UTS UNIVERSITY OF TECHNOLOGY SYDNEY

AN INTEGRATED CONCEPTUAL MODEL TOWARDS SUSTAINABLE RURAL WATER MANAGEMENT BASED REMOTE SENSING AND MACHINE LEARNING

by Thu Thuy Nguyen

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of Prof. Huu Hao Ngo, Prof. Wenshan Guo, and Dr. Yiwen Liu

University of Technology Sydney

Faculty of Engineering and Information Technology

May 2022

CERTIFICATION OF ORIGINAL AUTHORSHIP

I, Thu Thuy Nguyen declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Civil and Environmental Engineering/Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note:
Signature removed prior to publication.

Date: 10 May 2022

ACKNOWLEDGEMENTS

First of all, it is a pleasure to express my deepest gratitude to my supervisors, Prof. Huu Hao Ngo, Prof. Wenshan Guo, and Dr. Yiwen Liu. This Ph.D. thesis would not have been achievable without their continuous support and guidance. Especially it is my privilege to express a deep sense of gratitude to my principal supervisor, Prof. Huu Hao Ngo, for his unending patience, endless support, and brilliant ideas during my Ph.D. study and research. His DIPP philosophy: the spirit of Dedication, Inspiration, Passion, and Persistence has made me inspired. I humbly appreciate him for what he has provided me and for the time he has spent to help me revise my manuscripts and thesis to complete this study. I am highly grateful to him for his steadfast encouragement, unwavering guidance, and great sense of humor so that I can overcome challenges during my Ph.D. life. I would like to extend my heartfelt thanks to my co-supervisor, Prof. Wenshan Guo, and Dr. Yiwen Liu for their valuable comments and suggestions for completing my Ph.D. study.

I am deeply indebted to the Vietnam International Education Development (VIED) and the University of Technology Sydney (UTS) for providing me VIED-UTS Scholarship for my Ph.D. program. I am thankful to the Australian Postgraduate Research Intern (APR. Intern), Astron Environmental Services Company, and Food Agility CRC for data-sharing and supporting my Ph.D. research.

My special thanks to Dr. Tien Dat Pham, Dr. Kinh Bac Dang, and Dr. Chinh Luu for their data sharing and valuable suggestions. I am thankful to administrative staff and academic officers from the Graduate Research School of UTS, and the Institute for Agricultural Environment, especially Ms. Van Le from the School of Civil and Environmental Engineering, for their unparalleled support. Thanks should also go to my excellent schoolmates, including Loan, Hang, Dora, Khan, Jerry, and Phong, who shared their experience and knowledge. I appreciate my friends, An Le, Chelsey Vu, Allie Nguyen, Quyen Nguyen, Thuy Nguyen, and Minh Vu, who have willingly provided me with assistance during my Ph.D. life.

Finally, the most tremendous and sincere gratitude must be sent to my parents and brother, who have been my constant source of love and strength all my Ph.D. years.

Especially my most profound appreciation and respect should go to my grandma for her encouragement, kindness, and unconditional love. I am also grateful to my uncles, ant, and relatives for their support and encouragement during my study. Last but not least, I would love to thank my husband and my son, who is always with me, support me, and give me love and encouragement. They are my love, strength, and motivation that make me overcome any difficulties in my life. Without their unparalleled love, constant support, and care, I could not pursue and completed my Ph.D. study successfully. This dissertation is dedicated to my extended family, husband, and little boy.

TABLE OF CONTENTS

CERTIFICATION OF ORIGINAL AUTHORSHIP ---------- **i**

ACKNOWLEDGEMENTS ---------- **ii**

TABLE OF CONTENTS ---------- **iv**

LIST OF TABLES ---------- **ix**

LIST OF FIGURES ---------- **xi**

LIST OF ABBREVIATIONS ---------- **xiv**

RESEARCH OUTCOMES ---------- **xix**

Ph.D. DISSERTATION ABSTRACT ---------- **xxii**

Chapter 1 ---------- ***1***

Introduction ---------- **1**

1.1 Research background ---------- **2**

1.2 Research scope and objectives ---------- **5**

1.3 Research significance ---------- **6**

1.4 Thesis structure ---------- **6**

Chapter 2 ---------- ***9***

Literature Review ---------- **9**

2.1 Introduction ---------- **10**

2.2 Evolution of water management practices ---------- **12**

2.2.1 Drivers for water management strategies ---------- 12

2.2.2 Evolution of water management practices ---------- 13

2.2.3 Recent water management initiative –Sponge City ---------- 20

2.3 Challenges and opportunities of water management strategies adoption 26

2.3.1 Technical and physical challenges ---------- 26

2.3.2 Financial challenges ---------- 29

2.3.3 Administrative fragmentation ---------- 29
2.3.4 Public awareness and acceptance challenges---------- 30
2.3.5 Opportunities and future perspective for water infrastructure implementation ---------- 30

