Enhancing the Precision Agriculture of AgriSenze™ by Predicting the Soil Temperature at Different Depths*

Tsegazeab Hailu Tedla[1], Sindre Søpstad[2] and Veralia G. Sanchez[1,**]

[1]University of South-Eastern Norway, Vestfold, Norway
[2]Zimmer and Peacok, Norway

Abstract

This work uses a big data set platform, *AgriSenze™*, to predict daily soil temperatures in Norway. The big dataset is primarily collected from the Norwegian University of Life Sciences (NMBU) to build an ensemble stacking regressor machine learning algorithm to predict daily soil temperatures at six different depths (2 cm, 5 cm, 10 cm, 20 cm, 50 cm, and 100 cm). This study has successfully developed a model-based soil temperature prediction with performance results comparable to most existing research and even better performance at some soil depths (R^2=0.0.9996, RMSE = 0.0 for.01286 for depth 5cm) .

Keywords

Machine learning, smart farming, artificial intelligence, cold weather, soil temperature prediction

1. Introduction

Modern agriculture faces two primary challenges: meeting rising food demand while simultaneously addressing growing environmental concerns. Smart Farming (SF) emerges as a potential solution to tackle these dual challenges in modern agricultural practices [1]. SF is a technological innovation that utilizes information and communication technology to integrate computing with the physical farming processes [2]. For the SF to make precise decisions and provide accurate information to farmers, the SF needs to understand the complex environmental ecosystems [1] which requires collecting various environmental data such as soil, weather, crop, fertilizer, farming practices, supply-chain and market analysis data. Precision agriculture (PA), a component of the SF process, employs specialized sensors and algorithms to ensure crops receive precisely what they need to maximize yield and promote sustainability [3]. PA involves gathering data on weather patterns, soil conditions, and crop data through Internet of Things sensors installed in the fields or remotely interconnected with the help of Wireless Sensor Networks. With PA, a farmer may determine which parameters are required for a healthy crop, when, and how much are required at any given time. The agricultural parameters that impact the production yield and sustainability should be collected from different sources and analysed by the PA system to produce agronomic recommendations [3].

PA is more accurate when the levels of primary nutrients (N, P, K), secondary nutrients (Ca, Mg, S, Na), and micronutrients (Co, Fe, Cu, Mn, Mo, and Zn) in plants and soils are precisely predicted, with nitrogen being the most essential nutrient [4] in agricultural production. Soil nutrient concentration management, especially for N, P, and K, has become critical in modern agriculture to maximize yields, minimize fertilizer expenditures, and minimize environmental impacts [4]. The Zimmer and Peacock AS's AgriSenze™ solution employs electrochemical sensors to measure soil nitrate-nitrogen (N-N) concentration [5] and transmits the data to *Djuli™*, an IoT cloud solutions. *Djuli™* enables real-time nutrient monitoring, advanced data analytics and collaboration environment to end users [6].

NAIS 2025: Symposium of the Norwegian AI Society, June 17–18, 2025, Tromsø, Norway
*This article was edited from a master thesis at USN
**Corresponding author.
tsegazeabtedla@gmail.com (T. H. Tedla); sindre@zimmerpeacock.com (S. Søpstad); veralia.g.sanchez@usn.no (V. G. Sanchez)
https://www.usn.no/english/about/contact-us/employees/veralia-gabriela-sanchez (V. G. Sanchez)
0000-0003-2649-7898 (V. G. Sanchez)

1.1. Aim

This work aims to use the agricultural data from AgriSenze™ platform and apply artificial intelligence to predict daily soil temperatures at six different depths (2 cm, 5 cm, 10 cm, 20 cm, 50 cm, and 100 cm).

2. Background

The desire for using artificial fertilizers to increase plant yield has grown since they were first introduced. However, this heightened consumption has led to concerns about overfertilization, and its environmental impacts attributed to nitrogen and phosphorus, prompting the implementation of fertilization policies since the 1970s to regulate the amount and application rates [7].

Overfertilization can result in runoff, causing environmental issues such as greenhouse gas emissions, eutrophication, and groundwater contamination. Conversely, under-fertilization can lead to decreased plant yields and nutrient deficiencies.In addition to the meteorological factors, various soil parameters such as soil temperature, soil moisture, soil pH, soil organic matter, cation exchange capacity, etc can impact the soil nitrate-nitrogen concentration. Depending on the soil depth, soil temperature impacts the dissolved organic matter, humification and nitrate concentration [8]. Similarly, the soil moisture influences the plant nutrient uptake depending on the wetness and dryness of the soil.

There was a limited number of research studies that employed data-driven approaches to predict soil temperature. Elsayed S. et al. [9] employed hourly data on humidity, solar radiation, rainfall, dew point, and air pressure from five locations in North Dakota as input factors to forecast soil temperature at a depth of 10 cm. The models investigated were Random Forest, Support Vector Machines (SVM), Deep Neural Networks (DNN), Linear Regression, and Long Short-Term Memory (LSTM). Mampitiya et al. [10] recently performed a comparative analysis of various models, such as XGBoost, CatBoost, LSTM, Artificial Neural Network (ANN), Bi-LSTM, Ridge, Lasso, and ElasticNet. They utilized monthly data on air temperature, relative humidity, and wind speed spanning from 1989 to 2018 for Nukus, Uzbekistan. These models were assessed for their ability to predict soil temperature at a surface and 10 cm depth.

