UNIVERSITY OF CALIFORNIA
RIVERSIDE

Advancing Urban Landscape Irrigation Management using Smart Controllers and Machine Learning-Based Models

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Environmental Sciences

by

Amninder Singh

March 2022

Dissertation Committee:
Dr. Amir Haghverdi, Chairperson
Dr. Laosheng Wu
Dr. Hoori Ajami

The Dissertation of Amninder Singh is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGMENTS

I would like to begin by thanking all who have been a part of my life and supported me during these four years at the University of California, Riverside. First and foremost, I would like to express my sincere gratitude to my Ph.D. advisor Dr. Amir Haghverdi for his continuous support, encouragement, and confidence in me. Over the years, he has helped me develop my skills as a critical thinker and dedicated his time and effort to improve my research. Additionally, I would like to acknowledge my committee members Dr. Laosheng Wu and Dr. Hoori Ajami, for serving as members of my dissertation committee. Your meaningful insights and continuous guidance throughout the research have helped me tremendously.

The support of Dr. Darren Haver from SCREC, Irvine, who assisted with collected weekly data from the research plots, is greatly appreciated. Additionally, thanks are necessary for Dr. Wolfgang Durner and Dr. Hasan Sabri Öztürk, the co-authors for the manuscripts published from chapter 1 and 2 of this dissertation, for sharing their expert advice. I would also like to thank Haghverdi lab members Anish Sapkota, Akanksha Garg, Somayeh Ghodsi, Jean Claude, and Zahra Amiri for their friendship, stimulating discussions, and help with collecting field data as well as lab analysis. Last but not least, my greatest gratitude goes to my family. This accomplishment would not have been possible without them.

Funding for this project was provided by National Water Research Institute (NWRI) and UC Agricultural and Natural Resources.

ABSTRACT OF THE DISSERTATION

Advancing Urban Landscape Irrigation Management using Smart Controllers and Machine Learning-Based Models

by

Amninder Singh

Doctor of Philosophy, Graduate Program in Environmental Sciences
University of California, Riverside, March 2022
Dr. Amir Haghverdi, Chairperson

Efficient urban landscape irrigation management is critical in California and depends on the reliable estimation of soil hydraulic properties and reference evapotranspiration (ET_o). Direct measurements of these properties are time-consuming, challenging, and often expensive. Therefore, it is desirable to estimate indirectly using readily available data. The ET-based smart irrigation controllers used for landscape irrigation often rely on temperature-based ET_o models; thus, a comprehensive evaluation of these models across climate regions is required in California. Furthermore, studies are needed to evaluate the response of turfgrass to soil moisture sensor (SMS) based deficit irrigation treatments and assess the efficacy of smart controller for autonomous irrigation scheduling in semi-arid conditions of California.

This dissertation addresses the first two challenges by developing artificial neural network (ANN) based models for accurate estimations of soil hydraulic properties and ET_o.

For the first time, we utilized an international high resolution dataset measure by evaporation methods using HYPROP^{TM} (Hydraulic Property Analyzer, Meter Group Inc., USA) to develop the pseudo-continuous neural network PTF (PC_{NN}-PTF) models. We assessed the accuracy and reliability of the PC_{NN}-PTF approach for estimating the soil water retention curve (SWRC) and soil hydraulic conductivity curve (SHCC) The best performing PC_{NN}-PTF shoed root mean square error (RMSE) of 0.043 cm^3 cm^{-3} for SWRC, and RMSE of 0.520 for SHCC estimation. The subsequent study evaluated eight temperature-based empirical ET_o models and four ANN models for ET_o estimation in California. A total of 101 active California Irrigation Management Information System (CIMIS) weather stations were selected for this study, with more than 725,000 data points expanding from 1985 to 2019. The ANN model outperformed the widely used Hargreaves and Samani (HS_a) model using the same input variables (i.e., air temperature and extraterrestrial solar radiation) with 11% lesser RMSE.

Lastly, a three-year (2019–20121) irrigation research trial was conducted to evaluate the response of bermudagrass to soil moisture sensor (SMS) based deficit irrigation treatments and assess the smart controller for autonomous irrigation scheduling using recycled water. By the end of the research period, turfgrass quality was below the acceptable NDVI of 0.5, suggesting that bermudagrass generally does not perform well when deficit-irrigated with recycled water in a long-term basis in semi-arid climate. Further investigation is needed to substantiate SMS-based autonomous deficit irrigation scheduling when recycled water is used.

Table of Contents

Chapter 1. Introduction 1

1.1 Motivation and Background 1

Estimating Soil Hydraulic Properties using Pseudo-Continuous Pedotransfer Functions (PC-PTFs) 4

Reference Evapotranspiration (ET_o) 5

Soil Moisture sensor based Irrigation Scheduling using recycled water 6

1.2. Research Objectives 7

1.3References 8

Chapter 2. Developing Pseudo Continuous Pedotransfer Functions for International Soils Measured with the Evaporation Method and the HYPROP System: I. The Soil Water Retention Curve 12

Abstract 12

2.1 Introduction 12

2.2 Materials and Methods 16

2.2.1 Soil Data Sets 16

2.2.2. ANN PC-PTFs Development 19

2.2.3. Modeling Scenarios 21

2.2.4. Model Evaluation 22

2.2.5. Domain of the Pedotransfer Functions 23

2.3. Results 24

2.3.1. Importance of the Input Predictors 24

2.3.2. Performance across Soil Textures 27

2.3.3. Performance at the Wet, Intermediate and Dry Parts of the SWRC 29

2.4. Discussion 30

2.4.1. Accuracy and Reliability of the Developed PTFs 30

2.4.2. Importance of Input Variables 35

2.4.3. Performance across Textural Classes and Tension Ranges 36

2.5. Conclusions **38**
2.5. References **39**
Chapter 3. Developing Pseudo Continuous Pedotransfer Functions for International Soils Measured with the Evaporation Method and the HYPROP System: II. The Soil Hydraulic Conductivity Curve **44**
Abstract: **44**
3.1. Introduction **45**
3.2. Materials and Methods **48**
3.2.1. Soil Data Sets 48
3.2.2. Unsaturated Hydraulic Conductivity Calculations 51
3.2.3. PC_{NN}-PTFs Development 53
3.2.4. Modeling Scenarios 53
3.2.5. Model Evaluation 55
3.3. Results **56**
3.3.1. Importance of the Input Predictors 56
3.3.2. Performance across Soil Textures 60
3.3.3. Performance at the Wet, Intermediate and Dry Parts of the SHCC 62
3.4. Discussion **64**
3.4.1. Accuracy and Reliability of the Developed PTFs 64
3.4.2. Importance of Input Variables 66
3.4.3. Performance across Textural Classes and Tension Ranges 67
3.5. Conclusions **68**
3.6. References **69**
Chapter 4. Evaluating the long term performance of empirical temperature-based and artificial neural network models for estimating reference evapotranspiration in California… **73**
Abstract **73**
4.1. Introduction **74**

4.2. Material and methods **76**
4.2.1. Study Region and Data Sources 76
4.2.2. Temperature-Based Empirical ET_o models 79
4.2.3. Artificial Neural Network Models 82
4.2.4. Performance Assessment 84
4.3. Results & discussion **85**
4.3.1 Overall performance of the ET_o models 85
4.3.2 Importance of the Input Parameters 89
4.3.3 Temporal Analysis of the ET_o models 90
4.3.4 Spatial Analysis of the ET_o models 97
4.5. Conclusion **102**
4.6. References **103**
Chapter 5. Autonomous Hybrid Bermudagrass Recycled Water Irrigation Management Using a Smart Soil Moisture Sensor-Based Controller **108**
Abstract **108**
5.1. Introduction **109**
5.2. Material and methods **110**
5.2.1 Study site 110
5.2.2. Irrigation system and instrumentation 111
5.2.3 Irrigation trial and treatment design 113
5.2.4 Handheld remote sensing data 114
5.2.5 Soil sampling and Infiltration data collection 115
5.2.6 Statistical analysis 117
5.3. Results **117**
5.3.1. Irrigation applications 117
5.3.2. NDVI and turf temperatures 125
5.3.3. Soil salinity and Infiltration 131
5.4. Discussion **134**
5.4.1. Autonomous SMS based irrigation 134
5.4.2. Response of bermudagrass to deficit irrigation with recycled water 136

5.4.3. Soil salinity, SAR, and infiltration rate under deficit recycled water irrigation 137
5.5. Conclusion 139
5.6. References 140
Chapter 6. Summary and Conclusion 144

List of Figures

Figure 1.1. Illustration of soil water balance in a semi-arid urban landscape assuming no runoff and deep percolation. 3

Figure 2.1. Number and origin of the undisturbed soil core samples for the international data set used in this study to develop pedotransfer functions. 17

Figure 2.2. Soil water retention data pairs (**a**), and soil textural distribution for the data sets used in this study (**b**). Red points depict the international data set from evaporation experiments (Schindler and Müller, 2017) and blue points represent the Turkish data set (Haghverdi et al., 2020a, 2018, 2014). 19

Figure 2.3. Development workflow of the pseudo continuous neural network pedotransfer functions (PC_{NN}-PTFs) for the soil water retention curve (SWRC) estimations. 21

Figure 2.4. Scatterplots of the measured versus estimated volumetric water content (VWC) via PC_{NN}-PTFs when the international data set was used to train and test the models (top), and for the Turkish soil samples when Turkish data set was not used for training (middle) and when Turkish data set was incorporated into the training data set (bottom). 25

Figure 2.5. The domain of the developed PC_{NN}-PTFs using Mahalanobis distance, indicating that the two data sets were independent with a slight overlap since only 8 Turkish soil samples (highlighted in red) fell below the cut-off limit (y-axis for Turkish data set is on an exponential scale). 34

Figure 2.6. Relationship between the number of data points for each textural class and the accuracy of the best performing PC_{NN}-PTF (Model 3 with SSC, BD, and pF as inputs) when both international and Turkish data sets were used to develop the models. 37

Figure 3.1. Experimental setup of the extended evaporation experiment using HYPROP system. 48

Figure 3.2. The soil hydraulic conductivity and tension pairs (**a**), and soil textural distribution for the datasets (**b**). Dark orange circles depict the international data set (Schindler and Müller, 2017) and blue circles represent the Turkish dataset [21,28,30]. pF is the logarithmic transformation of soil tension in cm of water and *K* is the unsaturated hydraulic conductivity. 50

Figure 3.3. Scatterplots of measured versus estimated log(*K*) using PC_{NN}-PTFs. S1: training and test: the international dataset, S2: Training and test: Turkish dataset, S3: training: the international dataset, test: Turkish dataset, S4: training: international + Turkish dataset, test: Turkish dataset. Model 1 inputs: sand, silt and clay percentages (SSC), bulk density (BD), and soil organic matter content (SOM); Model 2 inputs: SSC; Model 3 inputs: SSC, BD; Model 4 inputs: SSC and SOM. 57

Figure 3.4. The root of mean squared error (RMSE) as a function of bulk density (BD) and organic matter content (SOM) for the PC_{NN}PTF Model 1 with SSC, BD, SOM as inputs. The model was developed using combined international and Turkish data sets and tested

using the Turkish data set (scenario 4). The error was calculated for each soil sample separately. 59

Figure 3.5. The root of mean squared error (RMSE) as a function of the number of measured hydraulic conductivity data pairs for each textural class for the PC_{NN}-PTF Model 1 with SSC, BD, SOM as inputs. The model was developed using combined international and Turkish data sets. 67

Figure 4.1. Distribution of the CIMIS stations evaluated in this study across the state of California. The aridity index values were obtained from the CGIAR-CSI Global-Aridity Database, and the classes were mapped based on the recommendations by United Nations Environmental Programme (UNEP). 77

Figure 4.2. Histograms of the meteorological variables (top) and availability of the data (bottom) across CIMIS stations used in this study. 78

Figure 4.3. The architecture of the feed-forward backpropagation neural network models that were developed in this study. T_{min}, T_{max}, and T_a are the daily minimum, maximum, and mean air temperature [°C], respectively; RH: Relative Humidity [%]; U: daily mean windspeed [m/s]; R_a: extraterrestrial solar radiation [mm/d]. 83

Figure 4.4. Scatterplots of the CIMIS ETo versus estimated daily ETo by the 12 temperature-based empirical and ANN models evaluated in this study. The dashed orange line is the 1:1 line. 87

Figure 4.5. Comparison between long-term year-round ET_o obtained from CIMIS against the 12 temperature-based empirical and ANN ET_o models evaluated in this study. The solid lines show mean ET_o and the shaded bands depict the standard deviation of ET_o across all CIMIS stations. DOY: day of the year. 91

Figure 4.6. Variation in annual mean absolute error, MAE [mm d^{-1}] values against meteorological variables - wind speed, relative humidity (RH), vapor pressure deficit (VPD), and global solar radiation (R_s). 96

Figure 4.7. Variation in annual mean bias error, MBE [mm d^{-1}] values against meteorological variables - wind speed, relative humidity (RH), vapor pressure deficit (VPD), and global solar radiation (R_s). 97

Figure 4.8. Comparison between the long-term ET_o maps obtained from CIMIS against the estimated maps by the 12 temperature-based empirical and ANN models evaluated in this study. 102

Figure 5.1. Daily reference evapotranspiration (ET_o) and precipitation during the experimental period obtained from the CIMIS station #75. Shaded green region represents the period when NDVI and turf temperature measurements were taken during each irrigation season. 111

Figure 5.2. Soil moisture and irrigation runtime for restricted (3d/week) and on-demand (7 d/week) irrigation treatments implemented for 2019. Solid red and dark blue lines are the lower (LL) and upper (UL) soil moisture thresholds levels, respectively. 120

Figure 5.3. Soil moisture and irrigation runtime for restricted (3d/week) and on-demand (7 d/week) irrigation treatments implemented for 2020. Solid red and dark blue lines are the lower (LL) and upper (UL) soil moisture thresholds levels, respectively. 122

Figure 5.4. Soil moisture and irrigation runtime for restricted (3d/week) and on-demand (7 d/week) irrigation treatments implemented for 2021. Solid red and dark blue lines are the lower (LL) and upper (UL) soil moisture thresholds levels, respectively. 124

Figure 5.5. Changes in normalized difference vegetation index (NDVI) values over time across the irrigation treatments for the restricted (3 d/week) and on-demand (7 day/week) irrigation treatments imposed in 2019 (top), 2020 (middle), and 2021 (bottom). 129

Figure 5.6. Changes in turfgrass temperature values over time across the irrigation treatments for the restricted (3 d/week) and on-demand (7 day/week) irrigation treatments imposed in 2019 (top), 2020 (middle), and 2021 (bottom)..... 130

Figure 5.7. Soil salinity (EC_e) distribution in the soil profile from soil samples collected before (Spring) and after (Fall) of the summer irrigation seasons for 2019, 2020, and 2021. 133

List of Tables

Table 2.1. Characteristics of soils from both international and Turkish data sets used in this study to develop and test pseudo continuous neural network pedotransfer functions (PC_{NN}-PTFs). 19

Table 2.2. Combinations of input attributes (scenarios) that were used in this study to develop the pseudo continuous neural network pedotransfer functions (PC_{NN}-PTFs). 22

Table 2.3. Comparison between the performance of the PC_{NN}-PTFs trained using different data sets to estimate the volumetric water content (cm^3 cm^{-3}) of the international and Turkish soil samples. 26

Table 2.4. Soil texture-based performance of the PC_{NN}-PTFs (inputs: SSC, BD, OM, pF) developed and tested using the international data set to estimate the volumetric water content (cm^3 cm^{-3}). 28

Table 2.5. Soil texture based performance of the PC_{NN}-PTFs (Model 1 with SSC, BD, SOM, and pF as inputs) developed using the international data set and the international plus Turkish data sets to estimate the volumetric water content (cm^3 cm^{-3}) of the Turkish soil samples. 28

Table 2.6. Performance of the PC_{NN}-PTFs (inputs: SSC, BD, OM, and pF as) developed using the international data set and the international plus Turkish data sets to estimate the volumetric water content (cm^3 cm^{-3}) at wet ($pF \leq 2$) intermediate ($2 < pF \leq 3$) and dry ($pF > 3$) parts of the SWRC. 30

Table 2.7. Comparison of the pseudo-continuous pedotransfer functions (PC-PTFs) developed in the literature to the PC_{NN}-PTF developed in this study. 32

Table 3.1. Characteristics of soils from international and Turkish data sets used in this study to develop and test pseudo continuous neural network pedotransfer functions (PC_{NN}-PTFs). 51

Table 3.2. Combinations of input attributes used in this study to develop PC_{NN}-PTFs... 54

Table 3.3. Different data partitioning scenarios used in the study to train, test, and validate PC_{NN}-PTFs. 55

Table 3.4. Performance of the PC_{NN}-PTFs estimating log-transformed soil hydraulic conductivity data (cm d−1) across four modeling scenarios 58

Table 3.5. Performance of PC-PTFs on main textural classes of the international and Turkish data sets for estimating log(K). 62

Table 3.6. Performance of the PC_{NN}-PTFs (inputs: SSC, BD, OM, and pF) developed to estimate the log(K) at wet ($pF \leq 2$) intermediate ($2 < pF \leq 3$) and dry ($pF > 3$) parts of the SHCC. 63

Table 4.1. Overall performance of the temperature-based ET_o models compared to CIMIS ET_o. 88

Table 4.2. Monthly root mean square error (RMSE) values for the temperature-based ET_o equations evaluated in this study against CIMIS ET_o..... 92

Table 4.3. Monthly mean absolute error (MAE) values for the temperature-based ET_o equations evaluated in this study against CIMIS ET_o..... 93

Table 4.4. Monthly mean bias error (MBE) values for the temperature-based ET_o equations evaluated in this study against CIMIS ET_o. 94

Table 4.5. Climate-division-specific calibration equations and performance statistics for the temperature-based ETo equations evaluated in this study against CIMIS ET_o. 99

Table 5.1. Treatments imposed with the Soil moisture sensors in the instrumented plots. 114

Table 5.2. Percent of ET_o applied to each treatment during the experimental period. .. 118

Table 5.3. Statistical analysis of the bermudagrass response in terms of normalized difference vegetation index (NDVI) and turf temperature to irrigation treatments imposed in years 2019, 2020 and 2021 (each year was analyzed separately). 127

Table 5.4. Statistical analysis of the soil samples collected before (Spring) and after (Fall) the irrigation season in 2019, 2020 and 2021. 134

Table 5.5. Mean infiltration rate (cm/s) for the entire study period measured by SATURO before (Spring) and after (Fall) the irrigation season..... 134

Chapter 1. Introduction

1.1 Motivation and Background

Irrigation demand is a significant component of total water use in the urban sector in California. About 50% of water dedicated to urban water use goes toward landscape irrigation (California Water Plan Update, 2013), which could be up to 90% (Cooley and Gleick, 2009) during summer months in southwestern US. This is due to the high evapotranspiration (ET) demand and low precipitation in summer not sufficient to provide landscapes with enough water needed to achieve the level of aesthetic appearance that is desired by home and business owners. However, urban landscape irrigation is important for maintaining urban landscapes and is considered valuable in planning processes in many urban areas.

Reducing the urban water use has gained increased attention considering the limiting freshwater resources, projected increase in extreme wet and dry seasons (Swain et al., 2018), urbanization (Alig et al., 2004), and semi-arid climate of the state. Water supplies for Southern California are projected to diminish as a result of more extreme hydrological events in southwestern US (Pagán et al., 2016). The projected rise in population of Southern California from what is already the highest in the state ("CSDO," 2020), calls for new approaches to outdoor urban water conservation that can alleviate pressure on limited water resources. Turfgrass, considered the largest irrigated crop in the US (Milesi et al., 2005), forms a major component of the landscapes planted in recreational fields, golf courses, and public parks. A study by (Litvak et al., 2017) done in city of Los Angeles showed that the irrigated turfgrass contributed for 64–84% of total urban landscape ET. On

the other hand, turfgrasses also have benefits such soil carbon (C) sequestration (Milesi et al., 2005), mitigation of the urban heat island effect (Shashua-Bar et al., 2009), and enhanced infiltration (Harivandi et al., 2009). To alleviate pressure on the limiting freshwater resources, studies are needed to develop efficient irrigation practices for nontraditional water resources such as recycled water to understand how irrigation management with these resources impacts landscape quality and soil health (Assouline et al., 2015).

Advancements in irrigation technologies have led to opportunities in affordable and efficient irrigation scheduling by adoption of smart controllers. Homeowners are likely to prefer smart irrigation controllers to conventional automated systems because of potential savings on annual water bills (Khachatryan et al., 2019). Irrigation-scheduling requires calculating soil water balance through a comprehensive system of soil and crop monitoring system, to make real time irrigation management decisions. The objective is to determine the precise amount and timing of this irrigation to replace soil moisture to a desired level while maintaining optimum yield or quality of the crop (Dane et al., 2006). Fulfilling the ET demand using soil moisture sensors (SMSs) requires selection of a lower threshold value to which the soil is allowed to dry before the next irrigation event, typically called management allowed depletion (MAD), and an upper threshold (field capacity) beyond which no irrigation is required, as depicted in Figure 1.1. Smart irrigation controllers (SICs) receive feedback from an onsite weather (ET) or SMS, and thus can be broadly classified into two categories – SMS-based and ET-based SICs (Dukes, 2012). Since SICs make irrigation decisions based on the real-time measurements of soil water with an SMS and/or

by calculating a soil water balance with feedback from weather sensor(s), accurate information about soil hydraulic properties and reference evapotranspiration (ET_o) is essential. Accurate estimation of soil hydraulic properties remains a critical component of efficient irrigation management in urban settings mainly because direct measurement of these properties is time-consuming and expensive. Furthermore, new developments in machine-learning techniques hold great potential to unlock the value of big data for irrigated agriculture (Adeyemi et al., 2017). Research is needed in semi-arid regions towards use of smart controllers and their potential to save freshwater resources.

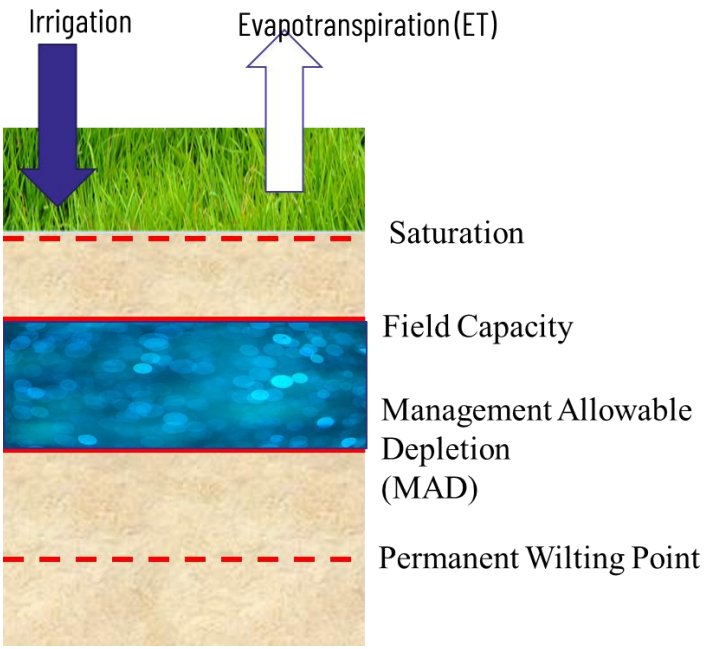


Figure 1.1. Illustration of soil water balance in a semi-arid urban landscape assuming no runoff and deep percolation.