2.4 Water management monitoring and modelling----------37
2.4.1. Overview of current water modelling ---------- 37
2.4.2 Remote sensing and machine learning application for water monitoring -- 48
2.4.3 Barriers of existing integrated water planning models ---------- 50

2.5 Conclusions----------52

Chapter 3 ---------- 54

Research Methodology ----------54

3.1 Materials----------55
3.1.1 Spatial datasets ---------- 55
3.1.2 Statistical data ---------- 57
3.1.3 Field survey data ---------- 57

3.2 Methods----------58
3.2.1 Spatial data processing ---------- 58
3.2.2 Vulnerability assessment---------- 65
3.2.3 Field data collection ---------- 67
3.2.4 Machine learning algorithms ---------- 68
3.2.5 Genetic Algorithm (GA) optimizer for optimal feature selection ---------- 70
3.2.6 Model evaluation---------- 70

Chapter 4 ---------- 72

A new integrated conceptual model for sustainable rural water infrastructures implementation ----------72

4.1 Introduction----------73

4.2. Conceptual model principles----------74
4.2.1 Selection of model's features---------- 74

4.2.2 Selection of relevant sub-models ---------- 76
4.2.3 Selection of model variables ---------- 78

4.3 Validation and calibration process of integrated model ----------79

4.4 Mapping model uncertainties ----------80

4.5 The proposed model's structure ----------82

4.5 Conclusions ----------84

Chapter 5 ---------- 86

New approach of water quantity vulnerability assessment using satellite images and GIS-based model ----------86

5.1 Introduction----------87

5.2 Materials and methods ----------89

5.2.1 Study area ---------- 89

5.2.2 Data acquisition ---------- 91

5.2.3 Data analysis ---------- 91

5.3 Results and Discussion----------91

5.3.1 Water quantity vulnerability assessment framework ---------- 91

5.3.2 Selection of water vulnerability indicators ---------- 93

5.3.3 Mapping of satellite data ---------- 99

5.3.4 The weights of indicators ---------- 104

5.3.5 Spatial distribution of water vulnerability ---------- 107

5.4 Overall discussion ---------- 111

5.5 Conclusions ---------- 113

Chapter 6 ---------- 115

Exploring next generation spatial modelling of soil moisture ---------- 115

6.1 Introduction---------- 116

6.2 Materials and methods ---------- 119

6.2.1 Study area and soil sample collection ---- 119
6.2.2 Research framework ---- 120
6.2.4 Machine learning algorithms ---- 121
6.2.5 Genetic algorithm (GA) for feature selection ---- 121
6.2.6 Model performance evaluation ---- 122

6.3 Results and discussion ---- 122
6.3.1 Correlation analysis of predictor indicators and measured SM ---- 122
6.3.2 Evaluation and comparison of scenarios and different ML models ---- 124
6.3.3 Spatial distribution patterns of SM maps ---- 126
6.3.4 Relative importance of SM prediction indicators ---- 127

6.4 Conclusions ---- 129

Chapter 7 ---- 131

A novel intelligence approach for agricultural soil organic carbon prediction ---- 131

7.1 Introduction ---- 132

7.2 Materials and methods ---- 133
7.2.1 Study area ---- 133
7.2.2 Soil samples collection ---- 134
7.2.3 Research framework ---- 134
7.2.4 Remote sensing data acquisition and image processing ---- 135
7.2.5 Scenarios development ---- 135
7.2.6 Machine learning techniques ---- 136
7.2.7 Model performance evaluation ---- 136

7.3 Results ---- 136
7.3.1 Land-use binary mapping ---- 136
7.3.2 Evaluation and comparison of scenarios and different ML models ---- 141
7.3.3 Spatial distribution patterns of agricultural SOC maps ---- 143