Kisi O. et al. [11] compared models including multi-layer perceptron (MLP), radial basis neural networks (RBNN), generalized regression neural networks (GRNN), and multiple linear regression (MLR) to predict monthly soil temperatures at depths of 5 cm, 10 cm, 50 cm, and 100 cm at Mersin Station, Turkey. Other works have used ANN such as Taheri M. et al. [12]. Zare Abyaneh H. et al [13] developed an ANN and Co-active neuro-fuzzy inference system (CANFIS). Farhangmehr V. et al. [14], have manipulated CNN model to predict daily maximum soil temperature under normal, hot and cold weather conditions at Ottawa area of Canada (45.250 N latitude, 75.500 W longitude). Bilgili M. et al. [15] evaluated ANFIS network with fuzzy c-means (ANFIS-FCM), subtractive clustering (ANFIS-SC), feed-forward neural network (FNN), grid partition (ANFIS-GP), Elman neural network (ENN), and LSTM models for predicting one-day ahead soil temperature at three depths (5, 50 and 100 cm) from the previously measured time-series daily soil temperature data from 2010-2020 in the Central Anatolia Region of Turkey. Ebtehaj Isa, et al.[16] have demonstrated emotional neural network using both meteorological variable and time-series based modelling to predict daily soil temperature at different depths (10 and 20 cm) for Springfield and Champaign stations, in Illinois, USA.

Xing L. et al. [17] have proposed an SVM and combined SVM models to predict annual average and daily average soil temperatures at 5, 10, 20, 50 and 100cm using air temperature, rainfall, wind speed, solar radiation and relative humidity input data collected for 130 years from 16 different sites located in the USA with dry, warm and snowy climates. Wang X. et al. [18] investigated an integrated network prediction model utilizing gated recurrent unit (GRU) to predict soil temperature at varying depths (5, 10, and 15 cm) and different time intervals (6, 12, and 24 hours). They utilized time-series soil temperature data gathered from two meteorological stations (Laegern and Fluehli) in Switzerland spanning from 2006 to 2014.

The range of machine learning models proposed by various researchers suggest that the best performing model is highly contingent upon the specific dataset and climatic region under consideration. Therefore, there is no single soil temperature prediction model that has been universally adopted, not even at the country level. Besides, it becomes even more difficult when developing models for cold climates at national level. Imanian H. et al.[19] recommended that machine learning models should consider the cold climate's unique features, such as snow depth or other during exceptionally cold conditions, for better prediction accuracy.

Therefore, this study contributes to the research in soil temperature prediction for agriculture in cold climates, specifically using data available from Norway.

3. Methods

3.1. Data collection

The data used in this study is gathered and pre-processed with help of a web API platform integrated to the Zimmer and Peacock (ZP) AgriSenze™ solution. The data was collected manually from NMBU and MET Norway using the data-gathering web API. The web API platform is integrated to the AgriSenze solution. The Meteorological data for ÅS - BIOKLIM collected by NMBU from January 01, 2000, to April 01, 2024, was used as the basis for building the agricultural dataset in this research.

ÅS - BIOKLIM research site is located on Søråsjordet in Ås i Viken weather station with geographic coordinates 59 39' 37'' N, 10 46' 54'' E, 93.3 m above sea level. The NMBU's data collection process is discussed in detail in the official website and scientific report [20, 21]. This dataset was used to build a machine learning model to predict the daily soil temperatures at 2cm, 5cm, 10cm, 20cm, 50cm and 100cm soil depths. The data contains different meteorological parameters recorded on a daily basis, including date, minimum temperature (°C), maximum temperature (°C), mean temperature (°C), relative humidity (%), precipitation (mm), evaporation (mm), daily air pressure (mbar), wind speed (m/s), wind direction (degrees), snow depth (cm), global radiation (W/m^2), reflected radiation (W/m^2), radiation balance (W/m^2), diffuse radiation (W/m^2), photosynthetic active radiation – PAR (mE/m^2) and soil temperature at various depths (2cm, 5cm, 10cm, 20cm, 50cm, 100cm in °C).

The data parameters pertaining to the nitrate-nitrogen sensors of AgriSenze™ are obtained through the DjuliTM IoT cloud service API developed by Zimmer and Peacock AS. Although this API provides nitrate concentration data, it has not yet been deployed on the ÅS - BIOKLIM field station; hence, the data gathered is not specific to this location.

3.2. Data cleaning and preprocessing

The Jupyter Notebook and Google Colab were used for data processing and analytics. The AgriSenze™ dataset was loaded from CSV file and analysed for duplicates, missing values, anomalies, and non-numeric values. The dataset records are organized on daily values for the different parameters and duplicate date records were removed. Features from the original dataset that contained numerous missing values and were deemed to have minimal impact on daily soil temperature were removed from the dataset. Also, in the original snow depth data, there were instances where both manually recorded and automatically measured values were available for the same dates. In these cases, the manually recorded values were selected for their better reliability.

There were a large number of missing values for the snow and evaporation in the winter season (November to March). To input the missing values, time series data analysis was used to determine whether to impute the missing values or exclude the parameter from the dataset. Since 8 years of data was missing, the mean the average value for the same day across multiple years within the same month was calculated and inputted, providing a more representative estimate based on historical observations for that specific day and month.

Following this process, outlier were identified and removed using Z-score normalization. Then, Isolation Forest algorithm was used to identify anomalous data points. Finally, correlation analysis was used to identify highly correlated features aiming to avoid multicollinearity effect on the models.