Estimating Soil Hydraulic Properties using Pseudo-Continuous Pedotransfer Functions (PC-PTFs)

The primary soil hydraulic properties include the soil water retention and hydraulic conductivity curves (SWRC and SHCC) that define the volumetric water content's nonlinear relationships with the soil tension and the soil hydraulic conductivity, respectively. Direct measurements of soil hydraulic properties in the field and laboratory can be tedious, laborious, and often expensive due to their significant inherent spatial variability. Therefore, pedotransfer functions (PTFs) are often developed and used to indirectly estimate these properties by establishing empirical relationships based on the readily available soil properties such as soil texture, bulk density (BD), and soil organic matter content (SOM) (Bouma, 1989). Pedotransfer functions (PTFs) are statistical tools used in soil science to estimate soil hydraulic properties, mainly the soil water retention curve (SWRC), based on the easily collected basic soil properties, available from most regional and national databases (Vereecken et al., 2010). The pseudo-continuous neural network PTF (PC_{NN}-PTF) (Haghverdi et al., 2012) was introduced as an alternative approach for continuous estimation of the SWRC at any desired water retention. PC_{NN}-PTF utilizes machine learning techniques to estimate the shape of the SWRC based on actual measured data points, unlike parametric PTFs, where the curvature is dictated by the selected soil hydraulic equation.

Schindler and Müller (2017) published a soil hydraulic international dataset using the Evaporation method and HYPROP^{TM} (Hydraulic Property Analyzer, Meter Group Inc., USA) system, and is becoming the standard approach of measuring soil hydraulic

properties in the laboratory. No PC_{NN}-PTF has been developed to estimate the SWRC and SHCC using high-resolution international data measured in this dataset. Using an international dataset to develop PTFs have potential to provide generalized estimations in regions of the world where availability of soil hydraulic datasets is limited.

Reference Evapotranspiration (ET_o)

Evapotranspiration (ET) is well recognized as the demand for water that different crops require for optimal development and is the sum of vaporization by the combined processes of evaporation and transpiration. Reference ET (ET_o) is the rate at which soil water is converted to water vapor for a defined reference crop (generally cool-season grass or alfalfa) under given atmospheric conditions. The FAO-56 Penman-Monteith is the accepted standard method to estimate the reference evapotranspiration, ET_o (Allen et al., 2005). However, it requires a wide range of accurate meteorological data such as temperature, humidity, solar radiation, and wind speed. Therefore, its implementation in data scare situations such as landscape irrigation management in urban areas is challenging and limited.

ET-based controllers saw an average of 20 percent irrigation reduction compared to homes with homeowner-scheduled irrigation (Devitt et al., 2008) in Las Vegas, Nevada, and resulted in 24 percent less water application compared to a typical residential landscape irrigated with a timer-based controller in Southern California (Bijoor et al., 2014). Additionally, a study conducted on St. Augustine turfgrass showed an average irrigation savings of 43 percent in the summer compared to homeowner-scheduled irrigation, with no reduction in turfgrass quality ((Davis et al., 2009). Similarly, other studies have shown

the water conservation potential of the ET based smart irrigation controllers (Davis & Dukes, 2010; Davis et al., 2009). However, water savings were found to be less pronounced under sustained dry conditions (Cardenas-Lailhacar et al., 2010; Grabow et al., 2012) when using ET based irrigation scheduling. The ET-based smart irrigation controllers used for landscape irrigation often rely on temperature-based ET_o models; thus, a comprehensive evaluation of these models across climate regions is required in California. Furthermore, the ability of ANN temperature based models to provide generalized estimations of ET_o in semi-arid regions needs to be evaluated.

Soil Moisture sensor based Irrigation Scheduling using recycled water

Using recycled water for irrigating urban landscapes is a promising solution to alleviate the pressure on freshwater resources compared to other alternatives such as seawater desalination and imported surface waters. Opportunities for increasing the use of recycled water to irrigate urban landscapes exist as the proportion of total recycled urban water use in southern California varied from 26% in Santa Ana, 38% for Los Angeles to 94% for the San Diego region in 2015 (SWRCB, 2015).

SMS based smart irrigation controllers have shown significant water saving potential (Blonquist et al., 2006; Cardenas-Lailhacar and Dukes, 2012; Cardenas et al., 2021; Qualls et al., 2001). For example, under relatively dry conditions, SMS-based irrigation controllers reduced irrigation by 11-53 percent compared to a time-based schedule (McCready et al., 2009). However, previous studies have mostly focused on the implementation of smart landscape irrigation technologies that used potable water. High concentrations of the salt present in recycled water could negatively impact plant growth

and soil health (Gonçalves et al., 2007; Qian and Mecham, 2005). This problem is particularly important in arid and semi-arid climates of Southern California, where low precipitation in some years may not adequately leach soluble salts from the root zone. Being a conservative constituent, salts will accumulate over time in the crop root zone if water supply is insufficient to provide leaching of salts (Ayers and Westcot, 1985).

1.2. Research Objectives

This research has several major objectives: (I) to develop PC_{NN}-PTFs for SWRC and SHCC estimations by utilizing the abovementioned international (Schindler and Müller, 2017) and Turkish (Haghverdi et al., 2018) data sets measured via HYPROP. The HYPROP system works based on the extended evaporation method (Schindler et al., 2010a, 2010b) it has several advantages over the traditional equilibrium methods (i.e., pressure plate extractors and sandbox apparatus). The accuracy and reliability of the PC_{NN}-PTFs, developed with four combinations of the input attributes including soil texture (i.e., percentages of sand, silt, and clay; SSC), BD, and SOM, was also evaluated. This is discussed in more detail in chapters 2 and 3.

(II) In Chapter 4, we used long-term daily CIMIS (California Irrigation Management Information System) ET_o data across the state of California to evaluate the performance of eight temperature-based empirical ET_o models, to develop and evaluate the accuracy of artificial neural network (ANN) based ET_o models using different sets of input data.

(III) In chapter 5, a turfgrass plot study was conducted in Orange County, CA to determine the impact of different deficit irrigation strategies with recycled water on hybrid bermudagrass quality under in, and on soil salinity (EC_e), sodium adsorption ratio (SAR),

and soil infiltration rate. While efforts are needed to conserve water, homeowners still value high-quality turf and landscapes. Therefore, it is important that effectiveness in maintaining turf quality be included in any evaluation of the water conservation potential of smart irrigation controllers. Performance of an SMS-based smart irrigation controller for efficient automatic recycled water irrigation management while maintaining an optimum quality of hybrid bermudagrass was evaluated in this chapter.

1.3References

Adeyemi, O., Grove, I., Peets, S., Norton, T., 2017. Advanced monitoring and management systems for improving sustainability in precision irrigation. Sustain. 9, 1–29. https://doi.org/10.3390/su9030353

Alig, R.J., Kline, J.D., Lichtenstein, M., 2004. Urbanization on the US landscape: Looking ahead in the 21st century. Landsc. Urban Plan. 69, 219–234. https://doi.org/10.1016/j.landurbplan.2003.07.004

Allen, R.G., Walter, I.A., Elliott, R.L., Howell, T.A., Itenfisu, D., Jensen, M.E., Snyder, R.L. (Eds.), 2005. The ASCE Standardized Reference Evapotranspiration Equation. Technical Committee on Standardization of Reference Evapotranspiration, Books. American Society of Civil Engineers, Reston, VA. https://doi.org/10.1061/9780784408056

Assouline, S., Russo, D., Silber, A., Or, D., 2015. Balancing water scarcity and quality for sustainable irrigated agriculture. Water Resour. Res. 51, 3419–3436. https://doi.org/10.1002/2015WR017071

Ayers, R.S., Westcot, D.W., 1985. Water Quality for Agriculture. FAO UNITED NATIONS,Rome,italy 97.

Bijoor, N.S., Pataki, D.E., Haver, D., Famiglietti, J.S., 2014. A comparative study of the water budgets of lawns under three management scenarios. Urban Ecosyst. 17, 1095–1117. https://doi.org/10.1007/s11252-014-0361-4

Blonquist, J.M., Jones, S.B., Robinson, D.A., 2006. Precise irrigation scheduling for turfgrass using a subsurface electromagnetic soil moisture sensor. Agric. Water Manag. 84, 153–165. https://doi.org/10.1016/j.agwat.2006.01.014

Bouma, J., 1989. Using Soil Survey Data for Quantitative Land Evaluation., in: Stewart, B.. (Ed.), Advances in Soil Science, Vol 9. Springer, New York, NY. https://doi.org/https://doi.org/10.1007/978-1-4612-3532-3

California Water Plan Update, 2013. Volume 3 Resource Management Strategies, DWR.

Cardenas-Lailhacar, B., Dukes, M.D., 2012. Soil Moisture Sensor Landscape Irrigation Controllers: A Review of Multi- Study Results and Future Implications. Am. Soc. Agric. Biol. Eng. 55, 581–590.

Cardenas-Lailhacar, B., Dukes, M.D., Miller, G.L., 2008. Sensor-Based Automation of Irrigation on Bermudagrass, during Wet Weather Conditions. J. Irrig. Drain. Eng. 134, 120–128. https://doi.org/10.1061/(ASCE)0733-9437(2008)134:2(120)

Cardenas, B., Dukes, M.D., Breder, E., Torbert, J.W., 2021. Long-term performance of smart irrigation controllers on single-family homes with excess irrigation . AWWA Water Sci. 3, 1–14. https://doi.org/10.1002/aws2.1218

Cooley, H., Gleick, P.H., 2009. Urban water-use efficiencies: Lessons from United States cities. Island Press: Washington, DC, USA.

CSDO [WWW Document], 2020. . Popul. Proj. Methodol. (2019 Baseline). URL https://www.dof.ca.gov/forecasting/demographics/projections/ (accessed 11.29.21).

Dane, J.H., Walker, R.H., Bahaminyakamwe, L., Belcher, J.L., 2006. Tall fescue and hybrid bluegrass response to soil water matric head limits. Agric. Water Manag. 86, 177–186. https://doi.org/10.1016/j.agwat.2006.07.001

Davis, S.L., Dukes, M.D., 2010. Irrigation scheduling performance by evapotranspiration-based controllers. Agric. Water Manag. 98, 19–28. https://doi.org/10.1016/j.agwat.2010.07.006

Davis, S.L., Dukes, M.D., Miller, G.L., 2009. Landscape irrigation by evapotranspiration-based irrigation controllers under dry conditions in Southwest Florida. Agric. Water Manag. 96, 1828–1836. https://doi.org/10.1016/j.agwat.2009.08.005

Devitt, D.A., Carstensen, K., Morris, R.L., 2008. Residential Water Savings Associated with Satellite-Based ET Irrigation Controllers. J. Irrig. Drain. Eng. 134, 74–82.

Dukes, M.D., 2012. Water conservation potential of landscape irrigation smart controllers. Trans. ASABE 55, 563–569.

Gonçalves, R.A.B., Folegatti, M. V., Gloaguen, T. V., Libardi, P.L., Montes, C.R., Lucas, Y., Dias, C.T.S., Melfi, A.J., 2007. Hydraulic conductivity of a soil irrigated with treated sewage effluent. Geoderma 139, 241–248. https://doi.org/10.1016/j.geoderma.2007.01.021

Grabow, G.L., Ghali, I.E., Huffman, R.L., Miller, G.L., Bowman, D., Vasanth, A., 2013. Water application efficiency and adequacy of ET-based and soil moisture-based irrigation controllers for turfgrass irrigation. J. Irrig. Drain. Eng. 139, 113–123. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000528

Haghverdi, A., Cornelis, W.M., Ghahraman, B., 2012. A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. J. Hydrol. 442–443, 46–54. https://doi.org/10.1016/j.jhydrol.2012.03.036

Haghverdi, A., Öztürk, H.S., Durner, W., 2018. Measurement and estimation of the soil

water retention curve using the evaporation method and the pseudo continuous pedotransfer function. J. Hydrol. 563, 251–259. https://doi.org/10.1016/j.jhydrol.2018.06.007

Harivandi, M.A., Baird, J., Hartin, J., Henry, M., Shaw, D., 2009. Managing Turfgrasses during Drought. UC ANR Publ. 8395 1–9.

Khachatryan, H., Suh, D.H., Xu, W., Useche, P., Dukes, M.D., 2019. Towards sustainable water management: Preferences and willingness to pay for smart landscape irrigation technologies. Land use policy 85, 33–41. https://doi.org/10.1016/j.landusepol.2019.03.014

Litvak, E., Manago, K.F., Hogue, T.S., Pataki, D.E., 2017. Evapotranspiration of urban landscapes in Los Angeles, California at the municipal scale. Water Resour. Res. 53, 4236–4252. https://doi.org/10.1002/2016WR020254

McCready, M.S., Dukes, M.D., Miller, G.L., 2009. Water conservation potential of smart irrigation controllers on St. Augustinegrass. Agric. Water Manag. 96, 1623–1632. https://doi.org/10.1016/j.agwat.2009.06.007

Milesi, C., Running, S.W., Elvidge, C.D., Dietz, J.B., Tuttle, B.T., Nemani, R.R., 2005. Mapping and modeling the biogeochemical cycling of turf grasses in the United States. Environ. Manage. 36, 426–438. https://doi.org/10.1007/s00267-004-0316-2

Pagán, B.R., Ashfaq, M., Rastogi, D., Kendall, D.R., Kao, S.C., Naz, B.S., Mei, R., Pal, J.S., 2016. Extreme hydrological changes in the southwestern US drive reductions in water supply to Southern California by mid century. Environ. Res. Lett. 11. https://doi.org/10.1088/1748-9326/11/9/094026

Qian, Y.L., Mecham, B., 2005. Long-term effects of recycled wastewater irrigation on soil chemical properties on golf course fairways. Agron. J. 97, 717–721. https://doi.org/10.2134/agronj2004.0140

Qualls, R.J., Scott, J.M., DeOreo, W.B., 2001. Soil moisture sensors for urban landscape irrigation: Effectiveness and reliability. J. Am. Water Resour. Assoc. 37, 547–559. https://doi.org/10.1111/j.1752-1688.2001.tb05492.x

Rhoades, J.D., Bingham, F.T., Letey, J., Hoffman, G.J., Dedrick, A.R., Pinter, P.J., Replogle, J.A., 1989. Use of saline drainage water for irrigation: Imperial Valley study. Agric. Water Manag. 16, 25–36. https://doi.org/10.1016/0378-3774(89)90038-3

Schindler, U., Durner, W., von Unold, G., Mueller, L., Wieland, R., 2010a. The evaporation method: Extending the measurement range of soil hydraulic properties using the air-entry pressure of the ceramic cup. J. Plant Nutr. Soil Sci. 173, 563–572. https://doi.org/10.1002/jpln.200900201

Schindler, U., Durner, W., von Unold, G., Müller, L., 2010b. Evaporation Method for Measuring Unsaturated Hydraulic Properties of Soils: Extending the Measurement Range. Soil Sci. Soc. Am. J. 74, 1071. https://doi.org/10.2136/sssaj2008.0358

Schindler, U., Müller, L., 2017. Soil hydraulic functions of international soils measured with the Extended Evaporation Method (EEM) and the HYPROP device. Open Data J. Agric. Res. 3, 10–16.

Shashua-Bar, L., Pearlmutter, D., Erell, E., 2009. The cooling efficiency of urban landscape strategies in a hot dry climate. Landsc. Urban Plan. 92, 179–186. https://doi.org/10.1016/J.LANDURBPLAN.2009.04.005

Swain, D.L., Langenbrunner, B., Neelin, J.D., Hall, A., 2018. Increasing precipitation volatility in 21st century California. Nat. Clim. Chang. 8. https://doi.org/10.1038/s41558-018-0140-y

SWRCB, California State Water Resources Control Board. [WWW Document], 2015. . Munic. wastewater Recycl. Surv. URL https://www.waterboards.ca.gov/water_issues/programs/grants_loans/water_recycling/munirec.shtml (accessed 10.31.18).

Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M.G., Genuchten, M.T. van, 2010. Using Pedotransfer Functions to Estimate the van Genuchten–Mualem Soil Hydraulic Properties: A Review. Vadose Zo. J. 9, 795. https://doi.org/10.2136/vzj2010.0045

Chapter 2. Developing Pseudo Continuous Pedotransfer Functions for International Soils Measured with the Evaporation Method and the HYPROP System: I. The Soil Water Retention Curve

Abstract

Direct measurements of soil hydraulic properties are time-consuming, challenging, and often expensive. Therefore, their indirect estimation via pedotransfer functions (PTFs) based on easily collected properties like soil texture, bulk density, and organic matter content is desirable. This study was carried out to assess the accuracy of the pseudo continuous neural network PTF (PC_{NN}-PTF) approach for estimating the soil water retention curve of 153 international soils (a total of 12,654 measured water retention pairs) measured via the evaporation method. In addition, an independent data set from Turkey (79 soil samples with 7729 measured data pairs) was used to evaluate the reliability of the PC_{NN}-PTF. The best PC_{NN}-PTF showed high accuracy (root mean square error (RMSE) = 0.043 cm^3 cm^{-3}) and reliability (RMSE = 0.061 cm^3 cm^{-3}). When Turkish soil samples were incorporated into the training data set, the performance of the PC_{NN}-PTF was enhanced by 33%. Therefore, to further improve the performance of the PC_{NN}-PTF for new regions, we recommend the incorporation of local soils, when available, into the international data sets and developing new sets of PC_{NN}-PTFs.

2.1 Introduction

Pedotransfer functions (PTFs) are statistical tools used in soil science to estimate soil hydraulic properties, mainly the soil water retention curve (SWRC), based on the easily collected basic soil properties, available from most regional and national databases

(Vereecken et al., 2010). The field-scale applications of the water flow and solute transport models, and calculations of soil available water content, a widely used parameter in agronomic models, is greatly facilitated by the development of PTFs. The SWRC provides critical information about the soil moisture dynamics ("movement", "flow" and "transport") in unsaturated soils and has a wide range of applications including estimation of field capacity and soil available water (Githinji et al., 2009), hydraulic conductivity (van Genuchten, 1980), horizontal and vertical infiltration (Prevedello and Armindo, 2016), and modeling-related problems in porous media (Gallipoli et al., 2003; Ghaffaripour et al., 2019).

Point PTFs (Gupta and Larson, 1979; Haghverdi et al., 2012; Pachepsky et al., 1996; Rawls et al., 1982) estimate soil moisture at specific points of the SWRC, such as field capacity or wilting point. Parametric PTFs (Børgesen and Schaap, 2005; Haghverdi et al., 2020b; Minasny et al., 1999; Wösten and van Genuchten, 1988) estimate the parameters of a soil hydraulic function that describes the water retention across a wide range of pressure heads. Parametric PTFs are more prevalent because of their continuous representation of SWRC and their ability to provide soil hydraulic parameter estimates for use in hydrological models. Developing parametric PTFs involves fitting a soil hydraulic model to individual water-retention points and subsequently estimating the parameters of that model using basic soil properties. The widely used parametric PTFs such as Rosetta (Schaap et al., 2001; Zhang and Schaap, 2017) and Neuro-m (Minasny and McBratney, 2002) use artificial neural networks (NNs) to estimate the parameters of the van Genuchten

water retention model (van Genuchten, 1980), which are then used to estimate the entire SWRC.

The pseudo-continuous NN PTF (PC_{NN}-PTF) (Haghverdi et al., 2012) was introduced as an alternative approach for continuous estimation of the SWRC at any desired water retention. PC_{NN}-PTF utilizes statistical data mining techniques to estimate the shape of the SWRC based on actual measured data points, unlike parametric PTFs, where the curvature is dictated by the selected soil hydraulic equation. Haghverdi et al. (2018, 2015, 2014) and Nguyen et al. (Nguyen et al., 2017) reported high accuracy for the pseudo-continuous pedotransfer function (PC-PTF) approach and showed that it could provide similar and in some cases better performance than parametric PTFs mainly as it generates continuous water retention estimations without the use of any soil hydraulic equations.

In a recent study, Haghverdi et al. (2018) used HYPROP (Hydraulic Property Analyzer, Meter Group Inc., Pullman, WA, USA) automated evaporation-based benchtop laboratory system to generate a high-resolution water retention data set and subsequently developed water retention PC_{NN}-PTFs. They reported promising results and concluded that more attention should be given to the development of PC_{NN}-PTFs using HYPROP data for SWRC estimations. The HYPROP system works based on the extended evaporation method (Schindler et al., 2010a, 2010b) and is becoming the standard approach of measuring soil hydraulic properties in the laboratory since it has several advantages over the traditional equilibrium methods (i.e., pressure plate extractors and sandbox apparatus). First, it generates high-resolution water retention data (approximately 100 water retention data points in the 0–100 kPa range), which is of particular importance when developing

data-driven PTFs such as PC_{NN}-PTFs. In addition, depending on the soil type, it can generate WRC in wet and intermediate ranges in a few days versus months using traditional equilibrium based methods (Schelle et al., 2013). In this study, only the drying path data were used since HYPROP measurements are taken during natural evaporation-based drying of undisturbed soil samples.

Haghverdi et al. (2018) utilized a Turkish data set to develop their PC_{NN}-PTFs, and no study has been done to evaluate the performance of PC_{NN}-PTFs using a more comprehensive international data set from evaporation experiments. Recently, Schindler and Müller (2017) published a high-resolution soil hydraulic international data set using the evaporation method and HYPROP system, making it possible to evaluate the efficacy of PC_{NN}-PTFs for estimations of the SWRC with a large data set—the main objective of this study. The empirical nature of PTFs typically restricts their use to a specific region and any extrapolation must be preceded by validation of the PTFs (Patil and Singh, 2016). In practice, however, PTFs are applied to soils different than their development data sets since sufficient data to derive new PTFs are lacking in many regions around the world. Therefore, when developing new international PTFs, it is crucial to evaluate both the accuracy (testing) and reliability (validation) of the models (Haghverdi et al., 2012; Patil and Singh, 2016; Vereecken et al., 2010; Wösten et al., 2001). The accuracy, typically, shows the performance of PTF for a randomly selected subset of the development data set that was not used to derive the PTF. The reliability, however, indicates the performance of PTF beyond their statistical training limits and their geographical training area for data sets independent from the ones used to develop the PTF. Consequently, the specific objectives

of this paper are to (I) develop water retention PC_{NN}-PTFs by utilizing the international data set from evaporation experiments, (II) evaluate the accuracy and reliability of the PC_{NN}-PTFs using the international data set from evaporation experiments and an independent Turkish data set and (III) determine whether incorporating the Turkish soils into the development data set improves the reliability of the PTFs.