7.4 Discussion ---- 145

7.4.1 Performance of agricultural SOC prediction models ---------- 145
7.4.2 Relative importance of predictor variables ---------- 146
7.5 Conclusions ---------- 147
Chapter 8 ---------- 149
Conclusions and recommendations ---------- 149
8.1 Conclusions ---------- 150
8.2 Recommendations ---------- 152
References ---------- 154
Appendix ---------- 186

LIST OF TABLES

Table 2.1 A comparison of different drainage systems from 3000BC to the present 14
Table 2.2 Challenges and outlook for future development of water management strategies..... 32
Table 2.3 Applying different integrated water modelling for water strategies assessment..... 41
Table 2.4 Prediction performance of agricultural SOC in the recent literature. 48
Table 3.1 Spatial variables and datasets used in water quantity vulnerability assessment..... 55
Table 3.2 Remote sensing data acquisition for the study areas..... 57
Table 3.3 Vegetation, soil, and water predictor variables derived from Sentinel 2 (modified from (Pham et al., 2020)) 60
Table 4.1 Key model features of the rural water model..... 74
Table 5.1 The selected indicators for water vulnerability assessment 93
Table 5.2 The Classification of climate types based on aridity index (adapted from (Trabucco and J. Zomer, 2018))..... 101
Table 5.3 Results of the weight calculation..... 105
Table 5.4 Results of water quantity vulnerability assessment for the ecological zones..... 107
Table 6.2 Lists of developed scenarios for soil moisture estimation 121
Table 6.3 Pearson's correlation analysis of input variables and measured SM 122
Table 6.5 Performance comparison of ML algorithms on agricultural SM estimation..... 125
Table 7.1 Model's performance of land-use binary mapping using S-2 dataset..... 137
Table 7.2 Pearson's correlation analysis of S-2 derived predictor indicators and measured SOC..... 139
Table 7.3 Pearson's correlation analysis of S-1 derived predictor indicators and measured SOC..... 140

Table 7.4 Model performance of the XGBoost technique in five scenarios 141
Table 7.5 Performance comparison of ML algorithms on agricultural SOC estimation 142

LIST OF FIGURES

Figure 1.1 Summary of water infrastructure implementation process 4
Figure 1.2 Thesis structure 7
Figure 2.1. The main causes of flooding 13
Figure 2.2 Differences between conventional and sustainable approaches of water systems 17
Figure 2.3 Schematic design of the Sponge City and Sponge City Ecological Services 19
Figure 2.4 Sponge City's principle concepts and objectives 21
Figure 2.5 Bio-retention design 24
Figure 2.6 Difference between impervious surfaces and pervious surfaces 26
Figure 2.7 Required data for water management design 28
Figure 2.8 Key aspects to successful implementation of water management practices 31
Figure 2.9 Sub-models of integrated water modelling. 38
Figure 2.10 Classification of integrated water models (adapted from Bach et al., 2014). 39
Figure 2.11 The limitation of existing integrated water management models... 40
Figure 3.1 Incorporation of spatial datasets, statistical data set and the study site 58
Figure 3.2 The steps of Sentinel images processing using SNAP Toolbox 59
Figure 3.3 Steps of Sentinel 1 pre-processing and processing 62
Figure 3.4 Indices generated from ALOS DSM: (a) DEM and (b) SLOPE 64
Figure 3.5 TWI mapping in the study site 65
Figure 3.6 Flow chart of land-use binary mapping and SOC samples selection using an active learning method. 67
Figure 3.7 Study areas and digitizing point selection: (a) Wests, and (b) Cookies 68
Figure 4.1 Overall development proposal of the sustainable rural water management model 78