Table 1
The different pre-processed time-independent shuffled and time-dependent unshuffled dataset cases used for the soil temperature prediction modelling input

Dataset Type	dataset Features (E.g. For Target ST100)
Case 1:Z-score outlier filtered dataset	mean air temperature, min air temperature, max air temperature, relative humidity, air pressure, precipitation mm, evaporation mm, earth heat flux, radiation balance, photosynthetic active radiation, albedo, snow depth, ST2, ST5, ST10, ST20, ST50, month, day
Case 6: time-series prediction	date, mean air temperature, min air temperature, max air temperature, relative humidity, air pressure, precipitation, evaporation, earth heat flux, radiation balance, photosynthetic active radiation, albedo, snow depth, ST2, ST5, ST10, ST20, ST50

3.3. Model selection, training and evaluation

The initial step in developing the machine learning model involves shuffling and splitting the pre-processed dataset into training, validation, and test subsets. The dataset with 8252 sample size after Z-score outlier filtering was divided into 70% training, 15% validation, and 15% test sets. This 70/30 splitting proportion proved to be more performant compared to the 80% for training and 20% for testing considered as common split as discussed in Géron A. [22]. The shuffling process ensures that each subset (training, validation, and test) is representative of the overall dataset. This preventive measure mitigates potential biases that might arise from the original order of the data, ultimately leading to reliable cross-validation results in subsequent steps [22].

The next step involves organizing the features, targets, and evaluation metrics to be utilized for the daily soil temperature prediction model. The targets are the daily soil temperatures at various depths, specifically 2 cm, 5 cm, 10 cm, 20 cm, 50 cm, and 100 cm. The features were initially used as independent input variables, a correlation and feature importance analysis were subsequently conducted to identify the most influential features for each target variable. Since the problem requires a regression model to predict daily soil temperature, the main evaluation metrics adopted for the prediction model includes the coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE) and root mean squared error (RMSE).

The regression models tested and compared consist of Stacking Regressor, CatBoost Regressor, HistGradientBoosting Regressor, RandomForest Regressor, AdaBoost Regressor, Ridge regressor, Support Vector Regression with both linear and non-linear kernels, Lasso Regressor, and ElasticNet Regressor. This comprehensive set of models were manipulated for a comparative analysis to identify the most suitable algorithm for the daily soil temperature prediction task. The models were fitted with the best params. Then cross-validation was performed followed by residual analysis. Residuals analysis was used as a valuable tool to evaluate the model's fit to the data. Finally, feature importance and backward feature elimination technique were employed.

4. Results

There are fourteen basic features (predictor variables) and six target (soil temperature) variables. To generate an optimized dataset for predicting soil temperature at six different depths, various preprocessing methods were employed. Table 1 summarizes the two main cases and the resulting datasets produced after applying the preprocessing steps. The dataset cases have normalized and denormalized versions and the performance of the predictions for each dataset case was evaluated. The dataset cases are generally categorized into time-independent shuffled meteorological datasets (Case 1) and time-series meteorological datasets (Case 6). The prediction performance analysis was done for the all datasets cases to identify which dataset case fits best for the models. Other dataset cases were also tested and found less performative. Among the dataset cases, Case 1 and Case 6 were identified as having superior performance.

Table 2
ST2, ST5, ST10, ST20, ST50 and ST100 prediction performance results of 5 models for the training set's 5-fold cross-validation, validation set's 1-fold validation and test set phases for the Z-score outlier filtered dataset (Case 1)

Target	Model	Evaluation Phases								
		Train set CV (5-Fold)			Validation set			Test set		
		R^2	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE
ST2	RF-R	0.9791	0.7309	0.9960	0.9767	0.6883	0.9476	0.9816	0.7650	1.0311
	XGB-R	0.9807	0.7049	0.9571	0.9800	0.6519	0.8856	0.9833	0.7271	0.9564
	CATB-R	0.9812	0.7018	0.9437	0.9802	0.6702	0.8862	0.9835	0.7268	0.9522
	HGB-R	0.9808	0.7001	0.9560	0.9797	0.6768	0.9219	0.9826	0.7258	0.9625
	STACK-R	0.9821	0.6795	0.9230	0.9811	0.6499	0.8766	0.9843	0.7059	0.9286
ST5	RF-R	0.9971	0.2514	0.3646	0.9969	0.2355	0.3489	0.9974	0.2546	0.3681
	XGB-R	0.9996	0.0978	0.1377	0.9996	0.0897	0.1251	0.9997	0.0932	0.1269
	CATB-R	0.9995	0.1150	0.1537	0.9995	0.1060	0.1398	0.9996	0.1175	0.1497
	HGB-R	0.9996	0.0996	0.1419	0.9996	0.0906	0.1237	0.9997	0.0958	0.1295
	STACK-R	0.9996	0.0925	0.1286	0.9997	0.0840	0.1138	0.9997	0.0886	0.1194
ST10	RF-R	0.9984	0.1694	0.2643	0.9986	0.1609	0.3164	0.9978	0.1586	0.2398
	XGB-R	0.9987	0.1421	0.2374	0.9990	0.1330	0.2899	0.9981	0.1332	0.2105
	CATB-R	0.9985	0.1650	0.2566	0.9987	0.1595	0.3097	0.9979	0.1560	0.2300
	HGB-R	0.9987	0.1450	0.2358	0.9990	0.1340	0.2922	0.9981	0.1323	0.2075
	STACK-R	0.9988	0.1363	0.2282	0.9990	0.1276	0.2857	0.9982	0.1265	0.2011
ST20	RF-R	0.9974	0.1972	0.3221	0.9977	0.1930	0.3728	0.9966	0.1859	0.2969
	XGB-R	0.9974	0.1937	0.3189	0.9979	0.1916	0.4089	0.9969	0.1796	0.2799
	CATB-R	0.9972	0.2124	0.3332	0.9975	0.2134	0.3922	0.9962	0.2056	0.3113
	HGB-R	0.9972	0.2021	0.3297	0.9979	0.1927	0.3917	0.9962	0.1833	0.2833
	STACK-R	0.9975	0.1828	0.3061	0.9980	0.1758	0.3741	0.9966	0.1714	0.2746
ST50	RF-R	0.9932	0.3260	0.4486	0.9866	0.3071	0.4191	0.9944	0.3327	0.6441
	XGB-R	0.9950	0.3095	0.4233	0.9860	0.2940	0.3976	0.9949	0.3189	0.6333
	CATB-R	0.9944	0.3325	0.4468	0.9859	0.3157	0.4197	0.9949	0.3341	0.6377
	HGB-R	0.9936	0.3129	0.4334	0.9862	0.2954	0.4056	0.9947	0.3171	0.6294
	STACK-R	0.9942	0.3030	0.4149	0.9865	0.2855	0.3874	0.9952	0.3081	0.6228
ST100	RF-R	0.9941	0.2434	0.3367	0.9945	0.2226	0.3128	0.9951	0.2320	0.3193
	XGB-R	0.9951	0.2259	0.3069	0.9961	0.2105	0.2803	0.9961	0.2324	0.4086
	CATB-R	0.9948	0.2370	0.3164	0.9917	0.2228	0.2993	0.9955	0.2379	0.3911
	HGB-R	0.9953	0.2270	0.3017	0.9961	0.2034	0.2767	0.9962	0.2208	0.4149
	STACK-R	0.9965	0.2165	0.2950	0.9920	0.2016	0.2730	0.9963	0.2186	0.3655