2.2 Materials and Methods

2.2.1 Soil Data Sets

Two data sets were used in this study to develop PC_{NN}-PTFs and evaluate their accuracy and reliability. The primary data set was published by Schindler and Müller (Schindler and Müller, 2017), hereafter referred to as the international data set, consisting of 173 soils from 71 sites collected from over the world (Figure 2.1). The international data set contains measurements of water retention, unsaturated hydraulic conductivity, and several basic soil properties, including textural data, organic matter content (SOM), and dry bulk density (BD) (Schindler and Müller, 2017). The hydraulic properties for the samples collected before 2007 (n = 40) had been measured using the evaporation method (Schindler, 1980). A short, saturated soil column was placed on a balance and was exposed to evaporation while the water loss per volume and tension (measured with tensiometers placed at two depths) were monitored. For the samples collected after 2008 (n = 133), the water retention data were determined with the extended evaporation method (EEM) using the HYPROP system. Schindler et al. (2010a, 2010b) extended the measurement range of the evaporation method up close to the wilting point by utilizing improved tensiometers, maximal degassing of the tensiometers, and by considering the air-entry pressure of the

tensiometer's porous ceramic cup as an additional tension measurement. For more information about the HYPROP system, readers are referred to Schindler et al. (2016). The second data set (referred to as the Turkish data set) consisted of 79 repacked samples with 7729 hydraulic measured water retention data pairs using the HYPROP system. The samples were collected from areas surrounding Ankara and Anamur, Turkey. The SOM was estimated from measured soil organic carbon content using the modified method of Walkley and Black (Jackson, 2005). Soil texture (percentages of soil separates, including sand, silt, and clay) was measured using the hydrometer method (Gee and Bauder, 1986). For more details about the soil data set and the laboratory procedures, readers are referred to Haghverdi et al. (2020a).

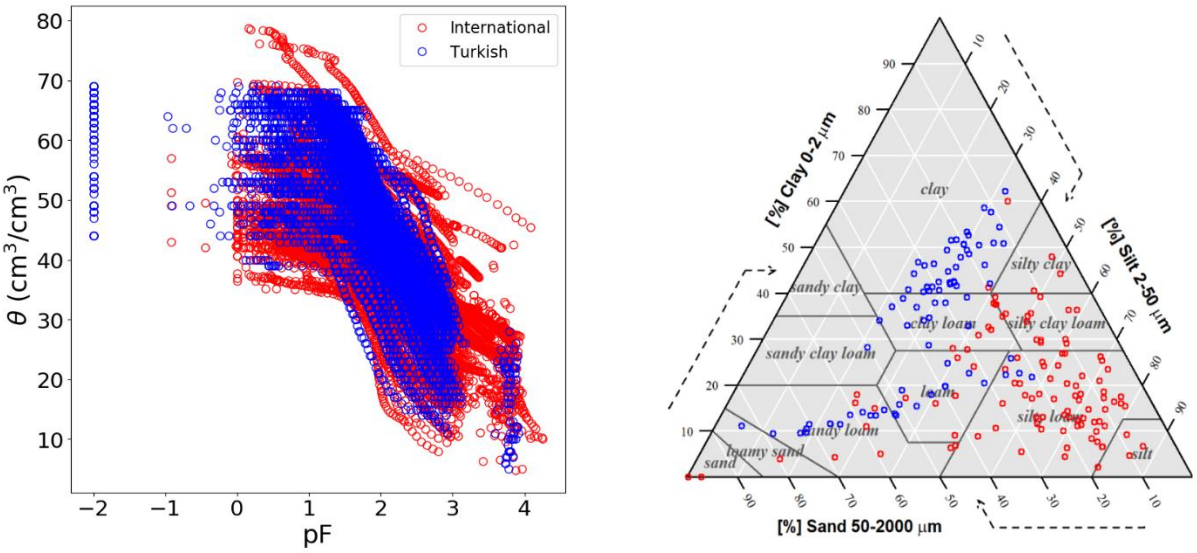
Figure 2.1. Number and origin of the undisturbed soil core samples for the international data set used in this study to develop pedotransfer functions.

The characteristics of the soils are shown in Table 2.1. The water retention data and the soil textural classification of the samples from the data sets used in this study are shown

in Figure 2.2. After screening the international data set, a subset of samples with water retention information (i.e., 153 soils with 12,654 total water retention data pairs) was selected for this study. The majority of the soil samples in the data set were from arable lands. However, samples were also collected from other land use types such as urban land, grassland, forests, fallow lands and riverbanks. These samples were collected from multiple soil horizons at depths ranging from surface to 310 cm (Schindler and Müller, 2017). The soil textural data were log-linear transformed to convert 63 μm silt-sand particle size limit used in the original data set to 50 μm silt-sand limit to match the USDA soil textural classification system. The most dominant texture in the international data set was silt loam constituting 79 soil samples (51.6% of the data set) followed by loam consisting of 19 samples (12.4% of the data set). The measured volumetric water content (VWC) ranged from 0.05 to 0.79 cm^3 cm^{-3} with an average of 0.38 cm^3 cm^{-3}. The logarithmic transformation of soil tension in cm of water (pF values) ranged from −0.9 to 4.3, with an average value of 2.0. The most dominant texture in the Turkish data set was clay constituting 38 soil samples (48.1% of the data set) followed by sandy loam consisting of 13 soil samples (16.5% of the data set). The measured water retention points of the Turkish data set ranged from full saturation (set to pF -2) to pF 3.9, with an average pF value of 1.8. The measured VWC varied between 0.05 and 0.69, with an average VWC value of 0.47 cm^3 cm^{-3}.

Table 2.1. Characteristics of soils from both international and Turkish data sets used in this study to develop and test pseudo continuous neural network pedotransfer functions (PC_{NN}-PTFs).

	International data set			Turkish dataset		
Attribute	Mean	Range	SD	Mean	Range	SD
Clay (%)	19.9	0.0–60.0	12.4	34.1	9.4–62.2	15.0
Silt (%)	56.7	0.2–86.8	17.2	30.7	5.2–57.6	8.7
Sand (%)	23.5	3.9–99.8	17.4	35.3	6.0–84.0	17.4
Bulk density (g cm^{-3})	1.33	0.55–1.69	0.23	0.98	0.69–1.33	0.14
Organic matter content (%)	3.0	0.00–12.0	2.5	1.2	0.0–3.1	0.6



(**a**) (**b**)

Figure 2.2. Soil water retention data pairs (**a**), and soil textural distribution for the data sets used in this study (**b**). Red points depict the international data set from evaporation experiments (Schindler and Müller, 2017) and blue points represent the Turkish data set (Haghverdi et al., 2020a, 2018, 2014).

2.2.2. ANN PC-PTFs Development

We developed a three-layer feed-forward perceptron NN model using MATLAB R2017a (Mathworks, 2017). The transfer functions were the "hyperbolic tangent sigmoid" and "linear" for the hidden and the output layers, respectively. The Levenberg–Marquardt

algorithm (Marquardt, 1963) was used for training the network. The maximum epoch (one cycle of a complete presentation of the training data set through the learning process) was set to 1000. The best weights were loaded automatically for testing.

Figure 2.3 illustrates the modeling workflow. Soil samples were randomly partitioned into five folds such that 80% of the data were used for the development of the PC_{NN}-PTFs and 20% as the test set. The development data set was further divided into 100 training and cross-validation subsets using a bootstrapping technique (random sampling with replacement). Each training subset was expected to have roughly 63% of the development soils (Hastie et al., 2009). The remaining development soils were used as a cross-validation subset. To eliminate the possibility of over-training, training was terminated when the root mean square error (RMSE) of the cross-validation subset either began to increase or showed no improvement. This process was repeated five times leaving aside a different fold as test such that all samples in the data set were used as a test set. The number of neurons of the hidden layer was iteratively changed from 1 to 14 to find the optimum topology of the models.

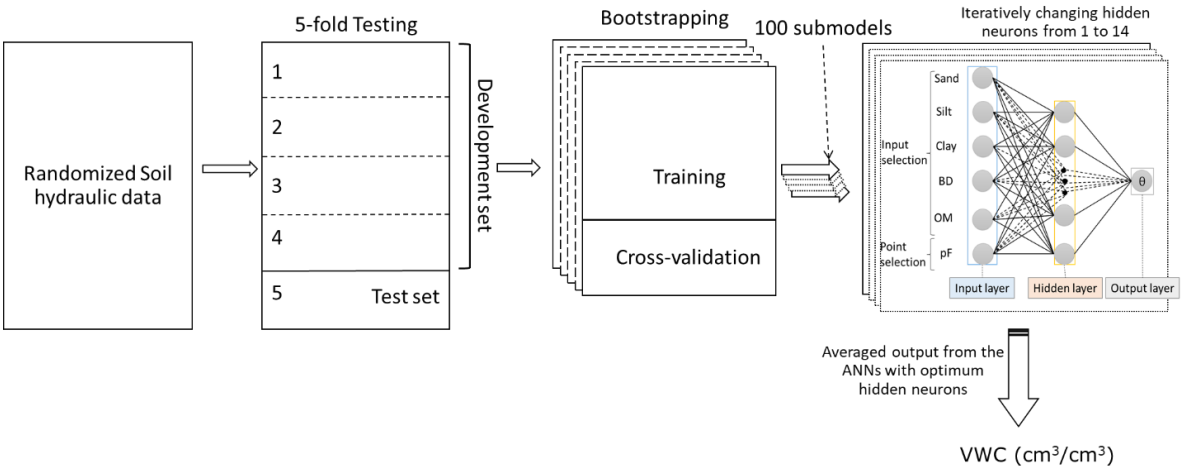


Figure 2.3. Development workflow of the pseudo continuous neural network pedotransfer functions (PC_{NN}-PTFs) for the soil water retention curve (SWRC) estimations.

The outputs of the 100 PC_{NN}-PTFs with optimum topology were averaged to obtain the water retention estimations. We then post-processed the raw outputs to make sure they are physically meaningful and water content does not increase as moving from the wet to the dry part of the SWRC. The computational cost of developing data-driven models becomes important when big data sets with a wide range of attributes are used. The data sets used for the development of PTFs (including the high-resolution evaporation-based data sets used in this study) are of relatively small size. Therefore, the computational cost of training PC_{NN}-PTFs is negligible and not discussed in this paper.

2.2.3. Modeling Scenarios

We evaluated the accuracy of the PC_{NN}-PTFs (developed using the international data set) with four combinations of the input attributes, including soil texture (i.e., percentages of sand, silt, and clay; SSC), BD, and SOM (Table 2.2). Using the logarithmic transformation of soil tension (pF) as an extra input predictor enables PC_{NN}-PTFs to

estimate VWC at any desired soil tension. The VWC is the output parameter corresponding to the input pF value. We estimated the water retention of Turkish soil samples to assess the reliability of the PC_{NN}-PTFs derived using the international data set. In addition, we developed new sets of PTFs after incorporating the Turkish soils into the training data set to determine whether including regional data into the international data set improves the reliability of the PTFs for that particular region.

Table 2.2. Combinations of input attributes (scenarios) that were used in this study to develop the pseudo continuous neural network pedotransfer functions (PC_{NN}-PTFs).

Model	Input Attributes
1	SSC, BD, SOM, pF
2	SSC, pF
3	SSC, BD, pF
4	SSC, SOM, pF

SSC: sand, silt, and clay percentages (%), BD: bulk density ($cm^3\ cm^{-3}$), SOM: soil organic matter content (%), pF: the logarithmic transformation of soil tension in cm of water.

2.2.4. Model Evaluation

The root mean square error (*RMSE*, Equation (1)), mean absolute error (*MAE*, Equation (2)), mean bias error (*MBE*, Equation (3)), and correlation coefficient (*R*, Equation (4)) were calculated to evaluate the performance of PC_{NN}-PTFs:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(E_i - M_i)^2} \quad (1)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|E_i - M_i| \quad (2)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(E_i - M_i) \quad (3)$$

$$R = \frac{\sum_{i=1}^{n}(E_i - \bar{E})(M_i - \bar{M})}{\sum_{i=1}^{n}(E_i - \bar{E})^2 \sum_{i=1}^{n}(M_i - \bar{M})^2} \quad (4)$$

where, E and M are the estimated and measured VWC ($cm^3\ cm^{-3}$), respectively, $\bar{E}$ and $\bar{M}$ are the mean estimated and measured VWC ($cm^3\ cm^{-3}$) and n is the total number of measured water retention points for each modeling scenario. In addition, the statistics were calculated separately for dominant soil textures and at the wet (pF ≤ 2), intermediate (2 < pF ≤ 3), and dry ranges (pF > 3) of the SWRC. These pF ranges were considered since a pF value of 2 (water potential of −9.8 kPa) is close to field capacity, the upper limit of available water content (Al Majou et al., 2008), and pF values greater than 3 are considered as dry ranges (Vereecken et al., 2010).

2.2.5. Domain of the Pedotransfer Functions

In most studies, the independent data set for validation of PTFs is typically described geographically or using the summary statistics of the data sets. We used the following approach to quantify the independence of the validation data set from the training set. We used Mahalanobis distance (d, Equation (5)) to evaluate which samples of the training and validation data sets belonged to the domain of applicability of the PC_{NN}-PTF (Model 1 with SSC, BD, and SOM as inputs) (Tranter et al., 2009).

$$d = \sqrt{(x - y)^T A (x - y)} \quad (5)$$

where A is the inverse of the training (international) data variance–covariance matrix, x is the individual data points in the validation data matrix, and y is the mean of the training (international) data set.

The means and covariance matrix of the predictor variables of the international data set (SSC, BD, SOM) were computed in order to calculate the Mahalanobis distance of all training points to the centroid of the training data set. Then, we computed the cut-off distance delineating the domain of the PTF as the 97.5% percentile of the cumulative $\chi 2$ distribution of the squared Mahalanobis distances (Rousseeuw and van Zomeren, 1990; Tranter et al., 2009). The Mahalanobis distance to the centroid of the training data set for all the samples of the Turkish (validation) data was computed to check if these points were within the domain of the training data set.

2.3. Results

2.3.1. Importance of the Input Predictors

Figure 2.4 illustrates the scatterplots of measured versus estimated VWC values and Table 2.3 summarizes the performance statistics for the PC_{NN}-PTFs developed and tested using different combinations of input predictors. Overall, all models showed acceptable performance, which is also demonstrated by the well-scattered data clouds (around 1:1 reference line) for all the models.

When the international data set was used for training and testing, Model 1 (inputs: SSC, BD, organic matter (SOM), pF) showed the best performance with an RMSE of 0.046 $cm^3\ cm^{-3}$ (MAE of 0.035 $cm^3\ cm^{-3}$) followed by Model 3 (inputs: SSC, BD, pF) with an RMSE of 0.047 $cm^3\ cm^{-3}$ (MAE of 0.036 $cm^3\ cm^{-3}$). Model 2, with only the soil textural components as input predictors, showed the lowest accuracy with an RMSE of 0.056 $cm^3\ cm^{-3}$ (MAE of 0.045 $cm^3\ cm^{-3}$). The low MBE values varying between 0.000 and 0.002

$cm^3\ cm^{-3}$ indicated no substantial over or underestimation. The R values were high for all the models ranging from 0.837 to 0.896, illustrating a good correlation between the measured and estimated VWC values.

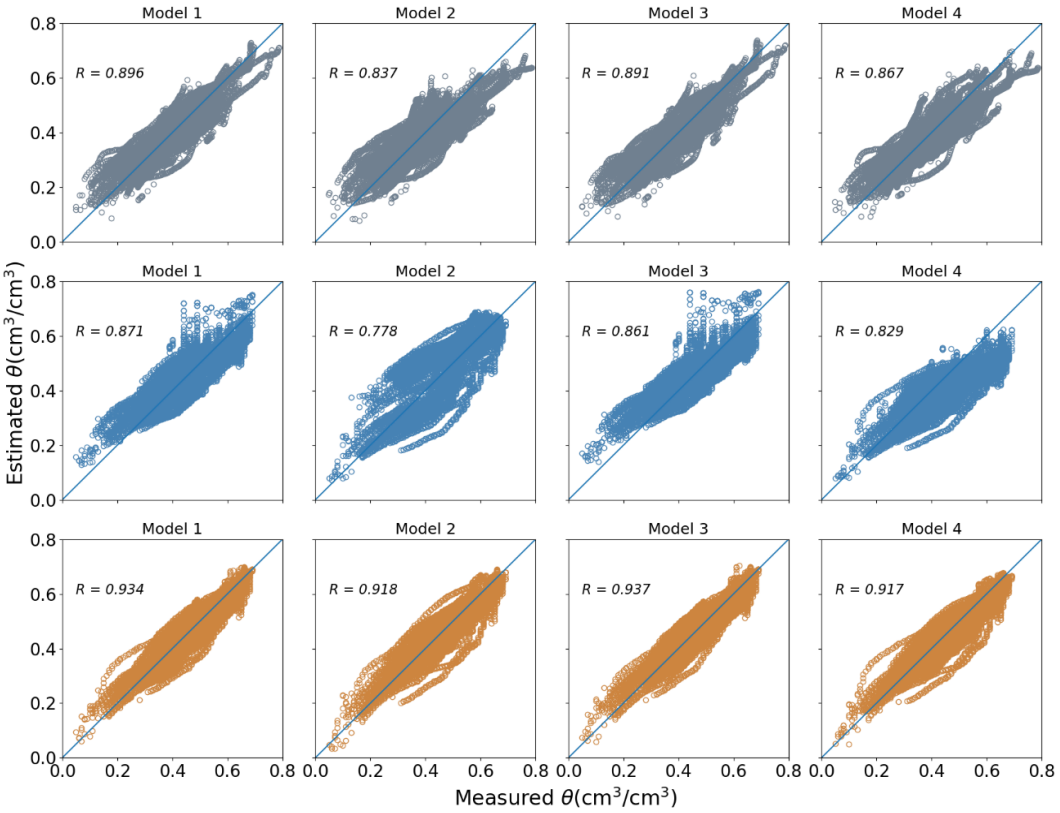


Figure 2.4. Scatterplots of the measured versus estimated volumetric water content (VWC) via PC_{NN}-PTFs when the international data set was used to train and test the models (top), and for the Turkish soil samples when Turkish data set was not used for training (middle) and when Turkish data set was incorporated into the training data set (bottom).

When the international data set was used for training and testing, Model 1 (inputs: SSC, BD, organic matter (SOM), pF) showed the best performance with an RMSE of 0.046 $cm^3\ cm^{-3}$ (MAE of 0.035 $cm^3\ cm^{-3}$) followed by Model 3 (inputs: SSC, BD, pF) with an RMSE of 0.047 $cm^3\ cm^{-3}$ (MAE of 0.036 $cm^3\ cm^{-3}$). Model 2, with only the soil textural components as input predictors, showed the lowest accuracy with an RMSE of 0.056 $cm^3\ cm^{-3}$ (MAE of 0.045 $cm^3\ cm^{-3}$). The low MBE values varying between 0.000 and 0.002

$cm^3\ cm^{-3}$ indicated no substantial over or underestimation. The R values were high for all the models ranging from 0.837 to 0.896, illustrating a good correlation between the measured and estimated VWC values.

Table 2.3. Comparison between the performance of the PC_{NN}-PTFs trained using different data sets to estimate the volumetric water content ($cm^3\ cm^{-3}$) of the international and Turkish soil samples.

	Training & Test: I				Training: I; Validation: T				Training: I + T; Test: T			
M	**RMSE**	**MAE**	**MBE**	**R**	**RMSE**	**MAE**	**MBE**	**R**	**RMSE**	**MAE**	**MBE**	**R**
1	0.046	0.035	0.002	0.896	0.061	0.051	−0.003	0.871	0.044	0.035	−0.002	0.934
2	0.056	0.045	0.001	0.837	0.081	0.066	0.010	0.778	0.049	0.039	−0.006	0.918
3	0.047	0.036	0.001	0.891	0.064	0.053	0.001	0.861	0.043	0.035	−0.002	0.937
4	0.051	0.040	0.000	0.867	0.092	0.078	−0.060	0.829	0.050	0.040	−0.012	0.917

M: Model, RMSE: Root mean square error ($cm^3\ cm^{-3}$), MAE: mean absolute error ($cm^3\ cm^{-3}$), MBE: mean biased error ($cm^3\ cm^{-3}$), R: correlation coefficient. I: international data set, T: Turkish data set.

When the Turkish data set was used as a validation set, Model 1 (inputs: SSC, BD, OM, pF) showed the best performance with an RMSE of 0.061 $cm^3\ cm^{-3}$ (MAE of 0.051 $cm^3\ cm^{-3}$) followed by Model 3 (inputs: SSC, BD) with an RMSE of 0.064 $cm^3\ cm^{-3}$ (MAE of 0.053 $cm^3\ cm^{-3}$). Model 4 (inputs: SSC, OM), showed the lowest performance with RMSE of 0.092 $cm^3\ cm^{-3}$ (MAE of 0.078 $cm^3\ cm^{-3}$). Model 4, with an MBE of −0.060 $cm^3\ cm^{-3}$, showed a tendency to underestimate the VWC, which is also depicted in Figure 2.4. The R values ranged from 0.778 to 0.871 with the lowest R observed for Model 2 and comparable values for the other models.

When the Turkish data set was incorporated into training and used as a test, Model 3 (inputs: SSC, BD) showed the best performance with an RMSE of 0.043 $cm^3\ cm^{-3}$ (MAE

of 0.035 cm^3 cm^{-3}) followed by Model 1 (inputs: SSC, BD, OM) with an RMSE of 0.044 cm^3 cm^{-3} (MAE of 0.035 cm^3 cm^{-3}). Model 4 (inputs: SSC, OM), showed the lowest accuracy with an RMSE of 0.050 cm^3 cm^{-3} (MAE of 0.040 cm3 cm^{-3}). The low MBE values ranging from −0.012 to −0.002 cm^3 cm^{-3} indicated no sign of systematic bias in any of the models. The R values were high for all the models ranging from 0.917 to 0.937, showing a good correlation between the measured and estimated VWC values.

2.3.2. Performance across Soil Textures

Table 2.4 summarizes the performance of the best performing model (i.e., Model 1 with SSC, BD, and OM as inputs) across dominant textures (textures constituting more than 10% percent of the data set) developed and tested using the international data set. The smallest error (RMSE: 0.04 cm^3 cm^{-3}; MAE: 0.028 cm^3 cm^{-3}) values belonged to silt clay loam and the greatest error belonged to clay loam (RMSE 0.052 cm^3 cm^{-3}; MAE 0.038 cm^3 cm^{-3}). The other textures (i.e., silt loam, loam, and sandy loam) showed similar performance with MAE varying from 0.033 to 0.034 cm^3 cm^{-3}. The MBE values of 0.016 and −0.016 cm^3 cm^{-3} suggested a slight tendency for over and underestimation for silty clay loam and clay loam textures, respectively. MBE values were negligible (close to zero) for other soil textures. The correlation coefficient values ranged from 0.824 to 0.935 among the textures with the greatest value observed for loam and lowest for sandy loam.

Table 2.5 shows the performance of the best performing model (i.e., Model 1 with SSC, BD, and OM as inputs) for the most dominant soil textures of the Turkish data set constituting roughly 92 percent of the data set.

Table 2.4. Soil texture-based performance of the PC_{NN}-PTFs (inputs: SSC, BD, OM, pF) developed and tested using the international data set to estimate the volumetric water content ($cm^3\ cm^{-3}$).

	Silt Loam	Loam	Silty Clay Loam	Clay Loam	Sandy Loam
RMSE	0.043	0.042	0.04	0.052	0.043
MAE	0.034	0.033	0.028	0.038	0.033
MBE	0.002	0.004	0.016	−0.016	0.009
R	0.888	0.935	0.824	0.926	0.882

RMSE: Root mean square error ($cm^3\ cm^{-3}$), MAE: mean absolute error ($cm^3\ cm^{-3}$), MBE: mean biased error ($cm^3\ cm^{-3}$), R: correlation coefficient.

Table 2.5. Soil texture based performance of the PC_{NN}-PTFs (Model 1 with SSC, BD, SOM, and pF as inputs) developed using the international data set and the international plus Turkish data sets to estimate the volumetric water content ($cm^3\ cm^{-3}$) of the Turkish soil samples.