Figure 4.2 A generic framework of the integrated rural water model development (modified from Sargent, 1991). 80
Figure 4.3 The main components of uncertainties in the integrated rural water model (modified from Deletic et al., 2012). 81
Figure 4.4 The possible structure of integrated rural water model 83
Figure 5.1 List of potential water vulnerability indicators (modified from (Plummer et al., 2012)) 88
Figure 5.2 Location map of the study area 90
Figure 5.3 The framework of water vulnerability assessment 92
Figure 5.4 The example of a GIS-based model for precipitation data processing 100
Figure 5.5 Mapping of spatial distribution for Vietnam's provincial remote sensing indicators 103
Figure 5.6 Spatial distribution for provincial exposure index (a), sensitivity index (b) and adaptive capacity index (c) 109
Figure 5.7 Spatial distribution of provincial water quantity vulnerability index (a), and ecological water quantity vulnerability index (b) 110
Figure 5.8 Linear regression plots 112
Figure 6.1 Location of the study sites and sampling points in Wests and Cookies area 120
Figure 6.3 Scatter plots of the measured SM and estimated SM using (a) XGBR, (b) CBR, (c) RFR, (d) and SVM. 126
Figure 6.4 Maps of SM content in study areas: (a) Wests and (b) Cookies using XGBR and multiple data fusion. 127
Figure 6.5 Variable importance of optimal features derived from multi-source EO data. 129
Figure 7.1 A novel established framework of agricultural SOC prediction using multi-sensor data fusion 135
Figure 7.2 Land use binary classification map derived from the XGBoost model using S-2 and sampling points selection: (a) Wests, and (b) Cookies 138
Figure 7.3 Scatter diagrams of the measured SOC and estimated SOC by (a)

XGBoost, (b) RF, (c) and SVM. 143
Figure 7.4 Spatial distribution characteristic of agricultural SOC in study areas: (a) Wests (a) and (b) Cookies using the proposed XGBoost combined data fusion. 144
Figure 7.5. Variable importance of optimal features derived from multi-source EO data. 147

LIST OF ABBREVIATIONS

AC	Adaptive capacity
AHP	Analytical hierarchy process
AI	Aridity index
ALOS	Advanced land observing satellite
ANN	Artificial neural networks
ASM	Advanced scatter meter
BI	Brightness index
BI2	Brightness index 2
BMPs	Best management practices
BMPs	Best management practices
BRT	Boosted regression trees
CBR	CatBoost gradient boosting regression
CCFSC	Central committee for flood and storm control
CH	Central Highland
CHIRPS	Climate hazards group infrared precipitation with station data
CI	Colour index
CV	Cross validation
DEM	Digital elevation model
DGPS	Differential global positioning system
DPSIR	Driver pressure state impact response
DSM	Digital surface model
E	Exposure
ELM	Extreme learning machine
EO	Earth observation
EPA	Environmental protection agency

ESA	Ecosystem services assessment
FullCAM	Full Carbon Accounting Model
GA	Genetic algorithm
GA	Grass pavement
GBI	Global international geosphere-biosphere programme
GIS	Geographic information system
GLMC	Grey level co-occurrence
GNDVI	Green normalized difference vegetation index
GSD	Ground sampling distance
HDI	Human development index
ICBMs	Integrated component-based models
IPCC	Inter-government panel on climate change
IRECl	Inverted red-edge chlorophyll index
IUDMs	Integrated urban drainage models
IUWCMs	Integrated urban water cycle models
IUWM	Integrated urban water management system
IUWSMs	Integrated urban water system models
KC	Kappa coefficient
LAI	Leaf area index
LCA	Life cycle assessment
LID	Low impact development
MCA	Multi-criteria analysis

MCARI	Modified chlorophyll absorption in reflectance index
MD	Mekong Delta
ML	Machine learning
MODIS	Moderate resolution imaging spectroradiometer
NCC	North Central Coast
NDI45	Normalized difference index using bands 4 & 5 of S-2
NDVI	Normalized difference vegetation index
NDWI	Normalized difference water index
NE	Northeast
NPP	Net primary production
NW	Northwest
OA	Overall accuracy
OA	Overall accuracy
P	Precision
PA	Permeable asphalts
PAC	Priestley–taylor alpha coefficient
PC	Permeable concretes
PICP	Permeable interlocking concrete pavers
PLSR	Partial least squares regression
PPPs	Private public partnerships
R	Recall
R^2	Coefficient of determination
RF	Random forest
RFR	Random forest regression
RI	Redness index