Z-score outlier filtered dataset (Case 1) results: Upon comprehensive performance evaluation of the time-independent datasets, the Z-score outlier filtered dataset (Case 1) emerged as the top performer out of the time-independent dataset cases. R^2, MAE and RMSE metrics were used as evaluation metrics. For a best model, the MAE and RMSE should be ideally the same. The performance results of the four base models (RF-R, XGB-R, CATB-R, HGB-R) and the meta-model (STACK-R) for the Case 1 dataset are listed in Table 2 for all targets.

Time-series dataset (Case 6) results: The dataset was divided into training, validation, and test sets sequentially in proportions of 70%, 15%, and 15%, respectively. As presented in Table 3, the performance results from the 10-Fold cross-validation (CV), 5-Fold CV, and test set evaluation were within a comparable range, except for ST50. In the case of ST50, there was a notable difference between the 10-Fold CV and 5-Fold CV RMSE values. Similarly, the test RMSE value significantly deviates from the CV RMSE value. For predicting soil temperatures at depths beyond 2 cm, historical soil temperatures data was used alongside the meteorological features.

Observed versus predicted values analysis: Scatter plots were employed to visualize the models' performance by comparing observed values against predicted values, examining the outliers between these two sets for each target variable. The scatter plot of the observed values against the STACK-R predicted values for soil temperatures at 2 cm, 5 cm, 10 cm, 20 cm and 100 cm depths are depicted in

Table 3
ST2, ST5, ST10, ST20, ST50 and ST100 prediction performance results of 5 models for the training set's 10-fold cross-validation, validation set's 5-fold cross-validation and test set phases for the time-series dataset (Case 6)

Target	Models	Evaluation Phases								
		Train set CV (10-Fold)			Valid. set CV (5-Fold)			Test set		
		R^2	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE
ST2	RF-R	0.9956	0.3079	0.4317	0.9941	0.3528	0.5119	0.9975	0.2457	0.3512
	XGB-R	0.9957	0.3086	0.4310	0.9951	0.3606	0.5158	0.9973	0.2597	0.3651
	CATB-R	0.9962	0.3318	0.4523	0.9718	0.7838	1.1207	0.9966	0.3065	0.4122
	HGB-R	0.9961	0.2868	0.4056	0.9953	0.3151	0.4598	0.9980	0.2296	0.3186
	STACK-R	0.9964	0.2791	0.3903	0.9963	0.3213	0.4573	0.9980	0.2260	0.3135
ST5	RF-R	0.9994	0.0961	0.1408	0.9982	0.1390	0.2771	0.9998	0.0576	0.0900
	XGB-R	0.9991	0.1112	0.1660	0.9982	0.1526	0.2788	0.9997	0.0800	0.1153
	CATB-R	0.9983	0.1822	0.2550	0.9861	0.5403	0.7719	0.9991	0.1501	0.2081
	HGB-R	0.9991	0.0971	0.1514	0.9981	0.1387	0.2887	0.9998	0.0658	0.0956
	STACK-R	0.9993	0.1003	0.1492	0.9983	0.1411	0.2727	0.9998	0.0659	0.0888
ST10	RF-R	0.9991	0.1182	0.1713	0.9974	0.1933	0.3267	0.9976	0.1321	0.3324
	XGB-R	0.9988	0.1297	0.1878	0.9972	0.2162	0.3371	0.9980	0.1204	0.3031
	CATB-R	0.9983	0.1774	0.2548	0.9906	0.4476	0.6149	0.9978	0.1607	0.3184
	HGB-R	0.9989	0.1131	0.1699	0.9975	0.1823	0.3187	0.9981	0.1065	0.2998
	STACK-R	0.9990	0.1137	0.1671	0.9976	0.1866	0.3196	0.9979	0.1122	0.3125
ST20	RF-R	0.9979	0.1426	0.2206	0.9983	0.1615	0.2518	0.9994	0.0918	0.1576
	XGB-R	0.9974	0.1649	0.2445	0.9981	0.1864	0.2673	0.9992	0.1018	0.1844
	CATB-R	0.9962	0.2264	0.3154	0.9916	0.4154	0.5694	0.9987	0.1563	0.2353
	HGB-R	0.9977	0.1482	0.2244	0.9983	0.1555	0.2496	0.9992	0.0908	0.1791
	STACK-R	0.9978	0.1429	0.2178	0.9985	0.1598	0.2374	0.9994	0.0889	0.1601
ST50	RF-R	0.9651	0.1905	0.7164	0.9982	0.1551	0.2270	0.8753	0.2450	2.1349
	XGB-R	0.9580	0.2282	0.8400	0.9981	0.1635	0.2359	0.9475	0.2848	1.8391
	CATB-R	0.9708	0.2884	0.6477	0.9893	0.4201	0.5584	0.8996	0.3444	1.9161
	HGB-R	0.9751	0.1690	0.5218	0.9986	0.1321	0.2021	0.9081	0.3890	1.8333
	STACK-R	0.9750	0.1935	0.5210	0.9986	0.1398	0.2014	0.9071	0.3634	1.8424
ST100	RF-R	0.9992	0.0662	0.0968	0.9990	0.0972	0.1389	0.9990	0.0596	0.1429
	XGB-R	0.9992	0.0695	0.1004	0.9986	0.1147	0.1611	0.9991	0.0632	0.1394
	CATB-R	0.9964	0.1553	0.2267	0.9862	0.3928	0.5049	0.9984	0.1132	0.1815
	HGB-R	0.9992	0.0628	0.0930	0.9991	0.0859	0.1258	0.9991	0.0533	0.1354
	STACK-R	0.9992	0.0689	0.0982	0.9987	0.1078	0.1353	0.9991	0.0595	0.1391