	Training: International				Training: International + Turkish			
	C	SL	CL	L	C	SL	CL	L
RMSE	0.060	0.069	0.052	0.060	0.039	0.047	0.044	0.042
MAE	0.052	0.055	0.042	0.048	0.032	0.037	0.035	0.034
MBE	−0.006	0.032	−0.019	−0.009	−0.001	0.001	−0.001	−0.004
R	0.879	0.813	0.905	0.820	0.938	0.895	0.910	0.907

RMSE: Root mean square error ($cm^3\ cm^{-3}$), MAE: mean absolute error ($cm^3\ cm^{-3}$), MBE: mean biased error ($cm^3\ cm^{-3}$), R: correlation coefficient. C: Clay, SL: Sandy Loam, CL: Clay loam, L: Loam.

When the Turkish data set was only used as a validation set, the lowest RMSE (0.052 $cm^3\ cm^{-3}$) and MAE (0.042 $cm^3\ cm^{-3}$) values belonged to clay loam, whereas sandy loam showed the highest values (RMSE = 0.069; MAE = 0.055). MBE values of −0.019 and 0.032 indicated slight underestimation and moderate overestimation for clay loam and sandy loam soil textures, respectively. The correlation coefficient varied from 0.813 for sandy loam to 0.905 for clay loam soil textures. When the Turkish soils were incorporated into the training phase, lowest RMSE (0.039 $cm^3\ cm^{-3}$) and MAE (0.032 $cm^3\ cm^{-3}$) belonged to clay, whereas the highest values were observed for sandy loam with RMSE

and MAE of 0.047 and 0.037 $cm^3\ cm^{-3}$, respectively. MBE values were close to zero (from −0.001 to 0.001), indicating no systematic bias for any of the models. The lowest and highest R values ranging from 0.895 to 0.938 were observed for sandy loam and clay textures, respectively.

2.3.3. Performance at the Wet, Intermediate and Dry Parts of the SWRC

Table 2.6 shows the performance of the best performing PC_{NN}-PTF (i.e., model 1 with SSC, BD, and OM as inputs) in wet ($pF \leq 2$), intermediate ($2 < pF \leq 3$) and dry ($pF > 3$) parts of the SWRC. When the international data set was used for training and testing, the lowest RMSE (0.041 $cm^3\ cm^{-3}$) and MAE (0.031 $cm^3\ cm^{-3}$) values were observed in the wet range of the SWRC. The intermediate range of the SWRC showed a relatively higher error with RMSE and MAE values of 0.05 and 0.039 $cm^3\ cm^{-3}$, respectively. The relatively higher and lower performances at the wet and intermediate parts were also evident by the R values of 0.868 and 0.733, respectively. MBE range of −0.008 to 0.007 suggested no bias for any of the models.

When the Turkish data set is used as a validation set, the lowest RMSE (0.061 $cm^3\ cm^{-3}$) and MAE (0.05 $cm^3\ cm^{-3}$) belonged to the wet range while the highest RMSE and MAE of 0.066 and 0.059 $cm^3\ cm^{-3}$, respectively, belonged to the dry range of the SWRC. Underestimation of the VWC was observed in the wet range as indicated by the negative MBE (−0.018 $cm^3\ cm^{-3}$) while overestimation was evident in intermediate (MBE: 0.021 $cm^3\ cm^{-3}$) and dry parts (MBE: 0.058 $cm^3\ cm^{-3}$) of the SWRC. The R values varied from 0.661 to 0.902, with the lowest and highest values belonging to the intermediate and dry ranges, respectively.

Table 2.6. Performance of the PC_{NN}-PTFs (inputs: SSC, BD, OM, and pF as) developed using the international data set and the international plus Turkish data sets to estimate the volumetric water content ($cm^3\ cm^{-3}$) at wet ($pF \leq 2$) intermediate ($2 < pF \leq 3$) and dry ($pF > 3$) parts of the SWRC.

	Training and Test: I			**Training: I; Validation: T**			**Training: I + T; Test: T**		
	Wet	**Mid**	**Dry**	**Wet**	**Mid**	**Dry**	**Wet**	**Mid**	**Dry**
RMSE	0.041	0.050	0.043	0.061	0.062	0.066	0.041	0.049	0.037
MAE	0.031	0.039	0.034	0.050	0.052	0.059	0.032	0.039	0.028
MBE	−0.001	0.007	−0.008	−0.018	0.021	0.058	−0.003	0.000	0.015
R	0.868	0.733	0.790	0.713	0.661	0.902	0.866	0.778	0.883

RMSE: Root mean square error ($cm^3\ cm^{-3}$), MAE: mean absolute error ($cm^3\ cm^{-3}$), MBE: mean biased error ($cm^3\ cm^{-3}$), R: correlation coefficient. I: International data set, T: Turkish data set.

When the Turkish soils were incorporated into the training phase, lowest RMSE (0.037 $cm^3\ cm^{-3}$) and MAE (0.028 $cm^3\ cm^{-3}$) values were observed in the dry range and highest values of 0.049 and 0.039 $cm^3\ cm^{-3}$, respectively, belonged to the intermediate range. MBE value of 0.015 $cm^3\ cm^{-3}$ suggested a tendency to overestimate VWC in the dry range. R values ranged from 0.778 to 0.883 and were higher and comparable in the wet and dry ranges, whereas the intermediate range showed the lowest correlation.

2.4. Discussion

2.4.1. Accuracy and Reliability of the Developed PTFs

Table 2.7 summarizes the performance of already published PC-PTFs and PTFs developed in this study. The accuracy of previous PC-PTFs developed to estimate water retention range from RMSE of 0.027 to 0.159 $cm^3\ cm^{-3}$, while the reliability ranges from RMSE of 0.036 to 0.088 $cm^3\ cm^{-3}$ (Table 2.7). As shown in Table 2.3, the high accuracy of PC_{NN}-PTF developed in this study (RMSE = 0.046 $cm^3\ cm^{-3}$) puts it in a good performance rank among already published PC-PTFs. Therefore, PC_{NN}-PTF is a reliable approach for developing accurate water retention models using international data from

evaporation experiments. The PC_{NN}-PTF developed by Haghverdi et al.(2018) was the only other PTF that was based on a data set with soil water retention points measured with the extended evaporation method, using the Turkish data set. Other studies used data sets where the soil water retention pairs were collected using equilibrium-based methods (i.e., pressure plate/sandbox). Not all the studies used a totally independent data set for validation except Haghverdi et al. (2012), whereas the validation data set in our study was independent of the international PTF-development data set.

Table 2.7. Comparison of the pseudo-continuous pedotransfer functions (PC-PTFs) developed in the literature to the PC_{NN}-PTF developed in this study.

Study	Method	Modeling Approach	Inputs	Origin, no. Samples/Datapoints	RMSE (cm^3 cm^{-3}) Test	Validation
(Haghverdi et al., 2012)	Iranian data from pressure plate and Australian data set using various equilibrium-based methods	NN	SSC	(Traing and Test- 122 soil samples from Iran) (772 soil samples for training from Australia, Validation- Iran)	0.029	0.037
			SSC, BD	-	0.028	0.037
			SSC, OC	-	0.028	0.036
			SSC, BD, OC	-	0.027	0.036
(Haghverdi et al., 2014)	sandbox/pressure plate	NN	SSC, BD, SOM	Turkey, 135 soil samples x 8 SWR points	0.047	-
				Belgium, (69 soil samples x 8 to 10 SWR points)	0.040	-
		SVM	SSC, BD, SOM	Turkey	0.054	
				Belgium	0.069	
(Moreira De Melo and Pedrollo, 2015)	different equilibrium-based methods (Pressure based, hanging water, tensiometer, and sand-box)	NN	SSC, particle density, total porosity, BD	UNSODA, (137 soil samples for training and 51 for validation)		0.088
(Nguyen et al., 2017)	sand-boxes and pressure chambers	NN	SSC, BD, OC	Vietnamese Mekong Delta, (1280 data points for training, 232 validation)	0.044	0.052
(Haghverdi et al., 2018)	evaporation	MLR		-	0.056	0.066
		SVM		-	0.036	0.068
		k-NN		-	0.056	0.050
		NN	SSC	Turkey, (81 soil samples)	0.129	
			SSC, BD	-	0.080	
			SSC, SOM	-	0.159	
			SSC, SA	-	0.107	
			SSC, SA, BD, SOM	-	0.061	
			SSC, BD, OM, SA, IWC	-	0.033	

SVM: support vector machine, MLR: multiple linear regression, NN: artificial neural network, *k*-NN: *k*-nearest neighbor, SSC: sand, silt, and clay percentages (%), BD: bulk density (cm^3 cm^{-3}), SOM: soil organic matter content (%), OC: organic carbon content (%), SA: percentage of stable aggregates, IWC: initial water content (cm^3 cm^{-3}).

The analysis of the Mahalanobis distances revealed that only eight soil samples from the validation data were below the cut-off limit (Figure 2.5), indicating that the two data sets used in this study were independent with a slight overlap. Despite the difference between the data sets, the PC_{NN}-PTF showed high reliability with an RMSE equal to 0.061 $cm^3\ cm^{-3}$ (Table 2.3). An RMSE of 0.043 $cm^3\ cm^{-3}$ was further achieved when Turkish data was included in the training of the PC_{NN}-PTF. Therefore, incorporation of local HYPROP data sets, if available, and retraining the PC_{NN}-PTF is recommended to further enhance the performance of the model for new regions. The ability of NNs to mimic the inputs–outputs relationship of the complex soil water system (Pachepsky and Schaap, 2004) can explain the adequate performance of PC_{NN}-PTFs in both training and validation phases.

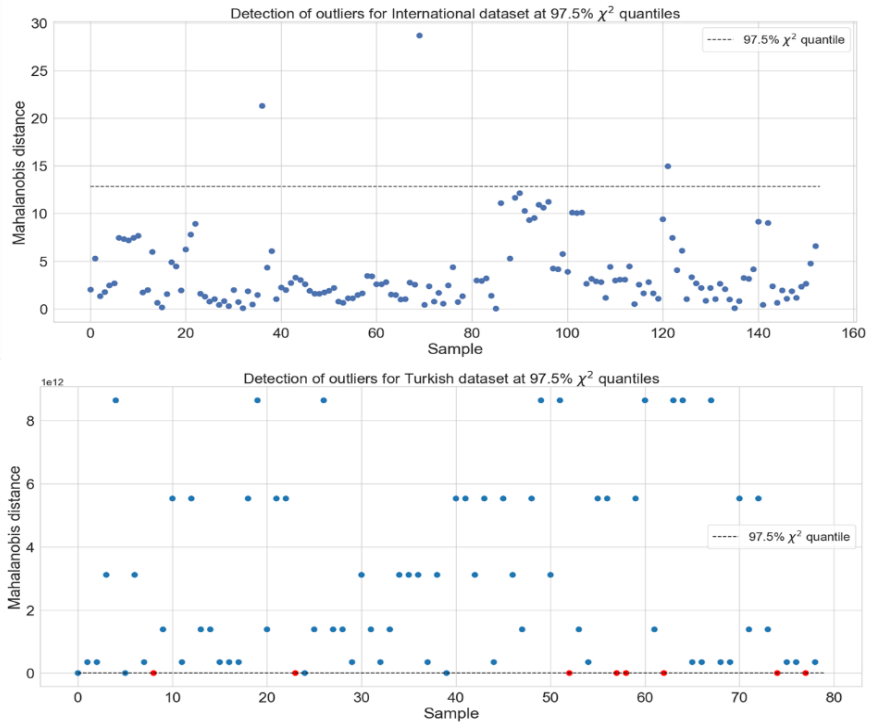


Figure 2.5. The domain of the developed PC_{NN}-PTFs using Mahalanobis distance, indicating that the two data sets were independent with a slight overlap since only 8 Turkish soil samples (highlighted in red) fell below the cut-off limit (y-axis for Turkish data set is on an exponential scale).

Several studies have recommended the use of local data set to develop PTFs instead of using larger data sets (McBratney et al., 2002; Nemes et al., 2003). Inconsistencies in the measurement techniques used in large data sets can introduce unexplained variance and negatively impact the performance of PTFs (Vereecken et al., 2010). The international data set used in our study contains soil samples collected on the continental scale, yet PTFs performed satisfactorily across modeling scenarios. This is in part because all measurements for both development and test data sets were done using the evaporation or extended evaporation methods. We recommend using the HYPROP system as a benchmark laboratory approach to maintain consistency in measurement techniques when adding local

samples to the international data set used in this study to develop new PC_{NN}-PTFs in the future.

2.4.2. Importance of Input Variables

Various studies have found that the addition of more input variables to the models did not necessarily result in better performance of the PTFs (Nemes et al., 2006; Schaap et al., 2004). The best performance in our study was observed for Model 1 using all the input predictors (SSC, BD, SOM) with RMSE of 0.046 cm^3 cm^{-3} for the test and RMSE of 0.061 cm^3 cm^{-3} for the validation sets. However, Model 3 also resulted in a comparable performance when using SSC and BD as inputs. Moreover, Model 3 was the best performing with RMSE of 0.043 cm^3 cm^{-3} when the Turkish data set was incorporated in the training set, which agrees with the results reported by Patil et al. (2013). Minasny and McBratney (2002) also found that adding BD improved the performance of the neuro-m model compared to using just the textural constituents. Moreover, the inclusion of BD as the input variable along with the soil texture resulted in better performance in both Neuro-m and Rosetta 3 PTFs to estimate water retention (Minasny and McBratney, 2002; Zhang and Schaap, 2017).

Including SOM as an input predictor did not improve the performance of the PC_{NN}-PTF in our study. Zacharias and Wessolek (2007) and Børgesen et al. (2008) also reported that SOM does not contribute to the model performance. Minasny and McBratney (2018) conducted a meta-analysis to conclude that an increase in the SOM only resulted in a small increase in the soil water content. Haghverdi et al. (2018) mentioned that the insignificant impact of SOM in their study could be due to its low concentration and narrow range in

most of the Turkish soil samples, which concur with the findings of our study despite having a larger range of SOM in the international data set.

While comparing to other PC-PTFs in literature, RMSE of 0.088 cm^3 cm^{-3} was observed by PC_{NN}-PTF of Moreira De Melo and Pedrollo (2015) using additional inputs such as particle density and porosity along with soil texture and bulk density. An accuracy with RMSE of 0.033 cm^3 cm^{-3} was observed by Haghverdi et al. (2018) when information about stable aggregates and initial water content was included in the training along with other inputs including SSC, BD, and SOM. Although adding more input predictors, if available, could enhance the performance of PTFs, our results indicate soil texture (SSC) and bulk density (BD) as the essential inputs required to develop accurate PC_{NN}-PTFs using evaporation data. These properties are also easily collected and are available in most data sets; thus, we recommended them to be included in future SWRC measurement campaigns using the HYPROP system.

2.4.3. Performance across Textural Classes and Tension Ranges

Generally, we observed that having more data points (due to having more soil samples) per textural classes in the training set improved the performance for that class (Figure 2.6). Khlosi et al. (2008) provided error statistics for 11 textural classes and found that the PTFs performed well in the relatively coarse-textured soils compared to heavy-textured soils. They observed better PTF performance for textural classes with somewhat larger sample size. Schaap et al. (1998) reported a relatively lower RMSE for the sandy loam and clay loam soil, which contributed 34% of their data set.

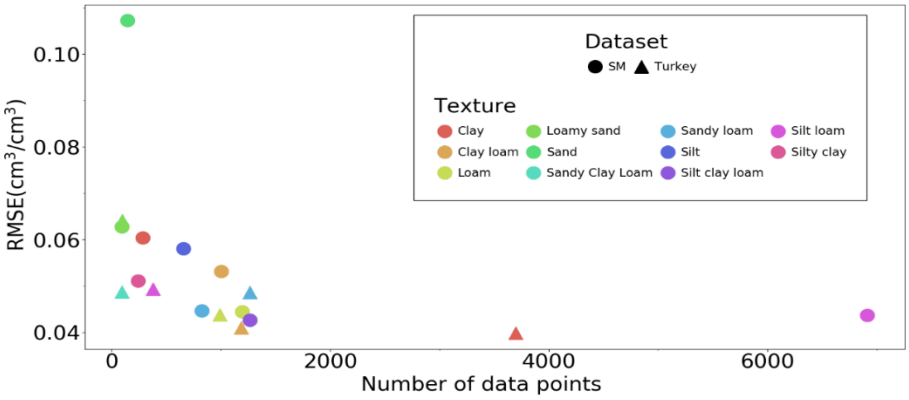


Figure 2.6. Relationship between the number of data points for each textural class and the accuracy of the best performing PC_{NN}-PTF (Model 3 with SSC, BD, and pF as inputs) when both international and Turkish data sets were used to develop the models.

For the international data set, silt loam was the dominant textural class followed by silty clay loam, and both showed a high agreement with the fitted curves. For the Turkish data set, clay as the dominant class shows a high agreement with the fitted curves compared to textural classes with lower data percentage share. Similar results were reported by (Schaap and Leij (1998) and Cornelis et al. (2001). Thus, it is possible to predict the SWRC accurately if enough data points are available for the soil texture in the training set (Haghverdi et al., 2014; Vereecken et al., 2010).

The best performance of the PC_{NN}-PTF was observed in the wet region of the SWRC for the test and validation sets, while the lowest accuracy was observed in the dry region for the validation data set, which concurs with the performance of parametric PTFs of Khlosi et al. (2008) and Børgesen and Schaap (2005). However, when Turkish data was included in the training of the models, the dry region had the best performance while the intermediate part showed a lower accuracy. Nonetheless, an improvement of 61%, 73%, and 49% in RMSE was observed in the wet, intermediate, and dry regions, respectively,

after incorporating Turkish data into training. Schaap et al. (2001) and Twarakavi et al. (2009) reported overestimation of soil water retention close to saturation (pF < 0.5) and between pF of 0.5 to 1, and underestimation beyond pF of 1.5, which is in contrast to what we observed in our study. This is, in part, attributed to the fact that the training data set used by Schaap et al. (2001) and Twarakavi et al. (2009) consisted of samples collected from several studies with a wide range of approaches used to measure water retention. Moreover, these studies developed parametric PTFs which means the shape of the curve was governed by the van Genuchten water retention model (van Genuchten, 1980). However, the PC_{NN}-PTF developed in our study learns the SWRC's shape from the measured water retention data without using any soil hydraulic model.

2.5. Conclusions

Literature suggests that PTFs developed from small local data sets perform better as compared to larger general sets of data (Nemes et al., 2003). However, in many parts of the world, there is a lack of soil hydraulic data to derive PTFs for accurate SWRC estimations. Most of the large data sets (e.g., UNSODA (Nemes et al., 2001), and HYPRES (Wösten et al., 1999) used in the past to develop international PTFs typically consist of smaller data sets with a wide range of measurement techniques applied to measure soil hydraulic properties. Having a PTF trained on an international data set with soil hydraulic properties measured with the same technique minimizes the inconsistency in the data set caused by variabilities in measurement techniques. We used an international data set from evaporation experiments (Schindler and Müller, 2017) to evaluate the accuracy and reliability of the PC_{NN}-PTF approach to estimate the SWRC. Evaporation based

measurement of water retention offers the advantage of producing a quasi-continuous description of the retention function in the tensiometric moisture range, i.e., up to pF 3. In practice, HYPROP measurements lead to roughly ten times more data points compared to the traditional method via sandbox/pressure plate instruments. We found that a neural network-based PC-PTF can provide accurate and reliable estimation of the SWRC. Moreover, the reliability was further improved by including the local data into the training of PC_{NN}-PTF. Therefore, we recommend retraining the models after incorporating local HYPROP data sets (if available) to enhance the performance of the PC_{NN}-PTFs developed in this study in different regions around the world.

2.5. References

Al Majou, H., Bruand, A., Duval, O., Le Bas, C., Vautier, A., 2008. Prediction of soil water retention properties after stratification by combining texture, bulk density and the type of horizon. Soil Use Manag. 24, 383–391. https://doi.org/10.1111/j.1475-2743.2008.00180.x

Børgesen, C.D., Iversen, B. V., Jacobsen, O.H., Schaap, M.G., 2008. Pedotransfer functions estimating soil hydraulic properties using different soil parameters. Hydrol. Process. 22, 1630–1639. https://doi.org/10.1002/hyp.6731

Børgesen, C.D., Schaap, M.G., 2005. Point and parameter pedotransfer functions for water retention predictions for Danish soils. Geoderma 127, 154–167. https://doi.org/10.1016/j.geoderma.2004.11.025

Cornelis, W.M., Ronsyn, J., Van Meirvenne, M., Hartmann, R., 2001. Evaluation of Pedotransfer Functions for Predicting the Soil Moisture Retention Curve. Soil Sci. Soc. Am. J. 65, 638–648. https://doi.org/10.2136/sssaj2001.653638x

Gallipoli, D., Gens, A., Sharma, R., Vaunat, J., 2003. An elasto-plastic model for unsaturated soil incorporating the effects of suction and degree of saturation on mechanical behaviour. Geotechnique 53, 123–135. https://doi.org/10.1680/geot.2003.53.1.123

Gee, G.W., Bauder, J.W., 1986. Particle-size Analysis, in: Klute, A. (Ed.), Methods of Soil Analysis. Part 1. SSSA Book Ser. 5. SSSA, Madison, WI. https://doi.org/10.2136/sssabookser5.1.2ed.c15

Ghaffaripour, O., Esgandani, G.A., Khoshghalb, A., Shahbodaghkhan, B., 2019. Fully

coupled elastoplastic hydro-mechanical analysis of unsaturated porous media using a meshfree method. Int. J. Numer. Anal. Methods Geomech. 43, 1919–1955. https://doi.org/10.1002/nag.2931

Githinji, L.J.M., Dane, J.H., Walker, R.H., 2009. Water-use patterns of tall fescue and hybrid bluegrass cultivars subjected to ET-based irrigation scheduling. Irrig. Sci. 27, 377–391. https://doi.org/10.1007/s00271-009-0153-4

Gupta, S.C., Larson, W.E., 1979. Estimating Soil Water Retention Characteristics From Particle Size Distribution, Organic Matter Percent, and Bulk Density. Water Resour. Res. 15, 1633–1635.

Haghverdi, A., Cornelis, W.M., Ghahraman, B., 2012. A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. J. Hydrol. 442–443, 46–54. https://doi.org/10.1016/j.jhydrol.2012.03.036

Haghverdi, A., Leib, B.G., Washington-Allen, R.A., Ayers, P.D., Buschermohle, M.J., 2015. High-resolution prediction of soil available water content within the crop root zone. J. Hydrol. 530, 167–179. https://doi.org/10.1016/j.jhydrol.2015.09.061

Haghverdi, A., Najarchi, M., Öztürk, H.S., Durner, W., 2020a. Studying Unimodal, Bimodal, PDI and Bimodal-PDI Variants of Multiple Soil Water Retention Models: I. Direct Model Fit Using the Extended Evaporation and Dewpoint Methods. Water 12, 16–23. https://doi.org/10.3390/w12030900

Haghverdi, A., Öztürk, H.S., Cornelis, W.M., 2014. Revisiting the pseudo continuous pedotransfer function concept: Impact of data quality and data mining method. Geoderma 226–227, 31–38. https://doi.org/10.1016/j.geoderma.2014.02.026

Haghverdi, A., Öztürk, H.S., Durner, W., 2020b. Studying Unimodal, Bimodal, PDI and Bimodal-PDI Variants of Multiple Soil Water Retention Models: II. Evaluation of Parametric Pedotransfer Functions Against Direct Fits. Water 12. https://doi.org/10.3390/w12030896

Haghverdi, A., Öztürk, H.S., Durner, W., 2018. Measurement and estimation of the soil water retention curve using the evaporation method and the pseudo continuous pedotransfer function. J. Hydrol. 563, 251–259. https://doi.org/10.1016/j.jhydrol.2018.06.007

Hastie, T., Tibshirani, R., Friedman, J., 2009. The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media.

