in rainfed agriculture—The need for a paradigm shift. Agric. Water Manag., Comprehensive Assessment of Water Management in Agriculture 97, 543–550. Doi: 10.1016/j.agwat.2009.09.009.
Roy, D.K., Sarkar, T.K., Kamar, S.S.A., Goswami, T., Muktadir, M.A., Al-Ghobari, H.M., Alataway, A., Dewidar, A.Z., El-Shafei, A.A., Mattar, M.A., 2022. Daily prediction and multi-step forward forecasting of reference evapotranspiration using LSTM and Bi-LSTM models. Agronomy 12, 594. https://doi.org/10.3390/agronomy12030594.
Salazar-Gutierrez, M.R., Johnson, J., Chaves-Cordoba, B., Hoogenboom, G., 2013. Relationship of base temperature to development of winter wheat. Int. J. Plant Prod. 7 https://doi.org/10.22069/ijpp.2013.1267.
Sarvia, F., De Petris, S., Borgogno-Mondino, E., 2020. A Methodological Proposal to Support Estimation of Damages from Hailstorms Based on Copernicus Sentinel 2 Data Times Series. In: Gervasi, O., Murgante, B., Misra, S., Garau, C., Blečić, I., Taniar, D., Apduhan, B.O., Rocha, A.M.A.C., Tarantino, E., Torre, C.M., Karaca, Y. (Eds.), Computational Science and Its Applications–. Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 737–751. https://doi.org/10.1007/978-3-030-58811-3_53.
Sarvia, F., Xausa, E., De Petris, S., Cantamessa, G., Borgogno-Mondino, E., 2021. A possible role of copernicus sentinel-2 data to support common agricultural policy controls in agriculture. Agronomy 11, 110. https://doi.org/10.3390/agronomy11010110.
Sauerbrei, W., Perperoglou, A., Schmid, M., Abrahamowicz, M., Becher, H., Binder, H., Dunkler, D., Harrell, F.E., Royston, P., Heinze, G., Abrahamowicz, M., Becher, H., Binder, H., Dunkler, D., Harrell, F., Heinze, G., Perperoglou, A., Rauch, G., Royston, P., Sauerbrei, W., for TG2 of the STRATOS initiative, 2020. State of the art in selection of variables and functional forms in multivariable analysis—outstanding issues. Diagn. Progn. Res. 4, 3. https://doi.org/10.1186/s41512-020-00074-3.
Sharma, A., Jain, A., Gupta, P., Chowdary, V., 2021. Machine Learning Applications for Precision Agriculture: A Comprehensive Review. IEEE Access 9, 4843–4873. https://doi.org/10.1109/ACCESS.2020.3048415.
Small, C., 2021. Grand Challenges in Remote Sensing Image Analysis and Classification. Front Remote Sens 1.
Soccolini, A., Vizzari, M., 2023. Predictive Modelling of Maize Yield Using Sentinel 2 NDVI, in: Gervasi, O., Murgante, B., Rocha, A.M.A.C., Garau, C., Scorza, F., Karaca, Y., Torre, C.M. (Eds.), Computational Science and Its Applications – ICCSA 2023 Workshops, Lecture Notes in Computer Science. Springer Nature Switzerland, Cham, pp. 327–338. Doi: 10.1007/978-3-031-37114-1_22.
Son, N.-T., Chen, C.-F., Chen, C.-R., Guo, H.-Y., Cheng, Y.-S., Chen, S.-L., Lin, H.-S., Chen, S.-H., 2020. Machine learning approaches for rice crop yield predictions using time-series satellite data in Taiwan. Int. J. Remote Sens. 41, 7868–7888. https://doi.org/10.1080/01431161.2020.1766148.
Stepchenko, A., Chizhov, J., 2015. NDVI short-term forecasting using recurrent neural networks. Environ. Technol. Resour. Proc. Int. Sci. Pract. Conf. 3, 180–185. https://doi.org/10.17770/etr2015vol3.167.
Szabo, S., Gácsi, Z., Bertalan-Balazs, B., 2016. Specific features of NDVI, NDWI and MNDWI as reflected in land cover categories. Landsc. Environ. 10, 194–202. https://doi.org/10.21120/LE/10/3-4/13.
Tetzner, D., Thomas, E., Allen, C., 2019. A validation of ERA5 reanalysis data in the southern antarctic peninsula—ellsworth land region, and its implications for ice core studies. Geosciences 9, 289. https://doi.org/10.3390/geosciences9070289.
Torres, J.F., Hadjout, D., Sebaa, A., Martínez-Álvarez, F., Troncoso, A., 2021. Deep learning for time series forecasting: a survey. Big Data 9, 3–21. https://doi.org/10.1089/big.2020.0159.
Vorobiova, N., Chernov, A., 2017. Curve fitting of MODIS NDVI time series in the task of early crops identification by satellite images. Procedia Eng., 3rd International Conference "Information Technology and Nanotechnology", ITNT-2017, 25-27 April 2017, Samara, Russia 201, 184–195. Doi: 10.1016/j.proeng.2017.09.596.
Xu, Q.-S., Liang, Y.-Z., 2001. Monte Carlo cross validation. Chemom. Intell. Lab. Syst. 56, 1–11. https://doi.org/10.1016/S0169-7439(00)00122-2.
Yamak, P.T., Yujian, L., Gadosey, P.K., 2020. A Comparison between ARIMA, LSTM, and GRU for Time Series Forecasting, in: Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence, ACAI '19.

Association for Computing Machinery, New York, NY, USA, pp. 49–55. Doi: 10.1145/3377713.3377722.

Yamoah, C.F., Walters, D.T., Shapiro, C.A., Francis, C.A., Hayes, M.J., 2000. Standardized precipitation index and nitrogen rate effects on crop yields and risk distribution in maize. Agric. Ecosyst. Environ. 80, 113–120. https://doi.org/10.1016/S0167-8809(00)00140-7.

Zhen, Z., Chen, S., Yin, T., Gastellu-Etchegorry, J.-P., 2023. Globally quantitative analysis of the impact of atmosphere and spectral response function on 2-band enhanced vegetation index (EVI2) over Sentinel-2 and Landsat-8. ISPRS J. Photogramm. Remote Sens. 205, 206–226. https://doi.org/10.1016/j.isprsjprs.2023.09.024.