Figure 4. For an ideal model, the scatter plot would show a straight diagonal line marked in red colour in the plots indicating perfect accuracy of the model prediction.

Feature Importances: Feature importance analysis evaluates the contribution of each feature in predicting a specific target. The feature importance analysis showed that precipitation, air pressure and day have the least impact on the upper layers of the soil temperatures at ST2, ST5, ST10 and ST20. Conversely, the impact of air temperature decreases as the depth of the soil increases whereas the influence of evaporation, earth heat flux and radiation balance increases.

Based on the feature importances analysis, backward elimination (sensitivity analysis) technique was tested for few features but not exhaustively tested for every possible combination because it takes time to simulate all possible combinations of features based on the feature importances for a specific target. The snow depth and evaporation features were removed to test their impact on the prediction performance for ST2. The impact of eliminating evaporation and snow depth did not affected the prediction of ST2 significantly. The importance of evaporation and snow depth in predicting ST2 is 9.68% and 3.74%.

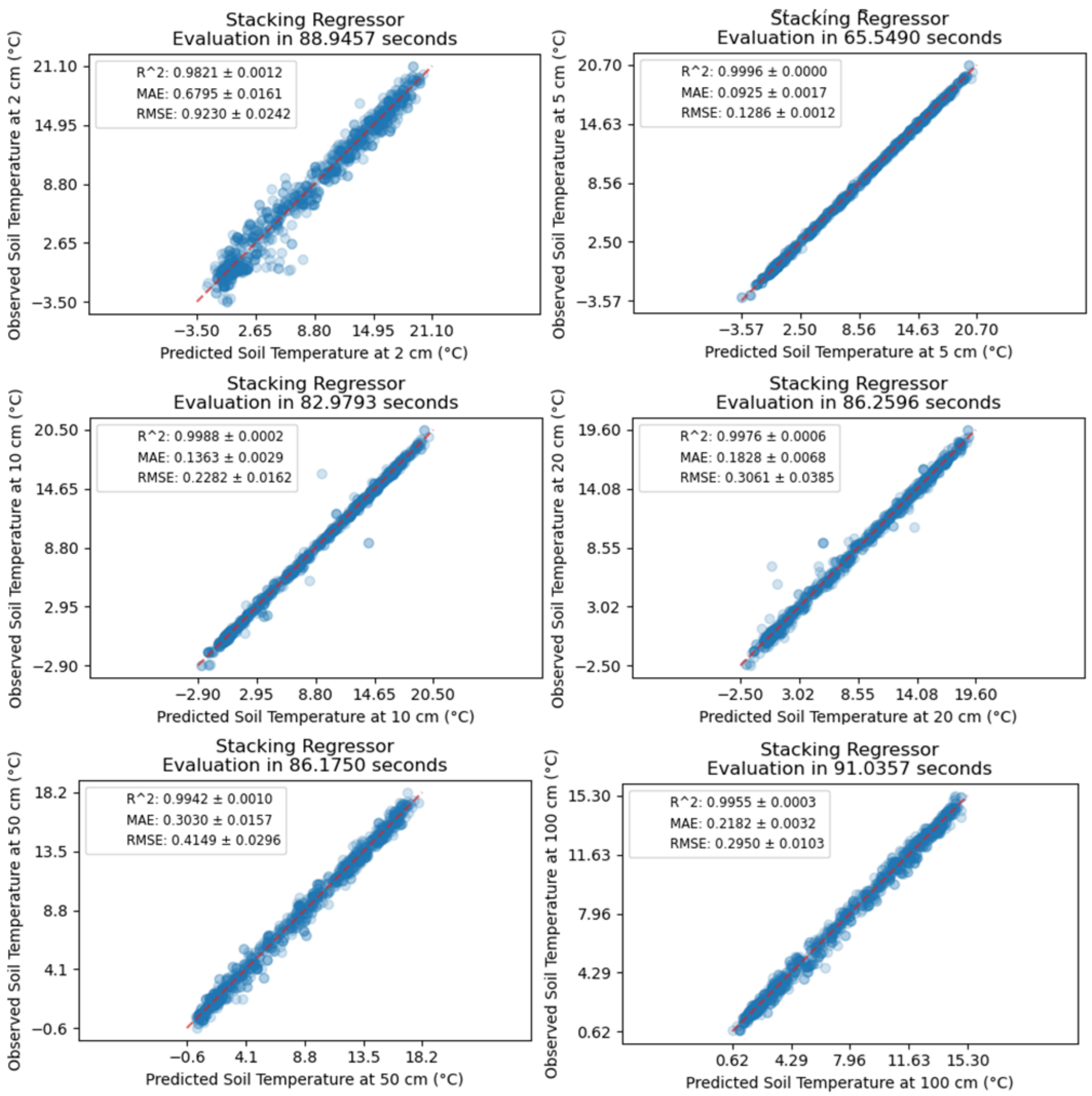

Figure 1: Scatter plots of observed values against predicted values of soil temperatures at 2, 5, 10 and 20 cm for STACK-R model for Case 1 dataset

5. Discussions

The data from AgriSenze™ solution was manipulated and valuable information was extracted to enhance decision-making. In this study, soil temperature was the primary focus of data analytics due to its major influence on plant growth and the available rich data source collected from NMBU. The soil temperature data analytics involved model-driven soil temperature predictions at six different depths: 2 cm, 5 cm, 10 cm, 20 cm, 50 cm, and 100 cm. The soil temperature prediction modelling process tested approximately 10 different models, including linear models (Lasso, ElasticNet, RidgeCV), non-linear model (SVR), tree-based models (RF-R, XGB-R, CATB-R, HGB-R), and meta-estimator (STACK-R). The study did not discuss the linear and SVR models in detail because in the preliminary testing they were found to be comparatively less efficient for the specific dataset used in this study.

According to the different pre-processing stages, various dataset cases were generated as shown previously in Table 1, and based on the preliminary assessment the candidate models (RF-R, XGB-R, CATB-R, HGB-R, STACK-R) were tested for each case. Case 1 proved to be superior to all other pre-processed datasets except the Case 6, time-series dataset which may need further validation.