Jackson, M.L., 2005. Soil chemical analysis: Advanced course. UW-Madison Libraries Parallel Press.

Khlosi, M., Cornelis, W.M., Douaik, A., van Genuchten, M.T., Gabriels, D., 2008. Performance Evaluation of Models That Describe the Soil Water Retention Curve between Saturation and Oven Dryness. Vadose Zo. J. 7, 87. https://doi.org/10.2136/vzj2007.0099

Marquardt, D.W., 1963. An Algorithm for Leat-Squares Estimation of Nnonlinear

Parameters. J. Soc. Indust. Appl. Math. 11, 431–441.

McBratney, A.B., Minasny, B., Cattle, S.R., Vervoort, R.W., 2002. From pedotransfer functions to soil inference systems. Geoderma 109, 41–73. https://doi.org/10.1016/S0016-7061(02)00139-8

Minasny, B., McBratney, A.B., 2018. Limited effect of organic matter on soil available water capacity. Eur. J. Soil Sci. 69, 39–47. https://doi.org/10.1111/ejss.12475

Minasny, B., McBratney, A.B., 2002. The Neuro-m Method for Fitting Neural Network Parametric Pedotransfer Functions. Soil Sci. Soc. Am. J. 66, 352–361. https://doi.org/10.2136/sssaj2002.3520

Minasny, B., McBratney, A.B., Bristow, K.L., 1999. Comparison of different approaches to the development of pedotransfer functions for water-retention curves. Geoderma 93, 225–253. https://doi.org/10.1016/S0016-7061(99)00061-0

Moreira De Melo, T., Pedrollo, O.C., 2015. Artificial neural networks for estimating soil water retention curve using fitted and measured data. Appl. Environ. Soil Sci. 2015. https://doi.org/10.1155/2015/535216

Nemes, A., Rawls, W.J., Pachepsky, Y.A., 2006. Use of the Nonparametric Nearest Neighbor Approach to Estimate Soil Hydraulic Properties. Soil Sci. Soc. Am. J. 70, 327. https://doi.org/10.2136/sssaj2005.0128

Nemes, A., Schaap, M.., Leij, F.., Wösten, J.H.., 2001. Description of the unsaturated soil hydraulic database UNSODA version 2.0. J. Hydrol. 251, 151–162. https://doi.org/10.1016/S0022-1694(01)00465-6

Nemes, A., Schaap, M.G., Wösten, J.H.M., 2003. Functional Evaluation of Pedotransfer Functions Derived from Different Scales of Data Collection. Soil Sci. Soc. Am. J. 67, 1093. https://doi.org/10.2136/sssaj2003.1093

Nguyen, P.M., Haghverdi, A., de Pue, J., Botula, Y.D., Le, K. V., Waegeman, W., Cornelis, W.M., 2017. Comparison of statistical regression and data-mining techniques in estimating soil water retention of tropical delta soils. Biosyst. Eng. 153, 12–27. https://doi.org/10.1016/j.biosystemseng.2016.10.013

Pachepsky, Y., Schaap, M.G.B.T.-D. in S.S., 2004. Data mining and exploration techniques, in: Development of Pedotransfer Functions in Soil Hydrology. Elsevier, pp. 21–32. https://doi.org/https://doi.org/10.1016/S0166-2481(04)30002-4

Pachepsky, Y.A., Timlin, D., Varallyay, G., 1996. Artificial Neural Networks to Estimate Soil Water Retention from Easily Measurable Data. Soil Sci. Soc. Am. J. 60, 727–733. https://doi.org/10.2136/sssaj1996.03615995006000030007x

Patil, N.G., Singh, S.K., 2016. Pedotransfer Functions for Estimating Soil Hydraulic Properties: A Review. Pedosphere 26, 417–430. https://doi.org/10.1016/S1002-0160(15)60054-6

Patil, N.G., Tiwary, P., Pal, D.K., Bhattacharyya, T., Sarkar, D., Mandal, C., Mandal, D.K.,

Chandran, P., Ray, S.K., Prasad, J., Lokhande, M., Dongre, V., 2013. Soil water retention characteristics of black soils of india and pedotransfer functions using different approaches. J. Irrig. Drain. Eng. 139, 313–324. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000527

Prevedello, C.L., Armindo, R.A., 2016. Generalization of the Green-Ampt Theory for Horizontal Infiltration into Homogeneous Soils. Vadose Zo. J. 15, vzj2016.04.0030. https://doi.org/10.2136/vzj2016.04.0030

Rawls, W.J., Brakensiek, D.L., Saxtonn, K.E., 1982. Estimation of Soil Water Properties. Trans. ASAE 25, 1316–1320. https://doi.org/10.13031/2013.33720

Rousseeuw, P.J., van Zomeren, B.C., 1990. Unmasking multivariate outliers and leverage points. J. Am. Stat. Assoc. 85, 633–639. https://doi.org/10.1080/01621459.1990.10474920

Schaap, M.G., Leij, F.J., 1998. Using neural networks to predict soil water retention and soil hydraulic conductivity. Soil Tillage Res. 47, 37–42. https://doi.org/10.1016/S0167-1987(98)00070-1

Schaap, M.G., Leij, F.J., van Genuchten, M.T., 2001. Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. J. Hydrol. 251, 163–176.

Schaap, M.G., Leij, F.J., van Genuchten, M.T., 1998. Neural Network Analysis for Hierarchical Prediction of Soil Hydraulic Properties. Soil Sci. Soc. Am. J. 62, 847. https://doi.org/10.2136/sssaj1998.03615995006200040001x

Schaap, M.G., Nemes, A., van Genuchten, M.T., 2004. Comparison of Models for Indirect Estimation of Water Retention and Available Water in Surface Soils. Vadose Zo. J. 3, 1455–1463. https://doi.org/10.2136/vzj2004.1455

Schelle, H., Heise, L., Jänicke, K., Durner, W., 2013. Water retention characteristics of soils over the whole moisture range: A comparison of laboratory methods. Eur. J. Soil Sci. 64, 814–821. https://doi.org/10.1111/ejss.12108

Schindler, U., 1980. Ein Schnellverfahren zur Messung der Wasserleitfähigkeit im teilgesättigten Boden an Stechzylinderproben. Arch. für Acker- und Pflanzenbau und Bodenkd.

Schindler, U., Durner, W., von Unold, G., Mueller, L., Wieland, R., 2010a. The evaporation method: Extending the measurement range of soil hydraulic properties using the air-entry pressure of the ceramic cup. J. Plant Nutr. Soil Sci. 173, 563–572. https://doi.org/10.1002/jpln.200900201

Schindler, U., Durner, W., von Unold, G., Müller, L., 2010b. Evaporation Method for Measuring Unsaturated Hydraulic Properties of Soils: Extending the Measurement Range. Soil Sci. Soc. Am. J. 74, 1071. https://doi.org/10.2136/sssaj2008.0358

Schindler, U., Mueller, L., von Unold, G., Durner, W., Fank, J., 2016. Emerging Measurement Methods for Soil Hydrological Studies, in: Novel Methods for

Monitoring and Managing Land and Water Resources in Siberia. Springer, Cham, pp. 345–363. https://doi.org/10.1007/978-3-319-24409-9_14

Schindler, U., Müller, L., 2017. Soil hydraulic functions of international soils measured with the Extended Evaporation Method (EEM) and the HYPROP device. Open Data J. Agric. Res. 3, 10–16.

Tranter, G., McBratney, A.B., Minasny, B., 2009. Using distance metrics to determine the appropriate domain of pedotransfer function predictions. Geoderma 149, 421–425. https://doi.org/10.1016/j.geoderma.2009.01.006

Twarakavi, N.K.C., Šimůnek, J., Schaap, M.G., 2009. Development of Pedotransfer Functions for Estimation of Soil Hydraulic Parameters using Support Vector Machines. Soil Sci. Soc. Am. J. 73, 1443–1452. https://doi.org/10.2136/sssaj2008.0021

van Genuchten, M.T., 1980. A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils1. Soil Sci. Soc. Am. J. 44, 892. https://doi.org/10.2136/sssaj1980.03615995004400050002x

Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M.G., Genuchten, M.T. van, 2010. Using Pedotransfer Functions to Estimate the van Genuchten–Mualem Soil Hydraulic Properties: A Review. Vadose Zo. J. 9, 795. https://doi.org/10.2136/vzj2010.0045

Wösten, J.H.M., Lilly, A., Nemes, A., Le Bas, C., 1999. Development and use of a database of hydraulic properties of European soils. Geoderma 90, 169–185. https://doi.org/10.1016/S0016-7061(98)00132-3

Wösten, J.H.M., Pachepsky, A., Rawls, J., 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. J. Hydrol. 251.

Wösten, J.H.M., van Genuchten, M.T., 1988. Using Texture and Other Soil Properties to Predict the Unsaturated Soil Hydraulic Functions. Soil Sci. Soc. Am. J. 52, 1762–1770. https://doi.org/10.2136/sssaj1988.03615995005200060045x

Zacharias, S., Wessolek, G., 2007. Excluding Organic Matter Content from Pedotransfer Predictors of Soil Water Retention. Soil Sci. Soc. Am. J. 71, 43–50. https://doi.org/10.2136/sssaj2006.0098

Zhang, Y., Schaap, M.G., 2017. Weighted recalibration of the Rosetta pedotransfer model with improved estimates of hydraulic parameter distributions and summary statistics (Rosetta3). J. Hydrol. 547, 39–53. https://doi.org/10.1016/j.jhydrol.2017.01.004

Chapter 3. Developing Pseudo Continuous Pedotransfer Functions for International Soils Measured with the Evaporation Method and the HYPROP System: II. The Soil Hydraulic Conductivity Curve

Abstract:

Direct measurement of unsaturated hydraulic parameters is costly and time-consuming. Pedotransfer functions (PTFs) are typically developed to estimate soil hydraulic properties from readily available soil attributes. For the first time, in this study, we developed PTFs to estimate the soil hydraulic conductivity (log(*K*)) directly from measured data. We adopted the pseudo continuous neural network PTF (PC_{NN}-PTF) approach and assessed its accuracy and reliability using two independent data sets with hydraulic conductivity measured via the evaporation method. The primary data set contained 150 international soils (6963 measured data pairs), and the second dataset consisted of 79 repacked Turkish soil samples (1340 measured data pairs). Four models with different combinations of the input attributes, including soil texture (sand, silt, clay), bulk density (BD), and organic matter content (SOM), were developed. The best performing international (root mean square error, RMSE = 0.520) and local (RMSE = 0.317) PTFs only had soil texture information as inputs when developed and tested using the same data set to estimate log(*K*). However, adding BD and SOM as input parameters increased the reliability of the international PC_{NN}-PTFs when the Turkish data set was used as the test data set. We observed an overall improvement in the performance of PTFs with the increasing number of data points per soil textural class. The PC_{NN}-PTFs consistently performed high across tension ranges when developed and tested using the international

data set. Incorporating the Turkish data set into PTF development substantially improved the accuracy of the PTFs (on average close to 60% reduction in RMSE). Consequently, we recommend integrating local HYPROP^{TM} (Hydraulic Property Analyzer, Meter Group Inc., USA) data sets into the international data set used in this study and retraining the PC_{NN}-PTFs to enhance their performance for that specific region.

3.1. Introduction

Direct measurements of soil hydraulic properties in the field and laboratory can be tedious, laborious, and often expensive due to their significant inherent spatial variability. Therefore, pedotransfer functions (PTFs) are often developed and used to indirectly estimate these properties by establishing empirical relationships based on the readily available soil properties such as soil texture, bulk density (BD), and soil organic matter content (SOM) (Bouma, 1989). The primary soil hydraulic properties include the soil water retention and hydraulic conductivity curves (SWRC and SHCC) that define the volumetric water content's nonlinear relationships with the soil tension and the soil hydraulic conductivity, respectively. The hydraulic conductivity decreases as the volumetric water content decreases because of a reduction in the cross-sectional area of water flow and increased tortuosity and drag forces (Assouline and Or, 2013; Vereecken et al., 2016).

The experimental determination of the SHCC is more complicated than the SWRC. Therefore, the SHCC is often derived from the SWRC and saturated hydraulic conductivity (K_s) information. A popular four-parameter expression developed by van Genuchten (1980) is widely used for SWRC parametrization, which coupled with Mualem-van Genuchten model (Mualem, 1976; van Genuchten, 1980) is often used for SHCC

parametrization using the K_s as a scaling factor. The SHCC can also be described by Gardner's empirical expression (Gardner, 1958), which in some cases works similarly or even better than the Mualem-van Genuchten model (Schaap and Leij, 1998).

The PTFs are mainly developed to only estimate K_s (point PTF) and parameters of the van Genuchten water retention model (parametric PTF), which are subsequently used for estimating the SHCC using the abovementioned approach (Børgesen et al., 2008; Parasuraman et al., 2006; Schaap et al., 1998; Weynants et al., 2009). For example, Schaap and Leij (2000) used PTF-based SWRC parameters of van Genuchten equation to estimate the SHCC. They observed that the best results are obtained when the parameters K_s and L (a term for the interaction between pore size and tortuosity) were flexible and not fixed as is the case in the classical Mualem–van Genuchten model. PTFs estimating unsaturated hydraulic conductivity at specific moisture tensions also exist (e.g., Moosavi and Sepaskhah (2012). However, little is known about the development and application of PTFs to directly estimate the SHCC (Wagner et al., 2001).

Multiple sources of error interact in a complicated manner when PTF-driven SWRC and K_s are used to estimate the SHCC. The first type of error is associated with estimating the parameters of the van Genuchten model and K_s using parametric and point PTFs, respectively. The second type of error is related to the Mualem–van Genuchten parameterization of the SWRC and SHCC, which is often fitted only using a few water retention data pairs measured by equilibrium approaches. The SHCC estimations via Mualem-van Genuchten model can result in poor performance near saturation because of the inability to account for water flow through macropores (Bormann and Klaassen, 2008;

Niemann and Rovey, 2009; Weynants et al., 2009). Furthermore, K_s is a highly variable soil hydraulic property dependent upon the pore geometry at the scale of interest (Niemann and Rovey, 2009) and seasonal variability (Bormann and Klaassen, 2008). Significant variabilities in K_s estimations might occur when using different PTFs modeling approaches (Baroni et al., 2010) and measurement techniques (Fodor et al., 2011), ultimately reflected in the SHCC estimations.

The pseudo-continuous PTF (PC-PTF) was introduced by Haghverdi et al. (2012) as a PTF development strategy for continuous estimation of the SWRC using machine learning approaches such as artificial neural networks (NN) and support vector machines (Haghverdi et al., 2014). Using high resolution measured data is recommended for developing robust PC-PTFs since PC-PTF learns the shape of the SWRC directly from the actual measured water retention data (Haghverdi et al., 2014).

Schindler and Müller (2017) published a soil hydraulic international dataset using the Evaporation method and HYPROP^{TM} (Hydraulic Property Analyzer, Meter Group Inc., USA) system. The HYPROP system (Figure 3.1) is an automated evaporation-based benchtop laboratory system that works based on the extended evaporation method (Schindler et al., 2010a, 2010b). The HYPROP has a relatively fast measurement cycle and provides high resolution reliable simultaneous measurements of soil water content and unsaturated hydraulic conductivity within a few days or weeks (Bezerra-Coelho et al., 2018; Peters et al., 2015; Peters and Durner, 2008). Haghverdi et al. (2018) used an HYPROP measured Turkish soil data set to develop water retention PC_{NN}-PTFs and reported promising results. In Chapter 2, we utilized the Schindler and Müller (2017)

dataset to develop water retention PC_{NN}-PTFs. However, no PC_{NN}-PTF has been developed to estimate the SHCC using high-resolution data measured via the evaporation method. Consequently, this study was carried out to (I) develop PC_{NN}-PTFs for SHCC estimations by utilizing the abovementioned international (Schindler and Müller, 2017) and Turkish (Haghverdi et al., 2018) data sets measured via the evaporation method, (II) determine the accuracy and reliability of the PC_{NN}-PTFs, and (III) assess the performance of the developed models across soil textures and different ranges of soil tension.

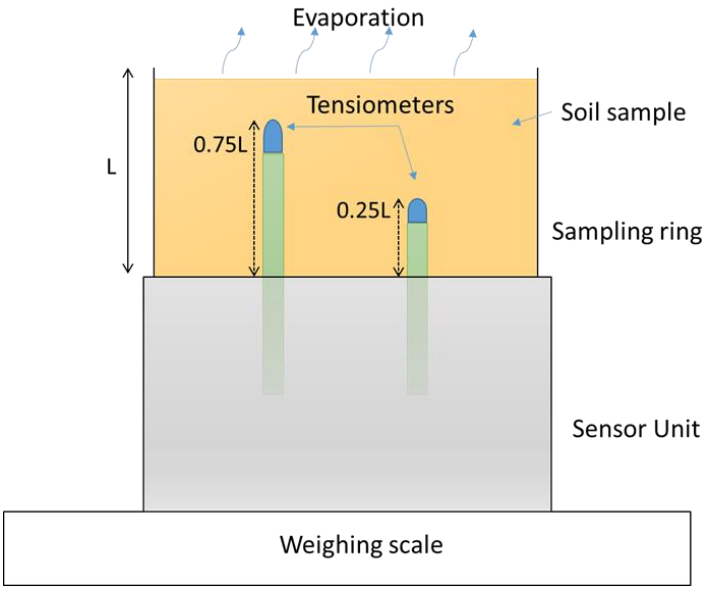


Figure 3.1. Experimental setup of the extended evaporation experiment using HYPROP system.

3.2. Materials and Methods

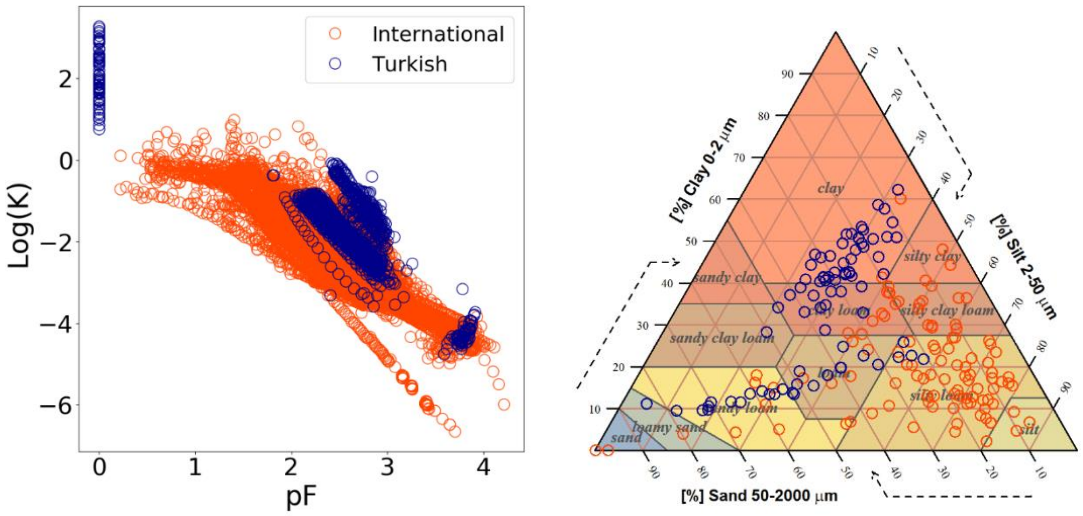
3.2.1. Soil Data Sets

In this study, two soil data sets were used to develop hydraulic conductivity PC_{NN}-PTFs and evaluate their accuracy and reliability. The measured hydraulic conductivity data and the soil textural classification of the samples for both data sets are shown in Figure 3.2.

The primary data set, hereafter referred to as the international data set, was published by Schindler and Müller (2017) and consisted of 173 soils collected from 71 sites from all over the world. This data set contains the measurements of water retention, unsaturated hydraulic conductivity, $K(h)$, and several basic soil properties, including textural data, organic matter content (SOM), and dry bulk density (BD). The soil hydraulic properties were measured using the evaporation experiments or the extended evaporation method via the HYPROP method. A majority of the soil samples in the data set were collected from arable lands, yet few samples from other land use types such as urban land, grassland, forests, fallow lands and riverbanks were also present. After screening the international data set, a subset of samples (i.e., 150 soils with 6963 total $K(h)$ data pairs) was selected to develop PC_{NN}-PTFs. The characteristics of the selected soils are shown in Table 3.1. The most dominant texture was silt loam; comprising 78 soil samples (52% of the data set), followed by loam; consisting of 18 soil samples (12% of the data set). Values of $K(h)$ were log-transformed because hydraulic conductivity data are generally log-normally distributed (Schaap and Leij, 2000). The measured $\log(K(h))$ values ranged from −6.64 to 0.98 (0 to 9.65 cm d^{-1}), with an average of −2.26 (0.073 cm d^{-1}). The pF (logarithmic transformation of soil tension in cm of water) values ranged from 0.22 to 4.21, with an average of 2.47.

The second data set (referred to as the Turkish data set herein) was mainly collected from areas surrounding Ankara, Turkey, and consisted of 79 repacked soil samples with 1340 $K(h)$ data pairs that were measured via the HYPROP system (Haghverdi et al., 2018). In this dataset, the K_s data (pF 0) were measured using the falling head method with the KSAT instrument (Meter Group Inc., Pullman, WA, USA). The $K(h)$ points were measured

for each sample with pF ranging from 1.80 to 3.91, with an average of 2.51, and log(*K*) ranging from −4.75 to 3.27 (0 to 1862 cm d^{-1}), with an average of −1.6 (0.03 cm d^{-1}). Clay was the dominant texture (38 soil samples or 48.1% of the data set), followed by sandy loam (13 samples or 16.5% of the data set). Further details about the laboratory procedures used to develop this data set are available in Haghverdi et al. (2020a, 2020b, 2018). More information about HYPROP's measurement principles is available in Schindler et al. (2016).



(**a**) (**b**)

Figure 3.2. The soil hydraulic conductivity and tension pairs (**a**), and soil textural distribution for the datasets (**b**). Dark orange circles depict the international data set (Schindler and Müller, 2017) and blue circles represent the Turkish dataset [21,28,30]. pF is the logarithmic transformation of soil tension in cm of water and *K* is the unsaturated hydraulic conductivity.

Table 3.1. Characteristics of soils from international and Turkish data sets used in this study to develop and test pseudo continuous neural network pedotransfer functions (PC_{NN}-PTFs).

	International Data (150 Soil Samples with 6963 Data Pairs)			**Turkish Data (79 Soil Samples with 1340 Data Pairs)**		
Attribute	Mean	Range	SD	Mean	Range	SD
Clay (%)	20.0	0.0–60.0	12.5	34.1	9.4–62.2	15.1
Silt (%)	56.4	0.2–86.8	17.1	30.7	5.2–57.6	8.7
Sand (%)	23.6	3.9–99.8	17.4	35.3	6.0–84.0	17.4
Bulk density (g cm^{-3})	1.3	0.6–1.7	0.2	1.0	0.7–1.3	0.1
Organic matter content (%)	3.1	0.0–12.0	2.5	1.2	0.0–3.1	0.6

SD: Standard deviation.