RMSE	Root mean square error
RRD	Red River Delta
RS	Remote sensing
RUE	Rain use efficiency
RVI	Ratio vegetation index
RWH	Rainwater harvesting
RWH	Rainwater harvesting
S	Sensitivity
SAR	Synthetic aperture radar
SAVI	Soil adjusted vegetation index
SCC	South Central Coast
Sentinel 1	S-1
Sentinel 2	S-2
SIs	Soil indices
SM	Soil moisture
SMAP	Soil moisture active passive
SMOS	Soil moisture and ocean salinity
SOC	Soil organic carbon
SuDS	Sustainable drainage system
SVM	Support vector machine
SW	South West
SWMM	Storm water management model
SWS	Soil water stress
TWI	Topographic wetness index
UASs	Unmanned aerial systems
UTM	Universal transverse Mercator
VIs	Vegetation indices
WA	Western Australia
WSC	Water sensitive city

WSE	Wrapper subset evaluator
WSUD	Water sensitive urban design
WUE	Water use efficiency
WVI	Water vulnerability index
XGBoost	Extreme gradient boosting
XGBR	Extreme gradient boosting regression

RESEARCH OUTCOMES

A. Peer-Reviewed Journal Articles

1. **Nguyen, T.T**., Ngo, H.H., Guo, W., Nguyen, H.Q., Luu, C., Dang, K.B., Liu, Y., Zhang, X. 2020a. New approach of water quantity vulnerability assessment using satellite images and GIS-based model: An application to a case study in Vietnam. *Science of The Total Environment*, **737**, 139784 (IF: **7.963**; SJR: **Q1**).
2. **Nguyen, T.T**., Ngo, H.H., Guo, W., Wang, X.C. 2020b. A new model framework for sponge city implementation: Emerging challenges and future developments. *Journal of Environmental Management*, **253**, 109689 (IF: **6.789**; SJR: **Q1**).
3. **Nguyen, T.T**., Ngo, H.H., Guo, W., Wang, X.C., Ren, N., Li, G., Ding, J., Liang, H. 2019. Implementation of a specific urban water management - Sponge City. *Science of The Total Environment*, **652**, 147-162 (IF: **7.963**; SJR: **Q1**).
4. **Nguyen, T.T**., Pham, T.D., Nguyen, C.T., Delfos, J., Archibald, R., Dang, K.B., Hoang, N.B., Guo, W., Ngo, H.H. 2022. A novel intelligence approach based active and ensemble learning for agricultural soil organic carbon prediction using multispectral and SAR data fusion. *Science of The Total Environment*, **804**, 150187 (IF: **7.963**; SJR: **Q1**).
5. Dang, K.B., **Nguyen, T.T.**, Ngo, H.H., Burkhard, B., Müller, F., Dang, V.B., Nguyen, H., Ngo, V.L., Pham, T.P.N. 2021. Integrated methods and scenarios for assessment of sand dunes ecosystem services. Journal of Environmental Management, 289, 112485 (IF: **6.789**; SJR: **Q1**).
6. **Nguyen, T.T.**, Ngo, H.H., Guo, W., Chang, S.W., Nguyen, D.D., Nguyen, C.T., Zhang, J., Liang, S., Bui, X.T., Hoang, N.B. 2022. A low-cost approach for soil moisture prediction using multi-sensor data and machine learning algorithm. *Science of The Total Environment*, 833, 155066 (IF: **7.963**; SJR: **Q1**).
7. **Nguyen, T.T**. 2021. Predicting agricultural soil carbon using machine learning. *Nature Reviews Earth & Environment*, **2**(12), 825-825.
8. **Nguyen, T.T**., Ngo, H.H., Guo, W. Fusion of feature selection optimizer and advance machine learning algorithm for improvement of soil carbon prediction. (in preparation).

B. Book chapters

1. Huu Hao Ngo, **Thu Thuy Nguyen**, Wenshan Guo, Dinh Duc Nguyen, Ashok Pandey, Xuan Thanh Bui, Sunita Varjani, Phuoc Dan Nguyen, Chapter 11: Circular bioeconomy for resource recovery from wastewaters using algae-based technologies, In the book series on Current Developments in Biotechnology and Bioengineering: Algae-based biomaterial for sustainable development: biomedical, environmental remediation and sustainability assessment, Huu Hao Ngo, Wenshan Guo, Ashok Pandey, Jo-Shu Chang, Duu-Jong Lee (Eds), Elservier, (In press).
2. Huu Hao Ngo, **Thu Thuy Nguyen**, Wenshan Guo, Lijuan Deng, Sunita Varjani, Yi Liu, Chapter 16: Sustainability assessment of biochar for climate change mitigation, In the book series on Current Developments in Biotechnology and Bioengineering: Biochar towards sustainable environment, Huu Hao Ngo, Wenshan Guo, Ashok Pandey, Sunita Varjani, Daniel CW Tsang (Eds), Elservier, (In press).