For the time-independent modelling, STACK-R demonstrated superior performance across all evaluation phases. This is because the stacking regressor leverages the strengths of the base models by learning from their performance. The other four base models (RF-R, CATB-R, XGB-R, HGB-R) have nearly similar performance except the RF-R has higher errors than the other models for almost all targets 2. The cause of the minimum error occurrence in different split sets is due to the nature of the data in the split sets. The prediction error for ST2, soil temperature at 2 cm, (RMSE = 0.9286) is higher compared to other soil depths. This is because the model for ST2 relies solely on meteorological variables, while the model for ST5 also incorporates the highly correlated soil temperature data at 2 cm. Similarly, the model ST10 considers the temperatures at the two shallower depths above it. This pattern continues for all higher depth soil temperature predictions.

The prediction accuracies were checked by dropping the historical soil temperature features from dataset resulting in weaker performance, for example for target ST5, when ST2 feature is dropped out, the prediction error increased from RMSE of 0.1286 to 0.9334. This is close to the prediction error of ST2 which was solely predicted from meteorological parameters only. The study by Bilgili M. et al. [15] supports this result that incorporating historical soil temperature data in the input helps to increase the prediction accuracy.

The model for ST2 is robust for outliers and has good generalization capabilities across wide range of data points according to the residual analysis- While the most stable prediction was achieved for ST50 with mean RMSE of 0.4015 and standard deviation of 2.6% between the cross-validation errors, the most accurate prediction was achieved for ST5 with mean RMSE of 0.1234 and standard deviation of 4%. Models can have good generalization capabilities across their wider range of data points but less stable (has higher variance in its prediction error) and less accurate (has higher mean prediction error) in the overall prediction error.

The time-dependent modelling with the time-series dataset (Case 6) exhibited a significant improvement in performance for some of the targets compared to the time-independent modelling. However, further validation is necessary to ensure that the prediction is consistent for any time-series data. While an exhaustive validation could not be performed due to time constraints, very promising and interesting results were observed. The modelling results are illustrated in Table 3. With the exception of the targets ST5 and ST50, all other targets demonstrated relative improvements in performance. The most crucial improvement in this time-series modelling was observed for ST2 with almost 50% decrease in RMSE, as all other soil temperatures depend on the impact of ST2 either directly or indirectly. This represents a promising result that warrants further consideration in the future work, involving more extensive validation and analysis on time-series modelling with more advanced deep learning models. This validation is highly required because a recent study by Imanian H. et al. [19] had suggested that time-series machine learning models face challenges in accurately predicting soil temperature especially for cold climates. The evaluation of the time-series modelling on the ST50 test set yielded somewhat strange results, necessitating further validation to determine whether the test set data is causing the extreme change in RMSE or if the model is not robust for all types of time-series data. This aspect will be considered in future work.