A statistical analysis conducted in Chapter 2 using Mahalanobis distance (Tranter et al., 2009) revealed that these two data sets were independent and most Turkish samples fall outside the domain of applicability of the international dataset.

3.2.2. Unsaturated Hydraulic Conductivity Calculations

During HYPROP measurements, saturated soil samples (closed from the base) were placed on a balance. Two tensiometers were positioned such that the tensiometers' tips were at depths of 0.25 L cm and 0.75 L cm, where L was the soil column height (which is typically 5 cm in laboratory evaporation experiments). The soil surface was open to the ambient atmosphere so that the soil water could evaporate. The medial pF value of the sample was calculated based on the average value of the two tensions measured by two tensiometers and corresponding water content was calculated based on the mass change of the soil sample.

The hydraulic conductivity was calculated using the water flow velocity (q_i[cm/d]) between time points t_{i-1} and t_i through a horizontal plane that laid exactly in the middle of the two tension-tips:

$$q_i = \frac{1}{2}\,\frac{(\Delta V_i/\Delta t_i)}{A} \tag{6}$$

where, ΔV_i is the change in water volume in the whole sample (cm^3), Δt_i is the time interval between two consecutive measurement points, and A the cross-sectional area (cm^2) of the column.

The data points for the hydraulic conductivity function were calculated by inverting Darcy's equation as:

$$K_i(h_i) = \frac{-q_i}{\{\left(\frac{\Delta h_i}{\Delta z}\right) - 1\}} \tag{7}$$

where, h_i (cm) is the time- and space-averaged tension, Δh_i is the difference of tensions between the two tensiometer tips, and Δz (cm) is the distance between the tensiometer tips. The calculations assume that moisture tension and water content distribute linearly through the column and, therefore, the arithmetic mean of the tensions at two points was used. This simplified assumption was shown to provide accurate results because linearity errors in fluxes and tensions cancel each other out (Peters et al., 2015). The effect of hysteresis on water flow and transport is well understood (Yang et al., 2014). However, since HYPROP measurements are taken during natural evaporation-based drying of soil samples, only drying hydraulic path was considered in this study.

3.2.3. PC_{NN}-PTFs Development

A three-layer feed-forward perceptron model was developed using MATLAB R2019a (Mathworks, 2019). The transfer functions were the "hyperbolic tangent sigmoid" and "linear" for the hidden and the output layers, respectively. The Levenberg–Marquardt algorithm (Marquardt, 1963) was used for training the models. The maximum epoch (one complete pass of the training data set through the learning process) was set to 1000 and the best weights were loaded automatically for testing.

Soil samples were randomly partitioned into 5 folds such that 80% of the data were used for the development of the PC_{NN}-PTF models and 20% for testing the models. The bootstrap technique was used on the development set to generate 100 replica datasets, each containing approximately 67% of the data. The rest of the development data (~33%) were used for cross-validation of the NN models. The training process was terminated when the root mean square error (RMSE) of the cross-validation subset began to increase or remain unchanged. To find the optimal topology of the neural network, the number of neurons of the hidden layer was iteratively changed from 1 to 14. This process was repeated five times leaving aside a different fold as the test set each time, such that all samples in the data set were used for testing the models. The outputs of the 100 PC_{NN}-PTFs with optimum topology were averaged to obtain the hydraulic conductivity estimations.

3.2.4. Modeling Scenarios

We evaluated the accuracy and reliability of the PC_{NN}-PTFs (developed using the international and the Turkish data sets) with four combination models of the input

attributes, including textural constituents—sand, silt, and clay (SSC), BD, and SOM (Table 3.2). Model 1 constituted all the input attributes and the logarithmic transformation of soil suction (pF). Model 2 included SSC and pF. Model 3 included SSC, BD, and pF. Model 4 included SSC, SOM, and pF. The log(*K* (cm/d)) was the output parameter corresponding to the input pF value.

Table 3.2. Combinations of input attributes used in this study to develop PC_{NN}-PTFs.

Model	Input Attributes
1	SSC, BD, SOM, pF
2	SSC, pF
3	SSC, BD, pF
4	SSC, SOM, pF

SSC: sand, silt, and clay percentages (%), BD: bulk density ($cm^3\ cm^{-3}$), SOM: soil organic matter content (%), pF: the logarithmic transformation of soil tension in cm of water.

Four data partitioning scenarios, as shown in Table 3.3, were considered when the international data set was used for training and testing (scenario 1), the Turkish data set was used for training and testing (scenario 2), the international data were used for training and the Turkish data for validation (scenario 3), and a combination of the two data sets was used for training and the Turkish dataset for testing (scenario 4). The accuracy of PTFs was assessed using a randomly selected subset of the development data set that was not used to derive the PTF. The reliability was evaluated based on the performance of PTFs on an independent data set beyond the statistical training limits and the geographical training area of the development dataset. For example, we estimated the log(*K*) of the Turkish soil samples to assess the reliability of the PC_{NN}-PTFs derived using the international data set. The results of the modeling scenarios were assessed to (i) quantify the improvements in

international PC_{NN}-PTFs for a specific region after incorporating local samples into the training data set and (ii) determining whether the international PC_{NN}-PTFs trained using the integrated data works as accurately as the local PC_{NN}-PTFs.

Table 3.3. Different data partitioning scenarios used in the study to train, test, and validate PC_{NN}-PTFs.

Scenario	Data Sets
S1	Training: International, Test: International.
S2	Training: Turkish, Test: Turkish.
S3	Training: International, Test: Turkish.
S4	Training: International + Turkish, Test: Turkish.

3.2.5. Model Evaluation

The root mean square error (*RMSE*, Equation (1)), mean absolute error (*MAE*, Equation (2)), mean bias error (*MBE*, Equation (3)), and correlation coefficient (*R*, Equation (4)) were calculated for the test data to evaluate the performance of PC_{NN}-PTFs:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(E_i - M_i)^2} \quad (3)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|E_i - M_i| \quad (4)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(E_i - M_i) \quad (5)$$

$$R = \frac{\sum_{i=1}^{n}(E_i - \bar{E})(M_i - \bar{M})}{\sum_{i=1}^{n}(E_i - \bar{E})^2 \sum_{i=1}^{n}(M_i - \bar{M})^2} \quad (6)$$

where, E and M are the estimated and measured log(K), respectively; $\bar{E}$ and $\bar{M}$ are the mean estimated and measured log(K), respectively; and n is the total number of measured water retention points for each model. In addition, the error statistics were calculated separately for dominant soil textures at the wet (pF ≤ 2), intermediate (2 < pF ≤ 3), and dry ranges (pF > 3) of the SHCC.

3.3. Results

3.3.1. Importance of the Input Predictors

Figure 3.3 shows the scatterplots of measured versus estimated log(K) values for the PC_{NN}-PTFs developed in this study using different combinations of input predictors. All models showed acceptable performance, demonstrated by the well-scattered data around the 1:1 reference line except for the K_s estimations in scenario 3 (training: the international dataset, test: Turkish datasets).

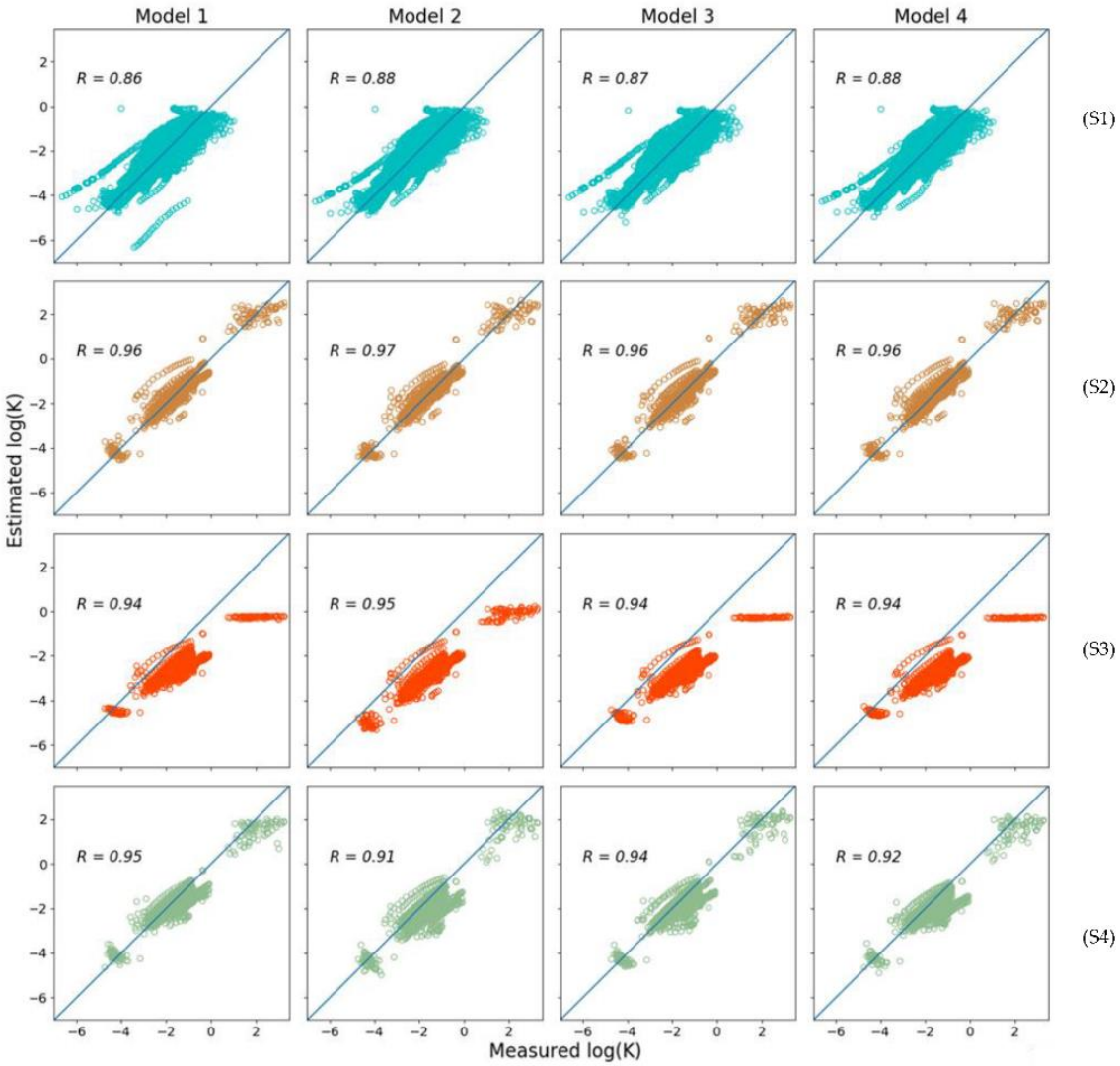


Figure 3.3. Scatterplots of measured versus estimated log(*K*) using PC_{NN}-PTFs. S1: training and test: the international dataset, S2: Training and test: Turkish dataset, S3: training: the international dataset, test: Turkish dataset, S4: training: international + Turkish dataset, test: Turkish dataset. Model 1 inputs: sand, silt and clay percentages (SSC), bulk density (BD), and soil organic matter content (SOM); Model 2 inputs: SSC; Model 3 inputs: SSC, BD; Model 4 inputs: SSC and SOM.

Table 3.4 summarizes the performance statistics of the models for all the scenarios. For scenario 1 (training: International, test: International), Model 2 (inputs: SSC, pF) resulted in the best performance with an RMSE of 0.520 and MAE of 0.406, followed by Model 4 (SSC, SOM, pF) where RMSE was 0.529 and MAE was 0.417. The lowest

performance was observed in model 1 where RMSE was 0.571 and MAE was 0.428. MBE varied from 0.013 to 0.033, demonstrating no substantial under or overestimation of log(*K*) for all models. The R values varied from 0.855 to 0.881, showing a high agreement between measured and estimated log(*K*) in all models.

Table 3.4. Performance of the PC_{NN}-PTFs estimating log-transformed soil hydraulic conductivity data (cm d−1) across four modeling scenarios.

	Training and Test: I				**Training and Test: T**				**Training: I.; Test: T**				**Training: I +T; Test: T**			
M	**RMSE**	**MAE**	**MBE**	**R**	**RMSE**	**MAE**	**MBE**	**R**	**RMSE**	**MAE**	**MBE**	**R**	**RMSE**	**MAE**	**MBE**	**R**
1	0.571	0.428	0.013	0.855	0.343	0.227	0.023	0.959	1.097	0.971	−0.959	0.935	0.429	0.312	−0.139	0.947
2	0.520	0.406	0.027	0.881	0.317	0.217	0.011	0.965	1.317	1.254	−1.249	0.954	0.613	0.456	−0.335	0.906
3	0.547	0.418	0.033	0.868	0.336	0.219	0.017	0.961	1.235	1.142	−1.133	0.942	0.453	0.308	−0.165	0.938
4	0.529	0.417	0.022	0.877	0.350	0.243	0.043	0.958	1.243	1.144	−1.132	0.943	0.554	0.400	−0.280	0.920

M: Model, RMSE: Root mean square error, MAE: mean absolute error, MBE: mean biased error, R: correlation coefficient. I: international data set, T: Turkish data set.

For scenario 2 (training: Turkish dataset, test: Turkish dataset), Model 2 (inputs: SSC, pF) resulted in the best performance with RMSE of 0.317 and MAE of 0.217, followed by Model 3 (inputs: SSC, BD) where RMSE was 0.336 and MAE was 0.219. MBE varied from 0.011 for model 2 to 0.043 for model 4, demonstrating no considerable under or overestimation of log(*K*). The R values were high (between 0.958 and 0.965) and similar among all models.

For scenario 3 (training: the international data set, test: the Turkish data set), model 1 with RMSE of 1.097 and MAE of 0.971 performed the best. Model 2 with RMSE of 1.317 and MAE of 1.254 showed lower accuracy compared to the other models. The estimated

log(*K*) values were highly correlated with the measured data (R: 0.935–0.954), yet all models showed an underestimation tendency (MBE ranging from −1.249 to −0.959). This is evident in Figure 3.3 as well, where data points are well scattered but located mainly below the 1:1 line.

For scenario 4 (training: combined international and Turkish data sets, test: the Turkish data set), the best performance was observed for Model 3 with RMSE of 0.453 and MAE of 0.308. Model 1 also had a similar performance. Slight underestimation of log(*K*) was observed with MBE ranging from −0.335 for model 2 to −0.139 for Model 1. Correlation between observed and estimated log(*K*) was high and similar among all models, with R values ranging from 0.906 to 0.947.

No distinct relationship was observed between BD and SOM with RMSE values except for the Turkish clay soils where RMSE declined as BD increased (Figure 3.4).

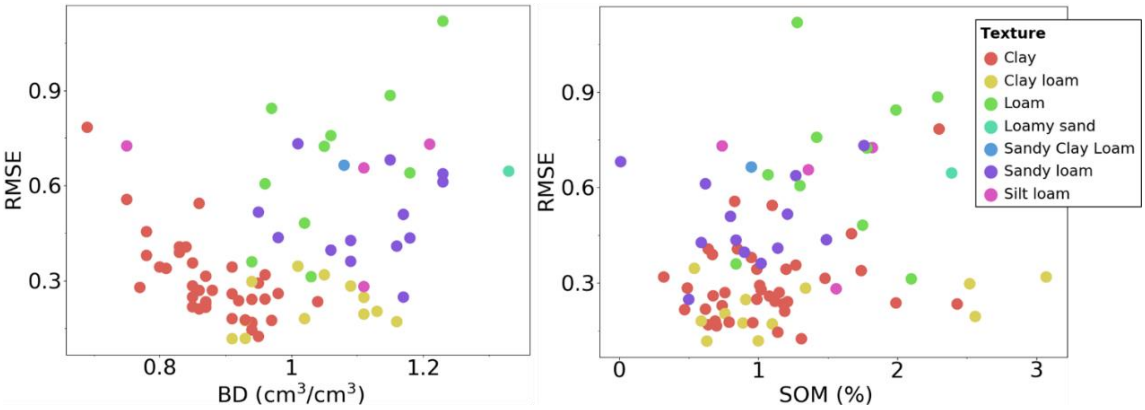


Figure 3.4. The root of mean squared error (RMSE) as a function of bulk density (BD) and organic matter content (SOM) for the $PC_{NN}PTF$ Model 1 with SSC, BD, SOM as inputs. The model was developed using combined international and Turkish data sets and tested using the Turkish data set (scenario 4). The error was calculated for each soil sample separately.

3.3.2. Performance across Soil Textures

The following analysis was only conducted using model 1 as the best performing PTF in the test phase. Table 3.5 shows the performance of the PC_{NN}-PTF models for the dominant soil textures, representing about 89% and 92% of the international and Turkish data sets, respectively. When the international data set was used as the training set (scenario 1), clay loam had higher RMSE and MAE values than other soil textures. RMSE values ranged from 0.517 to 1.124, MAE values ranged from 0.342 to 0.748, and MBE values ranged from 0.026 to 0.288 for all textures. Furthermore, the model showed a tendency to overestimate log(*K*) for all soil textures, except loam, where underestimation of log(*K*) was observed. The correlation coefficient (R) values varied between 0.603 in clay loam to 0.881 for silt loam.

When only Turkish data were used for training (scenario 2), RMSE and MAE values varied from 0.206 to 0.395 and 0.146 to 0.312, respectively. MBE values ranged from −0.096 for sandy loam to 0.018 for clay loam, and no substantial underestimation or overestimation of log(*K*) was observed. The agreement between the measured and estimated log(*K*) values was very high, indicated by high and similar R values (between 0.926 and 0.982) for all the models within each soil texture.

When the international data set was used for training and the Turkish data set for validation (scenario 3), RMSE and MAE values varied from 0.964 to 1.444 and 0.863 to 1.377, respectively. Underestimation of log(*K*) was observed for all the soil textures with MBE values ranging from −1.370 to −0.860. Loam had the highest error relative to other

soil textures, while clay had the lowest. High and similar correlation coefficient values (between 0.917 and 0.978) were observed for all the models and soil textures.

When the Turkish data set was used as a test and a combination of international and Turkish data sets were used for the training (scenario 4), the RMSE and MAE values varied from 0.230 to 0.745 and 0.173 to 0.683, respectively. The loam and sandy loam had higher RMSE and MAE values, while the errors for clay loam and clay were similar to when just the Turkish data were used for training (scenario 2). The MBE values ranged from −0.614 to −0.013, showing slight underestimation of log (K) for most soil textures except clay and clay loam where underestimation was not substantial. The agreement between the estimated and observed log(K) was high, as depicted by the high R values (ranging from 0.915 to 0.985) for all the models.

Table 3.5. Performance of PC-PTFs on main textural classes of the international and Turkish data sets for estimating log(K).

Data sets	Texture	RMSE	MAE	MBE	R
Training: I; Test: I	**SiL**	0.517	0.421	0.026	0.881
	L	0.612	0.485	−0.180	0.789
	SiCL	0.433	0.342	0.144	0.870
	CL	1.124	0.748	0.288	0.603
	SL	0.593	0.522	0.042	0.700
Training: T, Test: T	**C**	0.252	0.174	0.006	0.978
	SL	0.366	0.277	−0.096	0.966
	CL	0.206	0.146	0.018	0.982
	L	0.395	0.312	−0.048	0.926
Training: I, validation: T	**C**	0.986	0.863	−0.860	0.961
	SL	1.353	1.241	−1.223	0.96
	CL	0.964	0.894	−0.894	0.978
	L	1.444	1.377	−1.370	0.917
Training: I and T, Test: T	**C**	0.303	0.218	−0.013	0.972
	SL	0.535	0.446	−0.317	0.963
	CL	0.230	0.173	−0.094	0.985
	L	0.745	0.683	−0.614	0.915

RMSE: Root mean square error, MAE: mean absolute error, MBE: mean biased error, R: correlation coefficient. I: international data set, T: Turkish data set. SiL: Silt Loam: L: Loam, SiCL: Silty Clay Loam, CL: Clay loam, SL: Sandy Loam, C: Clay.

3.3.3. Performance at the Wet, Intermediate and Dry Parts of the SHCC

Table 3.6 shows the performance of different PC_{NN}-PTFs over three moisture ranges of the SHCC for the four data partitioning scenarios evaluated in this study. When the international data set was used for the training and testing of the models (scenario 1), the RMSE of Model 1 varied from 0.548 in the wet range to 0.603 in the dry range. The MAE

values varied from 0.420 in the wet range to 0.440 in the intermediate range of the SHCC. The MBE values varied between −0.060 in the wet range to 0.140 in the dry range. The R values ranged from 0.509 for the wet to 0.640 in the intermediate range.

Table 3.6. Performance of the PC_{NN}-PTFs (inputs: SSC, BD, OM, and pF) developed to estimate the log(*K*) at wet (pF ≤ 2) intermediate (2 < pF ≤ 3) and dry (pF > 3) parts of the SHCC.

	Training and Test: I			**Training and Test: T**			**Training: I, Validation: T**			**Training: I and T, Test: T**		
	Wet	**Mid**	**Dry**	**Wet**	**Mid**	**Dry**	**Wet**	**Mid**	**Dry**	**Wet**	**Mid**	**Dry**
RMSE	0.548	0.570	0.603	0.588	0.317	0.375	2.285	1.002	0.466	0.757	0.400	0.396
MAE	0.420	0.440	0.381	0.471	0.206	0.298	2.158	0.936	0.342	0.602	0.291	0.322
MBE	−0.060	0.007	0.140	0.031	0.016	0.109	−2.158	−0.926	−0.292	−0.520	−0.134	0.154
R	0.509	0.640	0.522	0.809	0.860	0.831	0.768	0.785	0.860	0.805	0.791	0.818

RMSE: Root mean square error, MAE: mean absolute error, MBE: mean biased error, R: correlation coefficient. I: International data set, T: Turkish data set.

When only Turkish data were used for the training and testing of the PC_{NN}-PTF models (scenario 2), the lowest error was observed in the intermediate range (RMSE = 0.317, MAE = 0.206, and MBE = 0.016) and the highest error belonged to the wet range (RMSE = 0.588, MAE = 0.471, and MBE = 0.031). The agreement between the observed and estimated log(*K*) was comparable among models with R ranging from a minimum of 0.809 in the wet range to a maximum of 0.860 in the intermediate range.

When the Turkish data were used as the validation data set (scenario 3), the PTFs showed their highest performance in the dry range (RMSE = 0.466 and MAE = 0.342) and their lowest performance in the wet range (RMSE = 2.285 and MAE = 2.158). Despite high R values (0.768–0.860), a tendency to underestimate log(*K*) was observed in all regions, as indicated by negative MBE ranging from −0.292 to −2.158.

When both international and Turkish data sets were used for the training of the models (scenario 4), the lowest error values were observed in the dry range (RMSE = 0.396 and MAE = 0.322) and the highest error in the wet range (RMSE = 0.757 and MAE = 0.602). RMSE values varied from 0.396 in the dry range to 0.757 in the wet range. Negative MBE values of −0.52 and −0.134 indicated a tendency to underestimate log(*K*) in the wet and intermediate ranges, respectively. The R values were high across all soil tension ranges.