C. Conference paper

1. Thu Thuy Nguyen, Huu Hao Ngo, Wenshan Guo, Xiaochang C. Wang. 2019. Developing a Conceptual Water Model for Sponge City. Green Technologies for Sustainable Water (GTSW). Ho Chi Minh City, Vietnam, 1 - 5, December.

D. Research awards

1. Higher Degree Top-Up Scholarship from Food Agility CRC, 2021 for the project: An integrated model for sustainable rural water management based on machine learning and remote sensing, 2021-2022, $10,000 per year.
2. Australian Postgraduate Research Intern offer by Australia Government and Astron Environmental Consulting Company, 2021 for carrying out an internship project: Investigating the use of satellite data for soil carbon monitoring within agricultural areas of Australia, 2021, $9000/ 3months.
3. UTS Thesis Equity Grant, 2021
4. Higher Degree Research Female Top-up Scholarship from the University of Technology, Sydney, 2019 for the outstanding achievement of female students, 2019,
5. PROM Program- International scholarship exchange of PhD students and

academics, 2019” from Poland Government, 2019.

6. FEIT-UTS International Conference Grant Approval, 2019

Ph.D. DISSERTATION ABSTRACT

Author: Thu Thuy Nguyen

Date: 23 February 2022

Thesis title: An integrated conceptual model towards sustainable rural water management based remote sensing and machine learning

Faculty: Faculty of Environmental and Information Technology

School: Civil and Environmental Engineering

Supervisors: Prof. Huu Hao Ngo (Principal supervisor)
Prof. Wenshan Guo (Co-supervisor)
Dr. Yiwen Liu (Co-supervisor)

Abstract

Recently, there have been some improvements in agricultural water supply systems. However, rural areas still face serious water deficiencies including droughts, poor water quality and floods due to inappropriate of water management systems and climate change. This critical issue points out the urgent need for developing an effective integrated rural water model, which can improve water monitoring in rural regions. This study therefore aims to develop the integrated conceptual model for rural sustainable water monitoring to help rural communities overcome the issues of water run-off, water pollution and lack of water for agricultural production.

This thesis presents a novel conceptual model framework including three sub-

models (water vulnerability quantity assessment model, soil moisture prediction model, and agricultural soil organic carbon model for supporting rural water modelling using the integration of free-of-charge satellite images including MODIS, Sentinel 1, Sentinel 2, and ALOS DSM imagery and different advanced machine learning algorithms. The framework firstly demonstrates a new approach of water quantity vulnerability assessment based on reliable and updated spatial-temporal datasets (soil water stress, aridity index, rain use efficiency and leaf area index), and the incorporation of the GIS-based model. Notably, this research devises a state-of-the-art machine-learning model for monitoring agricultural drought via predicting soil moisture (SM) using active and ensemble-based decision tree learning combined with multi-sensor data fusion at a national and world scale. This work explores the use of Sentinel-1, Sentinel-2, and an innovative machine learning (ML) approach using an integration of active learning for land-use mapping and advanced Extreme Gradient Boosting Regression (XGBR) for robustness of the SM estimates. The collected soil samples from a field survey in Western Australia were also used for the model validation and indicators including the coefficient of determination (R^2) and root - mean – square - error (RMSE) were applied to evaluate the model's performance. The proposed model XGBR with 21 optimal features obtained from GA was yielded the highest performance (R^2 = 0.891, RMSE = 0.875%). A combination of S1 and S2 sensors could also effectively estimate SOC in farming areas by using ML techniques. Satisfactory accuracy of the proposed XGBoost with optimal features was achieved the highest performance (R^2 = 0.870; RMSE = 1.818 tonC/ha) which outperformed random forest and support vector machine

Conclusively, the described conceptual model can further support precision agriculture, water management and drought resilience programs via water use efficiency, green infrastructure and smart irrigation management for agricultural production.

Keywords: Water vulnerability, spatial datasets, correlation analysis, GIS-based model, machine learning, soil moisture, soil organic carbon, feature selection, genetic algorithm