The other essential analysis done was the feature importances and backward feature elimination process. The prediction of soil temperature at 2 cm is fully dependent on the meteorological parameters and air temperature is the most crucial parameter with overall weight 56.61% (mean + max + min air temperatures) followed by the month, evaporation and snow depth at 14.68%, 9.68% and 3.74% respectively. The backward feature elimination showed that the ST2 prediction performance remained relatively similar regardless of whether evaporation and snow depth data were included or excluded in the Case 1 dataset. This could be due to some reasons: 1) it could be because majority of the snow depth data was missing in the original dataset, hinting that the impact of snow depth could be more pronounced if the actual snow depth data were accurately measured, 2) it could be because in cold climates the topsoil surface temperature is highly influenced by ice layer temperature during the winter season and that information is not well captured in snow depth dataset, 3) the impact of snow depth on topsoil temperature is relatively negligible. Similarly, the effect of evaporation, which was filled

Table 4
Summary of few recent studies on soil temperature predictions

Reference	Country	Dataset	Method	Depth (cm)	R^2	RMSE
Hao et al. (2021) [24]	Lagern, Oensingen, Fluhli Switzerland	Daily (2004 - 2014)	EEMD-CNN	5	0.9970	0.4660
				10	0.9980	0.3750
				25	0.9990	0.2760
Imanian H. et al. (2022) [23]	Ottawa, Canada	Hourly (01 June 2020 - 31 August 2020)	Stacking	0 - 7	0.9800	0.6600
			MLP	0 - 7	0.9800	0.6700
			Deep Learning	0 - 7	0.9800	0.6600
Bilgili M. et al. (2023) [15]	Sivas, Turkey	Daily (2010 - 2019)	ENN	5	0.9900	1.3186
			ANFIS-FCM	50	0.9988	0.1388
			MARS-GP	100	0.9999	0.0634
Imanian H. et al. (Dec 2023) [19]	Iqaluit, Canada	Hourly (2021)	MLP-shuffled	0 - 7	0.9498	1.4600
			MLP-shuffled	0 - 7	0.9660	1.1825
	Churchill, Canada					
Mampitiya L. et al. (2024) [10]	Nukus, Uzbekistan	Monthly (1989 – 2018)	Bi-LSTM	Soil surface	0.9450	4.1290
			LSTM	10	0.9630	2.5120
This Study (2024)	Ås, Norway (Cold climate)	Daily (01 Jan 2000 – 01 April 2024)	Stacking Regressor (RF-R, CATB-R, XGB-R, HGB-R)	2	0.9843	0.9286
				5	0.9996	0.1286
				10	0.9990	0.2857
				20	0.9980	0.3747
				50	0.9952	0.6228
				100	0.9963	0.3837

by using the daily average data across all the years and zero for the cold season, could be increased if the actual measured data was accurate in the dataset. The feature importance analysis revealed that mean air temperature has a major influence on predicting soil temperatures at shallower depths, and its influence reduces as the soil depth increases. Conversely, precipitation had the least impact on soil temperature predictions across all depths. The feature sensitivity analysis conducted by Imanian H. et al. [19], on soil temperature predictions, for cold climates at 0-7cm depth strengthened similar findings. The evaporation, earth heat flux and radiation have considerable impact on the prediction of deeper soil depths, supporting the assertions by Farhangmehr V. et al. [14] that features should not be disregarded without a thorough analysis of feature importance if prediction accuracy is the primary concern. The month and mean air temperature emerged as significant features for most target variables. The influence of the month feature increased as the soil depth increased, potentially due to the soil temperature at deeper depths taking longer to change. As a result, monthly variations in meteorological parameters are more pronounced than daily fluctuations at these deeper soil levels. The soil temperatures at greater depths are affected by the long-term averaged effects of meteorological parameters.

Even though direct one-to-one quantitative comparison is not recommended due to many research contexts, contextually the performance results in table 4 indicate that the modelling in this study has close performance metrics to the recent models especially on cold weather climates by Imanian H. et al. [19, 23] and for some of the soil depths even better performances are recorded in this study.

6. Conclusion

The dataset consisted of fourteen independent meteorological variables and six dependent soil temperature variables at six soil depths. Considering the soil temperature variables at shallower depths as inputs for predicting temperatures at deeper depths helped increase the prediction accuracy. The soil temperature prediction models yielded comparably good results, with the potential for future improvements expected to enhance the accuracy of certain models, such as ST2 and ST50.

Research on soil temperature prediction for cold and snowy climates is rare and relatively challenging compared to ordinary and hot climates due to the less impact of air temperature on soil temperature during colder periods. Nevertheless, this study has successfully developed a model-based soil temperature prediction with performance results comparable to most existing research and even better performance at some soil depths. Although it is difficult to compare the prediction performance results directly with the results from other research without contextualizing the dataset, location, season, depth of soil and scaling factors used by the research, some results taken from the recent research are presented here for contextual analysis.

Acknowledgments

The authors would like to thank Zimmer and Peacok for their involvement and guidance in the project.

Declaration on Generative AI

The author(s) have used Claude 3.7 Sonnet to generate the Latex code for the tables.

References

[1] A. Kamilaris, A. Kartakoullis, F. X. Prenafeta-Boldú, A review on the practice of big data analysis in agriculture, Computers and Electronics in Agriculture 143 (2017) 23–37. URL: https://www.sciencedirect.com/science/article/pii/S0168169917301230. doi:10.1016/j.compag.2017.09.037, publisher: Elsevier B.V.

[2] S. Wolfert, L. Ge, C. Verdouw, M. J. Bogaardt, Big Data in Smart Farming – A review, Agricultural Systems 153 (2017) 69–80. doi:10.1016/j.agsy.2017.01.023, publisher: Elsevier Ltd.

[3] U. Shafi, R. Mumtaz, J. García-Nieto, S. A. Hassan, S. A. R. Zaidi, N. Iqbal, Precision agriculture techniques and practices: From considerations to applications, Sensors (Switzerland) 19 (2019). doi:10.3390/s19173796, publisher: MDPI AG.

[4] F. Villalobos, E. Fereres, Principles of Agronomy for Sustainable Agriculture, Springer International Publishing, 2016. doi:10.1007/978-3-319-46116-8, publication Title: Principles of Agronomy for Sustainable Agriculture.