3.4. Discussion

3.4.1. Accuracy and Reliability of the Developed PTFs

In our study, the best international hydraulic conductivity PC_{NN}-PTF showed the accuracy (same data set for development and test) and reliability (independent data sets for development and test) of RMSE = 0.520 and 1.097, respectively. The local PC_{NN}-PTF developed and tested using the Turkish dataset showed even higher performance, as expected, with an RMSE of 0.317. Parasuraman et al. (2006) stated that better performance in estimating K_s is observed when a NN model is trained even on a small set of relevant data rather than a larger general dataset. Our study emphasizes that a local data set, when available, should be included in the training of PC_{NN}-PTF for a more accurate estimation of the SHCC.

Schaap and Leij (1998) reported RMSE values ranging from 1.12 to 1.76 for calibration subset and from 1.18 to 1.77 for an independent validation data set for their hydraulic conductivity PTFs, indicating lower accuracy and reliability than the PC_{NN}-PTF developed in this study. Børgesen et al. (2008) reported a reasonable accuracy for their

hydraulic conductivity PTFs with RMSE ranging from 0.598 to 1.196, yet most of the models showed underestimation. The above mentioned studies used the typical procedure to estimate the SHCC, which relies on estimated or measured K_s values and parametric SWRC. Therefore, the PC_{NN}-PTFs approach developed and tested for the first time in this study could be used as an alternative high-performance approach to estimate the SHCC.

Large data sets being used to develop international PTFs typically consist of smaller data sets, employing different techniques to measure soil hydraulic properties. The commonly used devices have discrepancies in K_s measurements because of factors such as sample size, soil conditions, flow geometry and installation procedures (Morbidelli et al., 2017). The same is true for the various devices used for unsaturated hydraulic conductivity measurements such as the steady-state pressure membrane method, tension disc infiltrometer, hot-air methods, and the widely used multistep outflow method (Benson and Gribb, 1997; Durner and Lipsius, 2005; Stolte et al., 1994). We recommend using HYPROP data sets for developing hydraulic conductivity PC_{NN}-PTFs. PC_{NN}-PTF takes advantage of the high resolution measured data provided by the HYPROP system to learn the shape of the SHCC directly from the actual measured data points, unlike the parametric PTFs where the relationships between the parameters and their predictors have to be known a priori. Furthermore, using only one method (evaporation experiment) for obtaining hydraulic conductivity data in the laboratory is expected to improve the performance of the PC_{NN}-PTFs by eliminating the variance related to employing multiple measurement techniques.

3.4.2. Importance of Input Variables

Studies have shown that the systematic variation in K_s is explained by properties like soil texture, porosity, SOM, and BD (Cosby et al., 1984; Wösten et al., 1999). According to Zhang and Schaap (2019), adding BD and SOM as input predictors improved the performance of PTFs in most studies estimating the K_s. Moosavi and Sepaskhah (2012) observed that the combination of inputs SSC, BD, SOM (Model 1 in this study) and SSC (Model 2 in this study) produced the best accuracy in estimating unsaturated hydraulic conductivity. In our study, considering SOM and BD as extra input attributes in addition to soil texture did not improve the accuracy of the international and local PTFs (scenarios 1 and 2). However, adding BD and SOM as input attributes noticeably enhanced the performance of the PC_{NN}-PTF in scenarios 3 and 4. This result differs from our observation in Chapter 2, where adding SOM as an extra input did not improve the performance of the water retention PC_{NN}-PTFs.

Except for scenario 1, BD was a more effective additional attribute than SOM in all modeling scenarios. Improvements in the estimation of log(K) were observed in other studies too when BD was included as an additional PTF input predictor (Minasny et al., 2004; Schaap and Leij, 1998). The BD, however, only provides limited information about soil structure as different preferential flow pathways may result in substantially different soil hydraulic conductivities (Romero-Ruiz et al., 2018). Hao et al. (2019) reported that K_s is influenced primarily by porosity and macro water-stable aggregates, which are not among typical PTF inputs. Further studies are needed to determine the impact of

considering additional soil structural input parameters on the performance of hydraulic conductivity PC_{NN}-PTF.

3.4.3. Performance across Textural Classes and Tension Ranges

The PC_{NN}-PTFs showed better performance for fine-textured (Clay and Clay Loam) than more coarse-textured (Loam and Sandy Loam) Turkish soils (scenarios 2, 3, and 4), which is attributed to a relatively higher number of fine-textured Turkish soil samples. We observed an overall improvement in the performance of PTFs with the increasing number of data points per soil textural class (Figure 3.5). These results concur with the results we reported in Chapter 2, where high performance was observed for the dominant soil textures for the SWRC estimations. Since PC_{NN}-PTFs are machine learning-based models, their performance is expected to improve as more data become available for training.

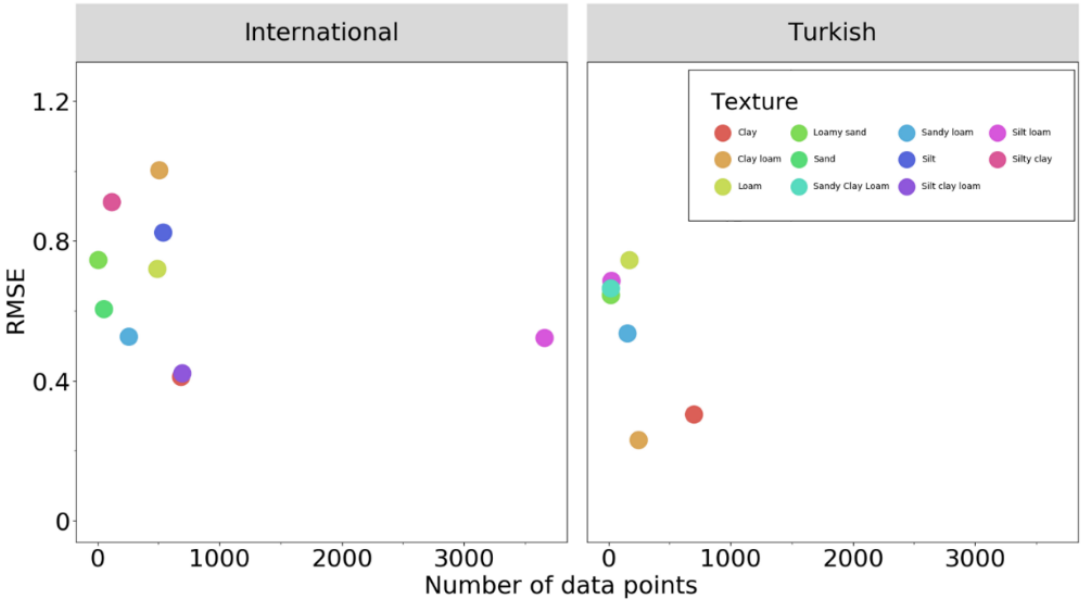


Figure 3.5. The root of mean squared error (RMSE) as a function of the number of measured hydraulic conductivity data pairs for each textural class for the PC_{NN}-PTF Model 1 with SSC, BD, SOM as inputs. The model was developed using combined international and Turkish data sets.

The performance of PC_{NN}-PTFs was consistent across tension ranges when developed and tested using the international data set, which concurs with the results observed in the Chapter 2. Moosavi and Sepaskhah (2012) reported a relatively lower accuracy of the NN-based PTFs to estimate hydraulic conductivity at the saturated and/or near-saturated tensions. We observed a somewhat higher error in the wet tension region ($pF \leq 2$) for the Turkish data set, primarily when PC_{NN}-PTFs were developed using the international data set. In Chapter 2, however, the performance of water retention PC_{NN}-PTFs was similar in three tension regions. The relatively higher error in the wet range in this study was because the Turkish data set only contained K_s data in the wet part (measured via the KSAT instrument), while K_s measurements were not available for the international data set.

3.5. Conclusions

We developed and evaluated PC_{NN}-PTFs to estimate the SHCC measured using the evaporation experiments, mainly via the HYPROP system. The PC_{NN}-PTF approach showed promising performance for continuous hydraulic conductivity estimation over a wide range of soil tensions. The HYPROP system offers the advantage of producing high-resolution soil hydraulic conductivity data over a wide range of soil tensions ($pF = 1.5$ to 3.5), which is critical for developing robust PC_{NN}-PTF models since this approach learns the shape of the SHCC directly from measured data. The KSAT instrument can be employed to measure the saturated hydraulic conductivity (K_s) that can be used along with HYPROP data. The water retention PC_{NN}-PTFs developed and validated in the second chapter also performed very well. Consequently, we recommend the PC_{NN}-PTF approach

to derive the next generation of water retention and hydraulic conductivity models using high-resolution data measured via the HYPROP system.

3.6. References

Assouline, S., Or, D., 2013. Conceptual and Parametric Representation of Soil Hydraulic Properties: A Review. Vadose Zo. J. 12. https://doi.org/10.2136/vzj2013.07.0121

Baroni, G., Facchi, A., Gandolfi, C., Ortuani, B., Horeschi, D., Van Dam, J.C., 2010. Uncertainty in the determination of soil hydraulic parameters and its influence on the performance of two hydrological models of different complexity. Hydrol. Earth Syst. Sci. 14, 251–270. https://doi.org/10.5194/hess-14-251-2010

Benson, C.H., Gribb, M.M., 1997. Measuring unsaturated hydraulic conductivity in the laboratory and field. Geotech. Spec. Publ. 113–168.

Bezerra-Coelho, C.R., Zhuang, L., Barbosa, M.C., Soto, M.A., Van Genuchten, M.T., 2018. Further tests of the HYPROP evaporation method for estimating the unsaturated soil hydraulic properties. J. Hydrol. Hydromechanics 66, 161–169. https://doi.org/10.1515/johh-2017-0046

Børgesen, C.D., Iversen, B. V., Jacobsen, O.H., Schaap, M.G., 2008. Pedotransfer functions estimating soil hydraulic properties using different soil parameters. Hydrol. Process. 22, 1630–1639. https://doi.org/10.1002/hyp.6731

Bormann, H., Klaassen, K., 2008. Seasonal and land use dependent variability of soil hydraulic and soil hydrological properties of two Northern German soils. Geoderma 145, 295–302. https://doi.org/10.1016/j.geoderma.2008.03.017

Bouma, J., 1989. Using Soil Survey Data for Quantitative Land Evaluation., in: Stewart, B.. (Ed.), Advances in Soil Science, Vol 9. Springer, New York, NY. https://doi.org/https://doi.org/10.1007/978-1-4612-3532-3

Cosby, B.J., Hornberger, G.M., Clapp, R.B., Ginn, T.R., 1984. A Statistical Exploration of the Relationships of Soil Moisture Characteristics to the Physical Properties of Soils. Water Resour. Res. 20, 682–690. https://doi.org/10.1029/WR020i006p00682

Durner, W., Lipsius, K., 2005. Determining Soil Hydraulic Properties, in: Anderson, M.G., Mcdonnell, J.. (Eds.), Encyclopedia of Hydrological Sciences. John Wiley & Sons, Ltd, Chichester, UK, pp. 1121–1144. https://doi.org/10.1002/0470848944.hsa077b

Fodor, N., Sándor, R., Orfanus, T., Lichner, L., Rajkai, K., 2011. Evaluation method dependency of measured saturated hydraulic conductivity. Geoderma 165, 60–68. https://doi.org/10.1016/j.geoderma.2011.07.004

Gardner, W.R., 1958. Some steady state solutions of the unsaturated moisture flow equation with application to evaporation from a water table. Soil Sci. 85, 228–232.

Haghverdi, A., Cornelis, W.M., Ghahraman, B., 2012. A pseudo-continuous neural

network approach for developing water retention pedotransfer functions with limited data. J. Hydrol. 442–443, 46–54. https://doi.org/10.1016/j.jhydrol.2012.03.036

Haghverdi, A., Najarchi, M., Öztürk, H.S., Durner, W., 2020a. Studying Unimodal, Bimodal, PDI and Bimodal-PDI Variants of Multiple Soil Water Retention Models: I. Direct Model Fit Using the Extended Evaporation and Dewpoint Methods. Water 12, 16–23. https://doi.org/10.3390/w12030900

Haghverdi, A., Öztürk, H.S., Cornelis, W.M., 2014. Revisiting the pseudo continuous pedotransfer function concept: Impact of data quality and data mining method. Geoderma 226–227, 31–38. https://doi.org/10.1016/j.geoderma.2014.02.026

Haghverdi, A., Öztürk, H.S., Durner, W., 2020b. Studying Unimodal, Bimodal, PDI and Bimodal-PDI Variants of Multiple Soil Water Retention Models: II. Evaluation of Parametric Pedotransfer Functions Against Direct Fits. Water 12. https://doi.org/10.3390/w12030896

Haghverdi, A., Öztürk, H.S., Durner, W., 2018. Measurement and estimation of the soil water retention curve using the evaporation method and the pseudo continuous pedotransfer function. J. Hydrol. 563, 251–259. https://doi.org/10.1016/j.jhydrol.2018.06.007

Hao, M., Zhang, J., Meng, M., Chen, H.Y.H., Guo, X., Liu, S., Ye, L., 2019. Impacts of changes in vegetation on saturated hydraulic conductivity of soil in subtropical forests. Sci. Rep. 9, 1–9. https://doi.org/10.1038/s41598-019-44921-w

Marquardt, D.W., 1963. An Algorithm for Leat-Squares Estimation of Nnonlinear Parameters. J. Soc. Indust. Appl. Math. 11, 431–441.

Minasny, B., Hopmans, J.W., Harter, T., Eching, S.O., Tuli, A., Denton, M.A., 2004. Neural Networks Prediction of Soil Hydraulic Functions for Alluvial Soils Using Multistep Outflow Data. Soil Sci. Soc. Am. J. 68, 417–429. https://doi.org/10.2136/sssaj2004.4170

Moosavi, A.A., Sepaskhah, A., 2012. Artificial neural networks for predicting unsaturated soil hydraulic characteristics at different applied tensions. Arch. Agron. Soil Sci. 58, 125–153. https://doi.org/10.1080/03650340.2010.512289

Morbidelli, R., Saltalippi, C., Flammini, A., Cifrodelli, M., Picciafuoco, T., Corradini, C., Govindaraju, R.S., 2017. In situ measurements of soil saturated hydraulic conductivity: Assessment of reliability through rainfall–runoff experiments. Hydrol. Process. 31, 3084–3094. https://doi.org/10.1002/hyp.11247

Mualem, Y., 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. Water Resour. Res. 12, 513–522. https://doi.org/10.1029/WR012i003p00513

Niemann, W.L., Rovey, C.W., 2009. A systematic field-based testing program of hydraulic conductivity and dispersivity over a range in scale. Hydrogeol. J. 17, 307–320. https://doi.org/10.1007/s10040-008-0365-3

Parasuraman, K., Elshorbagy, A., Si, B.C., 2006. Estimating Saturated Hydraulic Conductivity In Spatially Variable Fields Using Neural Network Ensembles. Soil Sci. Soc. Am. J. 70, 1851–1859. https://doi.org/10.2136/sssaj2006.0045

Peters, A., Durner, W., 2008. Simplified evaporation method for determining soil hydraulic properties. J. Hydrol. 356, 147–162. https://doi.org/10.1016/j.jhydrol.2008.04.016

Peters, A., Iden, S.C., Durner, W., 2015. Revisiting the simplified evaporation method: Identification of hydraulic functions considering vapor, film and corner flow. J. Hydrol. 527, 531–542. https://doi.org/10.1016/j.jhydrol.2015.05.020

Romero-Ruiz, A., Linde, N., Keller, T., Or, D., 2018. A Review of Geophysical Methods for Soil Structure Characterization. Rev. Geophys. 56, 672–697. https://doi.org/10.1029/2018RG000611

Schaap, M.G., Leij, F.J., 2000. Improved Prediction of Unsaturated Hydraulic Conductivity with the Mualem-van Genuchten Model. Soil Sci. Soc. Am. J. 64, 843–851.

Schaap, M.G., Leij, F.J., 1998. Using neural networks to predict soil water retention and soil hydraulic conductivity. Soil Tillage Res. 47, 37–42. https://doi.org/10.1016/S0167-1987(98)00070-1

Schaap, M.G., Leij, F.J., van Genuchten, M.T., 1998. Neural Network Analysis for Hierarchical Prediction of Soil Hydraulic Properties. Soil Sci. Soc. Am. J. 62, 847. https://doi.org/10.2136/sssaj1998.03615995006200040001x

Schindler, U., Durner, W., von Unold, G., Mueller, L., Wieland, R., 2010a. The evaporation method: Extending the measurement range of soil hydraulic properties using the air-entry pressure of the ceramic cup. J. Plant Nutr. Soil Sci. 173, 563–572. https://doi.org/10.1002/jpln.200900201

Schindler, U., Durner, W., von Unold, G., Müller, L., 2010b. Evaporation Method for Measuring Unsaturated Hydraulic Properties of Soils: Extending the Measurement Range. Soil Sci. Soc. Am. J. 74, 1071. https://doi.org/10.2136/sssaj2008.0358

Schindler, U., Mueller, L., von Unold, G., Durner, W., Fank, J., 2016. Emerging Measurement Methods for Soil Hydrological Studies, in: Novel Methods for Monitoring and Managing Land and Water Resources in Siberia. Springer, Cham, pp. 345–363. https://doi.org/10.1007/978-3-319-24409-9_14

Schindler, U., Müller, L., 2017. Soil hydraulic functions of international soils measured with the Extended Evaporation Method (EEM) and the HYPROP device. Open Data J. Agric. Res. 3, 10–16.

Stolte, J., Freijer, J.I., Bouten, W., Dirksen, C., Halbertsma, J.M., Van Dam, J.C., Van den Berg, J.A., Veerman, G.J., Wösten, J.H.M., 1994. Comparison of Six Methods To Determine Unsaturated Soil Hydraulic Conductivity. Soil Sci. Soc. Am. J. 58, 1596–1603. https://doi.org/https://doi.org/10.2136/sssaj1994.03615995005800060002x

Tranter, G., McBratney, A.B., Minasny, B., 2009. Using distance metrics to determine the

appropriate domain of pedotransfer function predictions. Geoderma 149, 421–425. https://doi.org/10.1016/j.geoderma.2009.01.006

van Genuchten, M.T., 1980. A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils1. Soil Sci. Soc. Am. J. 44, 892. https://doi.org/10.2136/sssaj1980.03615995004400050002x

Vereecken, H., Schnepf, A., Hopmans, J.W., Javaux, M., Or, D., Roose, T., Vanderborght, J., Young, M.H., Amelung, W., Aitkenhead, M., Allison, S.D., Assouline, S., Baveye, P., Berli, M., Brüggemann, N., Finke, P., Flury, M., Gaiser, T., Govers, G., Ghezzehei, T., Hallett, P., Hendricks Franssen, H.J., Heppell, J., Horn, R., Huisman, J.A., Jacques, D., Jonard, F., Kollet, S., Lafolie, F., Lamorski, K., Leitner, D., McBratney, A., Minasny, B., Montzka, C., Nowak, W., Pachepsky, Y., Padarian, J., Romano, N., Roth, K., Rothfuss, Y., Rowe, E.C., Schwen, A., Šimůnek, J., Tiktak, A., Van Dam, J., van der Zee, S.E.A.T.M., Vogel, H.J., Vrugt, J.A., Wöhling, T., Young, I.M., 2016. Modeling Soil Processes: Review, Key Challenges, and New Perspectives. Vadose Zo. J. 15, vzj2015.09.0131. https://doi.org/10.2136/vzj2015.09.0131

Wagner, B., Tarnawski, V.R., Hennings, V., Muller, U., 2001. Evaluation of pedo-transfer functions for unsaturated soil hydraulic conductivity using an independent data set.

Weynants, M., Vereecken, H., Javaux, M., 2009. Revisiting Vereecken Pedotransfer Functions: Introducing a Closed-Form Hydraulic Model. Vadose Zo. J. 8, 86–95. https://doi.org/10.2136/vzj2008.0062

Wösten, J.H.M., Lilly, A., Nemes, A., Le Bas, C., 1999. Development and use of a database of hydraulic properties of European soils. Geoderma 90, 169–185. https://doi.org/10.1016/S0016-7061(98)00132-3

Yang, H., Khoshghalb, A., Russell, A.R., 2014. Fractal-based estimation of hydraulic conductivity from soil-water characteristic curves considering hysteresis. Geotech. Lett. 4, 1–10. https://doi.org/10.1680/geolett.13.00071

Zhang, Y., Schaap, M.G., 2019. Estimation of saturated hydraulic conductivity with pedotransfer functions: A review. J. Hydrol. 575, 1011–1030. https://doi.org/10.1016/j.jhydrol.2019.05.058

Chapter 4. Evaluating the long term performance of empirical temperature-based and artificial neural network models for estimating reference evapotranspiration in California

Abstract

Efficient urban landscape irrigation management is critical in California and depends on the reliable estimation of reference evapotranspiration (ET_o). The ET-based smart irrigation controllers used for landscape irrigation often rely on temperature-based ET_o models; thus, a comprehensive evaluation of these models across climate regions is required in California. This study evaluated eight temperature-based empirical ET_o models, including Blaney and Criddle (BC), FAO-24 BC (BC_{fao}), Hamon (H), Hargreaves (H_s), Hargreaves and Samani (HS_a), Jenson and Haise (JH), Kharuffa (K), and Linacre (L) models. In addition, four artificial neural network models (ANN) were developed using raw weather data as inputs and the reconstructed signal obtained from wavelet transform as inputs. A total of 101 active California Irrigation Management Information System (CIMIS) weather stations were selected for this study, with more than 725,000 data points expanding from 1985 to 2019. The performance of the models was assessed against the standard CIMIS ET_o. The ranking of the empirical temperature-based models from best to worst performing was H_s > HS_a >BC_{fao} > JH > BC >K > H > L. The ANN model with temperature, solar radiation, relative humidity, wind speed information as input variables showed the best overall performance with RMSE equal to 0.51 mm d^{-1}. The H_s model with root mean square error (RMSE) of 0.61 mm d^{-1} was the most accurate empirical model, followed by the HS_a model with RMSE of 0.78 mm d^{-1}. The ANN model outperformed the

HS_a model using the same input variables (i.e., air temperature and extraterrestrial solar radiation) with 11% lesser RMSE. No improvement in the performance of ANN models was observed by using reconstructed signals obtained from the wavelet transform. Our results suggest using calibrated HS_a and ANN models that only require air temperature measurements for obtaining accurate estimations of ET_o in data-scarce conditions in urban areas across climate regions in CA.

4.1. Introduction

The FAO-56 Penman-Monteith is the accepted standard method to estimate the reference evapotranspiration, ET_o (Allen et al., 2005). It accounts for both energy and mass transfer processes to provide a reliable estimation of ET_o. However, it requires a wide range of accurate meteorological data such as temperature, humidity, solar radiation, and wind speed. Therefore, its implementation in data scare situations such as landscape irrigation management in urban areas is challenging and limited. Promising results and significant water savings have been reported for landscape irrigation using smart evapotranspiration-based controllers with on-site weather sensors (Cardenas et al., 2021; Serena et al., 2020). These irrigation products often employ temperature-based ET_o equations in their scheduling algorithms because of the high cost to install and maintain a complete weather station (Davis and Dukes, 2010). Furthermore, for smart controllers that rely on interpolated ET_o for irrigation scheduling, air temperature data is widely available and spatially interpolated more accurately than other weather parameters (Temesgen et al., 2005).