[5] Zimmer and Peacock AS, Nitrate Concentration Sensor, 2024.

[6] Zimmer and Peacock AS, Djuli cloud platform, 2024.

[7] F. J. Villalobos, A. Delgado, A. López-Bernal, M. Quemada, FertiliCalc: A Decision Support System for Fertilizer Management, International Journal of Plant Production 14 (2020) 299–308. doi:10.1007/s42106-019-00085-1, publisher: Springer.

[8] E. R. Toosi, J. P. Schmidt, M. J. Castellano, Soil temperature is an important regulatory control on dissolved organic carbon supply and uptake of soil solution nitrate, European Journal of Soil Biology 61 (2014) 68–71. doi:10.1016/j.ejsobi.2014.01.003.

[9] S. Elsayed, M. Gupta, G. Chaudhary, S. Taneja, H. Gaur, M. Gad, M. Hamdy Eid, A. Kovács, S. Péter, A. Gaagai, U. Schmidhalter, Interpretation the Influence of Hydrometeorological Variables on Soil Temperature Prediction Using the Potential of Deep Learning Model, Knowledge-Based Engineering and Sciences 4 (2023) 55–77. doi:10.51526/kbes.2023.4.1.55-77, publisher: Knowledge-based Engineering and Sciences.

[10] L. Mampitiya, K. Rozumbetov, N. Rathnayake, V. Erkudov, A. Esimbetov, S. Arachchi, K. Kantamaneni, Y. Hoshino, U. Rathnayake, Artificial intelligence to predict soil temperatures by development of novel model, Scientific Reports 14 (2024). doi:10.1038/s41598-024-60549-x, publisher: Nature Research.

[11] O. Kisi, M. Tombul, M. Z. Kermani, Modeling soil temperatures at different depths by using three different neural computing techniques, Theoretical and Applied Climatology 121 (2015) 377–387. doi:10.1007/s00704-014-1232-x, publisher: Springer-Verlag Wien.

[12] M. Taheri, H. K. Schreiner, A. Mohammadian, H. Shirkhani, P. Payeur, H. Imanian, J. H. Cobo, A Review of Machine Learning Approaches to Soil Temperature Estimation, Sustainability (Switzerland) 15 (2023). doi:10.3390/su15097677, publisher: MDPI.

[13] H. Zare Abyaneh, M. Bayat Varkeshi, G. Golmohammadi, K. Mohammadi, Soil temperature estimation using an artificial neural network and co-active neuro-fuzzy inference system in two different climates, Arabian Journal of Geosciences 9 (2016). doi:10.1007/s12517-016-2388-8, publisher: Springer Verlag.

[14] V. Farhangmehr, J. H. Cobo, A. Mohammadian, P. Payeur, H. Shirkhani, H. Imanian, A Convolutional Neural Network Model for Soil Temperature Prediction under Ordinary and Hot Weather Conditions: Comparison with a Multilayer Perceptron Model, Sustainability (Switzerland) 15 (2023). doi:10.3390/su15107897, publisher: MDPI.

[15] M. Bilgili, S. Ünal, A. Şekertekin, C. Gürlek, Machine Learning Approaches for One-Day Ahead Soil Temperature Forecasting, Tarim Bilimleri Dergisi 29 (2023) 221–238. doi:10.15832/ankutbd.997567, publisher: Ankara University.

[16] I. Ebtehaj, H. Bonakdari, P. Samui, B. Gharabaghi, Multi-depth daily soil temperature modeling: meteorological variables or time series?, Theoretical and Applied Climatology 151 (2023) 989–1012. doi:10.1007/s00704-022-04314-y, publisher: Springer.

[17] L. Xing, L. Li, J. Gong, C. Ren, J. Liu, H. Chen, Daily soil temperatures predictions for various climates in United States using data-driven model, Energy 160 (2018) 430–440. doi:10.1016/j.energy.2018.07.004, publisher: Elsevier Ltd.

[18] X. Wang, W. Li, Q. Li, A New Embedded Estimation Model for Soil Temperature Prediction, Scientific Programming 2021 (2021). doi:10.1155/2021/5881018, publisher: Hindawi Limited.

[19] H. Imanian, A. Mohammadian, V. Farhangmehr, P. Payeur, D. Goodarzi, J. Hiedra Cobo, H. Shirkhani, A comparative analysis of deep learning models for soil temperature prediction in cold climates, Theoretical and Applied Climatology 155 (2024) 2571–2587. doi:10.1007/s00704-023-04781-x, publisher: Springer.

[20] NBMU, Meteorological data for Ås - BIOKLIM, 2024. URL: https://www.nmbu.no/forskning/grupper/meteorologiske-data.

[21] Wolff. M, Grimenes A.A, Meteorologiske data for Ås, Scientific report, Norwegian University of Life Sciences (NMBU) 2000-2024, Technical Report, NMBU, ????

[22] A. Géron, Hands-On Machine Learning with Scikit-Learn and TensorFlow, Technical Report, 2017.

[23] H. Imanian, J. Hiedra Cobo, P. Payeur, H. Shirkhani, A. Mohammadian, A comprehensive study of artificial intelligence applications for soil temperature prediction (2022). URL: www.preprints.org. doi:10.20944/preprints202202.0101.v2.

[24] H. Hao, F. Yu, Q. Li, Soil Temperature Prediction Using Convolutional Neural Network Based on Ensemble Empirical Mode Decomposition, IEEE Access 9 (2021) 4084–4096. doi:10.1109/ACCESS.2020.3048028, publisher: Institute of Electrical and Electronics Engineers Inc.