Various studies have evaluated the temperature-based ET_o models (Djaman et al., 2015; Hope and Evans, 1993; Liu et al., 2017; Muhammad et al., 2019; Muniandy et al., 2016; Tabari et al., 2013; Xu and Singh, 2001). In California, data from CIMIS (California Irrigation Management Information System) stations have been widely used to evaluate the suitability of empirical temperature-based ET_o equations (Hargreaves & Allen, 2003; Hope & Evans, 1993; Temesgen et al., 2005). Our recent multiyear turfgrass field irrigation research trials in Southern and Central California revealed accurate estimation of the ET_o by Weathermatic SL4800 smart controller using Hargreaves and Samani equation (Hargreaves and Samani, 1985) compared to CIMIS ET_o (Haghverdi et al., 2021b, 2021a). However, the performance of ET_o estimation methods varies with climate and data availability (Djaman et al., 2015), and the data requirements vary among models. Therefore, it is vital to evaluate various temperature-based ET_o measurement methods and develop regional/site-specific calibration equations, which can help users determine the best approach depending on the availability of data and the climate conditions (Kukal et al., 2020; Long et al., 2013).

Machine learning methods have also been proposed to estimate ET_o based on temperature and solar radiation data (Fan et al., 2018; Kisi and Alizamir, 2018). Algorithms such as artificial neural networks (ANNs) and combination of wavelet analysis/transform (WA) with ANNs, referred to as WA-ANNs, have been explored in various hydrological studies (Adamala, 2018; Adamowski & Sun, 2010; Chen et al., 2020; Evrendilek, 2014; Falamarzi et al., 2014; Kisi & Alizamir, 2018; Partal, 2009; Traore et al., 2016; Traore et al., 2010). The WA-ANN was introduced by Zhang & Benveniste (1992) as an alternative

approach to feed-forward ANNs to denoise input data. The wavelets can also be used as a substitute to activation functions in feed-forward ANNs, referred to as WNNs (Alexandridis and Zapranis, 2013). CIMIS data has been used in multiple studies for ET_o estimations using ANN and WA-based models (Cobaner, 2013; Kişi, 2010; Partal, 2009). However, the reliability of the ANN models to estimate ET_o at new locations was seldom evaluated, and models were typically developed and tested using ET_o data obtained from a limited number of sites.

This study used long-term daily CIMIS ET_o data across the state of California to (I) evaluate the performance of eight temperature-based empirical ET_o models, (II) develop calibration equations for the empirical temperature-based models for all climate divisions in CA, (III) develop and evaluate the accuracy of ANN and WA-ANN ET_o models using different sets of input data and (IV) determine the generalization ability of the machine learning-based ET_o models developed in this study.

4.2. Material and methods

4.2.1. Study Region and Data Sources

This study was carried out in California using the data obtained from automated CIMIS weather stations managed by the California Department of Water Resources (https://cimis.water.ca.gov/). In general, ET_o rates peak during the summer months and are low during the winter under California's Mediterranean climate, while precipitation is restricted mainly to winter and spring. A total of 101 active CIMIS stations with available data ranging from 1985 to 2019 were selected. Around 50% of the stations had 25 to 35 years of data, and the remaining stations had at least ten years of data to ensure a wide

range of weather conditions and drought events were considered. Figure 4.1 shows the locations of the selected CIMIS stations within the California aridity index map. The distribution of the weather parameters across the selected weather stations is shown in Figure 4.2. Data quality was checked before modeling, and recordings with missing data or physically meaningless values were removed. The mean ratio of actual to possible sunshine hours was limited to fall between 0 and 1, and wind speed between 0 and 20 m s^{-1}.

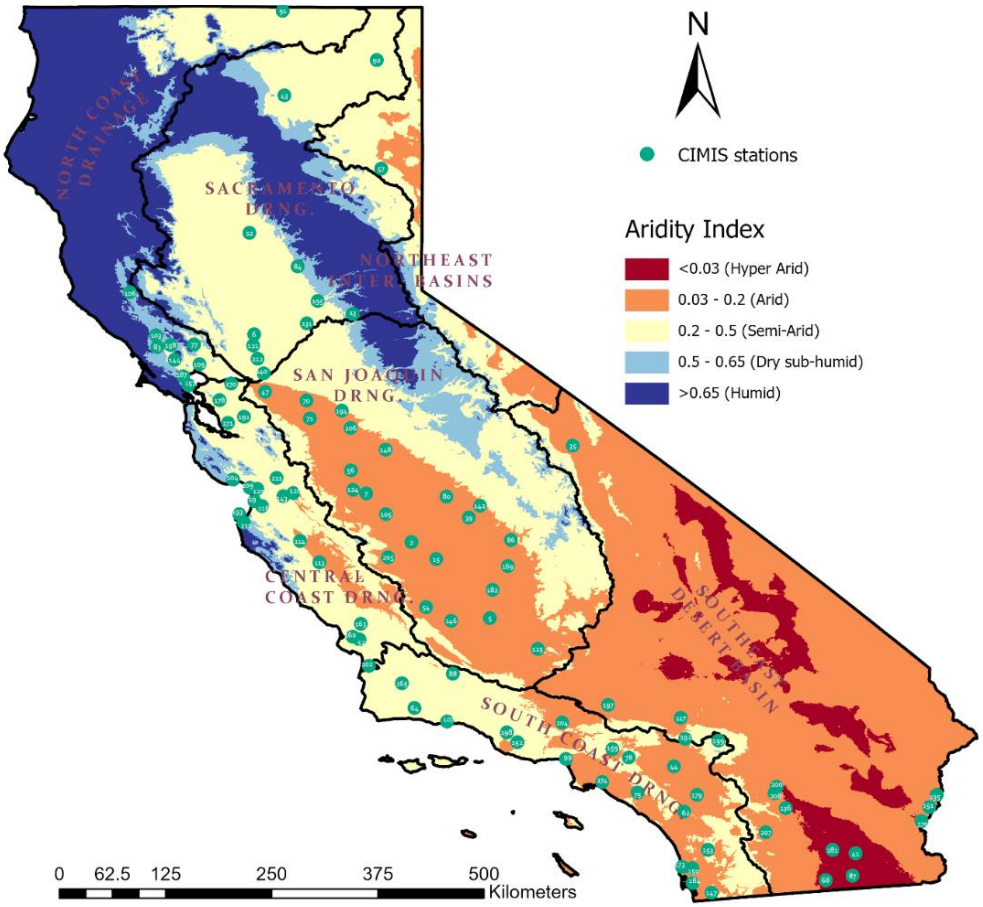


Figure 4.1. Distribution of the CIMIS stations evaluated in this study across the state of California. The aridity index values were obtained from the CGIAR-CSI Global-Aridity Database, and the classes were mapped based on the recommendations by United Nations Environmental Programme (UNEP).

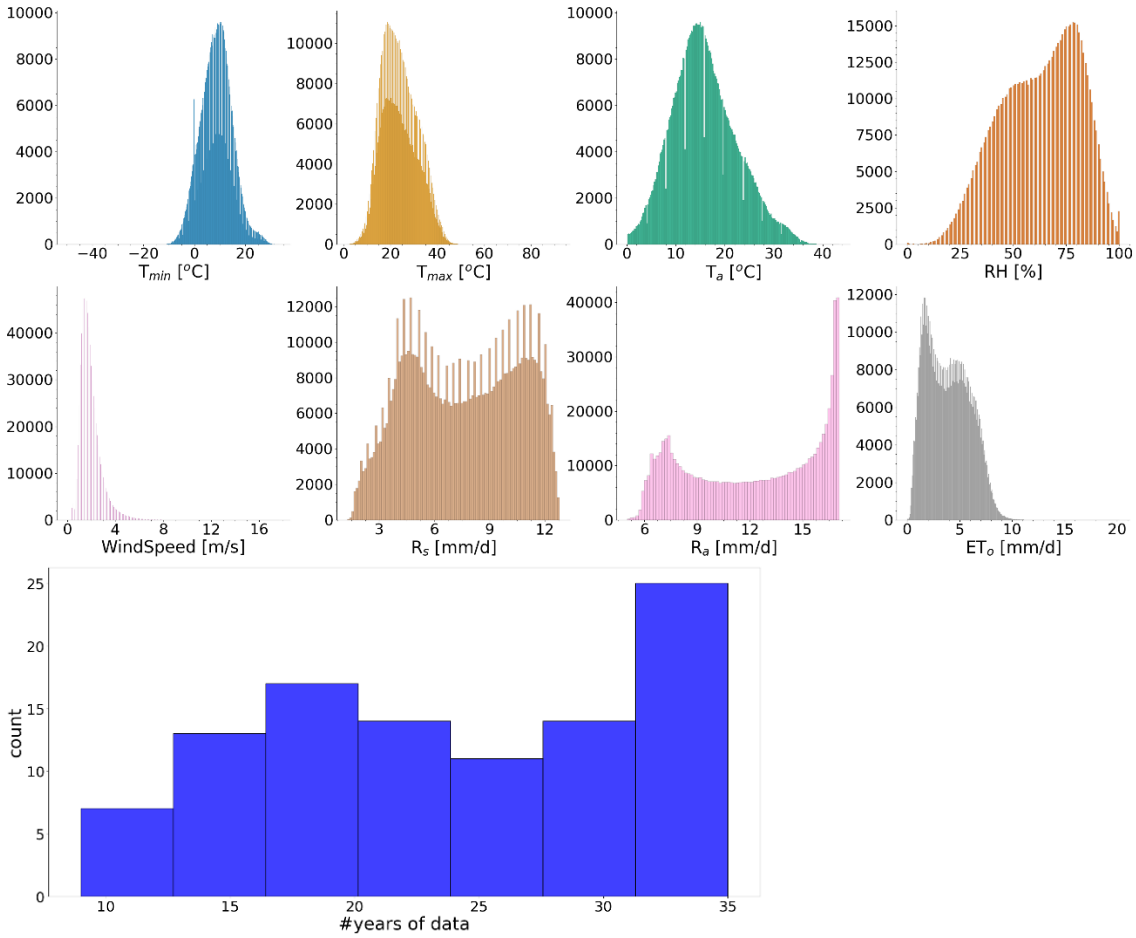


Figure 4.2. Histograms of the meteorological variables (top) and availability of the data (bottom) across CIMIS stations used in this study.

The aridity index, which can quantify precipitation availability over atmospheric water demand, was created following the classification recommended by United Nations Environmental Programme (UNEP) and using the CGIAR-CSI Global-Aridity Database (Trabucco and Zomer, 2018). The majority of the selected CIMIS stations belonged to arid (n = 48) and semi-arid (n = 39) classes, while 4, 9, and 1 CIMIS stations belonged to hyper-arid, dry sub-humid, and humid classes, respectively. The spatial dataset of the seven climate divisions of California was obtained from the National Climatic Data Center-National Oceanic and Atmospheric Administration (NCDC-NOAA). The climate division

data were used to perform a region-specific performance assessment of the models evaluated in this study.

4.2.2. Temperature-Based Empirical ET_o models

Following is a summary of the eight regression temperature-based empirical models used in this study. We used Python programming language version 3.8 (http://www.python.org) to calculate ET_o for all selected CIMIS stations. Furthermore, linear calibration equations were developed for the eight regression-based models using the daily ET_o data for each of the seven climate divisions.

The Blaney and Criddle equation, BC (Blaney and Criddle, 1950), is one of the earliest models of estimating ET_o involving the use of air temperature and is given as:

$$ET_o = p(0.46T_a + 8.13) \quad (1)$$

where ET_o is the reference evapotranspiration [mm d^{-1}], T_a is mean air temperature [°C], p is the percentage of total daytime hours out of total daytime hours of the year (365 × 12). The parameter p for a given latitude and month can be obtained from the tables provided in Doorenbos & Pruitt (1977).

The FAO 24 – Blaney and Criddle method, BC_{fao} (Doorenbos & Pruitt, 1977), is a modified version of the BC equation that includes correction factors to adjust for local weather or climatic conditions.

$$ET_o = A + B[p(0.46T_a + 8.13)] \quad (2)$$

$$A = 0.0043RHmin - \frac{n}{N} - 1.41 \quad (3)$$

$$B \quad (4)$$

$$= 0.82 - 0.0041(RHmin) + 1.07\left(\frac{n}{N}\right) + 0.066(Ud)$$

$$- 0.006(RHmin)(\frac{n}{N}) - 0.0006(RHmin)(Ud)$$

where *RHmin* is daily minimum relative humidity, *Ud* is the daytime wind speed, and *n/N* is the mean ratio of actual to possible sunshine hours that can be obtained from solar radiation estimates as:

$$\frac{n}{N} = 2\left(\frac{R_S}{R_a}\right) - 0.5 \quad (5)$$

where R_a is the extraterrestrial radiation expressed in equivalent evaporation units [mm d^{-1}] and R_s is the global solar radiation at the surface [mm d^{-1}]. Extraterrestrial solar radiation (R_a) was calculated following the equations given by Allen et al. (1998), using python library 'pyETo' (https://pyeto.readthedocs.io/en/latest/index.html). Daytime wind speed (*Ud*) was obtained by multiplying the 24-hr measurements of wind speed by 1.33 (Doorenbos and Pruitt, 1977).

The Hamon model, H (Hamon, 1961), is derived from the daily average air temperature and is given by:

$$ET_o = (0.1397N^2\, Pt)25.4 \quad (6)$$

where *N* is the maximum daylight hours possible for a given day [*h*], and *Pt* is a saturated water vapor density term calculated using the mean air temperature:

$$Pt = \frac{4.95e^{(0.062T_a)}}{100} \quad (7)$$

The original Hargreaves equation, H_s (Hargreaves, 1975), is based on the measured solar radiation:

$$ET_o = 0.0135R_s(T_a{}^{\circ}C + 17.8) \quad (8)$$

where R_s and T_a are the same as previously defined.

The Hargreaves and Samani method, HS_a (Hargreaves & Samani (1985), uses air temperature and estimated extraterrestrial solar radiation for a given latitude and day to estimate ET_o.

$$ET_o = 0.0023R_a(T_a{}^{\circ}C + 17.8)\sqrt{T_{max} - T_{min}} \quad (9)$$

where $T_{max} - T_{min}$ is the difference between the daily maximum and minimum air temperatures [°C].

The Jensen-Haise method, JH (Jensen & Haise, 1963), method is based on the global solar radiation information:

$$ET_o = R_s(0.025T_a + 0.08) \quad (10)$$

where R_s and T_a are the same as previously defined.

The Kharrufa method, K (Kharrufa, 1985), was developed based on the correlation of ET_o/p and T_a and is given as:

$$ET_o = 0.34pT_a^{1.3} \quad (11)$$

where p and T_a are the same as previously defined.

The Linacre method, L (Linacre, 1977), was developed for well-watered vegetation with an albedo of 0.25:

$$ET_o = \frac{500T_m/(100 - A) + 15(T_a - T_d)}{(80 - T_a)} \quad (12)$$

where T_m is the sea-level equivalent of the measured mean air temperature (T_a) calculated as $T_a + 0\ 006h$, h is the elevation [m], A is the latitude [degrees], and T_d is the mean dew point. The T_a, T_m, and T_d are in °C.

4.2.3. Artificial Neural Network Models

Two feed-forward backpropagation ANN models were developed with different combinations of input variables (Figure 4.3). One model used a wide range of inputs, including T_{min}, T_{max}, T_a, RH, U, and R_a, while the other model only had T_{min}, T_{max}, T_a, and R_a as inputs. A total of 725,849 data points from 101 CIMIS stations were divided into five folds such that four folds were used for training and the remaining group (~20% of the data) for testing the models. Cross-validation was done using 10% of the training set to terminate the training and avoid overfitting. The model development process was repeated five times to ensure that data from all the CIMIS stations had been in the test set. This approach assessed the generalization ability of the models since data from the same CIMIS station was never utilized in the training and test set simultaneously. The "adam" optimizer function was used to train the feed-forward ANN models and the best weights and biases were automatically loaded for testing. Activation functions were the "ReLu" and "linear" for the hidden and the output layers, respectively. The maximum epoch (i.e., one cycle of

a complete presentation of the training data set through the learning process) was set to 1000. The Python library "tensorflow" (TensorFlow Developers, 2021) was utilized to develop the ANN models.

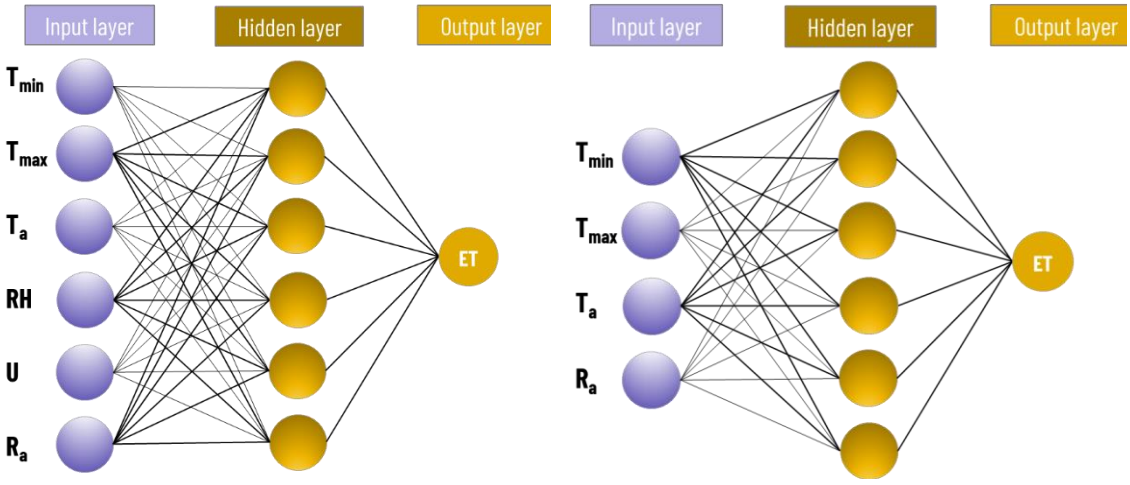


Figure 4.3. The architecture of the feed-forward backpropagation neural network models that were developed in this study. T_{min}, T_{max}, and T_a are the daily minimum, maximum, and mean air temperature [°C], respectively; RH: Relative Humidity [%]; U: daily mean windspeed [m/s]; R_a: extraterrestrial solar radiation [mm/d].

The same aforementioned steps were followed to develop WA-ANN models using the reconstructed signals obtained from wavelet transform. Discrete wavelet transform (DWT) was applied to decompose the climatic time series of input predictors by passing through a series of high-pass and low-pass filters, separated at different scales. The Daubechies (db10) family of wavelets with decompositions level of 10 (2–4–8–16–32–64–128–256–512–1024) was used. The original time series were decomposed into series of approximation(A) and detail(D) sub-series. The correlation coefficient between ET_o data series and the decomposed sub-series was computed. The sub-series with correlation coefficient values greater than ±0.1 were recognized as effective (Partal, 2009). The

effective wavelet subseries were used to reconstruct the time series used as inputs to the ANN model. The Python library "Pywavelets" (Lee et al., 2019) was used for this analysis.

4.2.4. Performance Assessment

The performance of the models was evaluated against CIMIS ET_o data, a modified version of the Penman equation (Pruitt and Doorenbos, 1977) that uses a wind function developed at the University of California, Davis and unique cloud factor values for each station location. Four statistical indices, including the root mean squared error (RMSE), mean absolute error (MAE), mean bias error (MBE), and coefficient of determination (R^2), were used to quantify and compare the performance of the temperature-based models.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(E_i - M_i)^2} \tag{5}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|E_i - M_i| \tag{6}$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(E_i - M_i) \tag{7}$$

$$R^2 = \left(\frac{\sum_{i=1}^{n}(E_i - \bar{E})(M_i - \bar{M})}{\sqrt{\sum_{i=1}^{n}(E_i - \bar{E})^2 \sum_{i=1}^{n}(M_i - \bar{M})^2}}\right)^2 \tag{8}$$

where E and M are the estimated and measured ET_o, respectively; $\bar{E}$ and $\bar{M}$ are the mean estimated and measured ET_o, respectively; and n is the total number of data points for each model.

Monthly and annual error metrics were also computed to understand the performance of the models on a temporal scale. The long-term ET_o data were interpolated using the inverse distance weighting approach in ArcGIS Pro 2.8.2 (ESRI Inc.) to assess the spatial performance of the models evaluated in this study.

4.3. Results & discussion

4.3.1 Overall performance of the ET_o models

Table 4.1 summarizes the overall performance statistics of all the empirical regression and ANN-based ET_o models evaluated in this study. Figure 4.4 depicts the scatter plots of the ET_o estimated by the models against CIMIS ET_o. The empirical temperature-based models can be ranked as H_s > HS_a >BC_{fao} > JH > BC >K > H > L based on their performance. The H_s (RMSE = 0.61 mm d^{-1}; MAE = 0.41 mm d^{-1}) and HS_a (RMSE = 0.78 mm d^{-1}; MAE = 0.56 mm d^{-1}) models showed the highest accuracy among the empirical temperature-based models. However, the H_s model (MBE = -0.28 mm d^{-1}) tended to underestimate ET_o, whereas the HS_a model resulted in the lowest magnitude of positive or negative bias among all empirical regression models. Kukal et al. (2020) reported underestimation of ET_o by the HS_a model in arid, semi-arid, and dry subhumid sites of the U.S. High Plains region, whereas overestimation of ET_o by the HS_a model was reported in the Sahelian climate by Djaman et al. (2015). The L model (RMSE = 1.55 mm d^{-1}; MAE = 1.17 mm d^{-1}) resulted in the poorest performance followed by the H model (RMSE = 1.29 mm d^{-1}; MAE = 1.02 mm d^{-1}), the K model (RMSE = 1.18 mm d^{-1}; MAE = 0.91 mm d^{-1}), and the BC model (RMSE = 1.13 mm d^{-1}; MAE = 0.91 mm d^{-1}). Performance of the BC_{fao} (RMSE = 0.91 mm d^{-1}; MAE = 0.67 mm d^{-1}) and the JH models (RMSE = 0.90 mm

d^{-1}; MAE = 0.70 mm d^{-1}) was comparable however resulting in slight overestimation and considerable underestimation of ET_o, respectively. The R^2 values ranged from 0.60 for the L model to 0.93 for the H_s model. Most models showed a tendency to underestimate the CIMIS ET_o except for the BC, BC_{fao}, and L models, which showed various degrees of overestimation, as illustrated in Figure 4.4.

The ANN models can be ranked as ANN_all > ANN_T > WA-ANN_all > WA-ANN_T based on their performance. They showed high accuracy and reliability (generalization ability), given that data from the same CIMIS station was never present in the training and test set simultaneously. The ANN_all model (inputs: T_{min}, T_{max}, T_a, R_a, RH, U) showed the best performance (RMSE = 0.51 mm d^{-1}; MAE = 0.35 mm d^{-1}) followed by ANN-T (RMSE = 0.69 mm d^{-1}; MAE = 0.49 mm d^{-1}). The WA-ANN models using reconstructed signals obtained from the wavelet transform resulted in 69% higher error than ANN models using raw data as inputs (Table 4.1). The MBE values were negligible in all ANN models indicating no substantial over or underestimation of the ET_o, as shown in Figure 4.4. The R^2 values ranged from 0.77 for the WA-ANN_T to 0.94 for the ANN_all model. Models using reconstructed signal obtained from wavelet transform performed reasonably well with WA-ANN_all (RMSE = 0.96 mm d^{-1}; MAE = 0.70 mm d^{-1}) performing slightly better than WA-ANN_T (RMSE = 0.99 mm d^{-1}; MAE = 0.73 mm d^{-1}). Using the raw input data was a better approach than using the reconstructed signal in our study. This result agrees with the findings of Falamarzi et al. (2014) but differs from the study by Kisi & Alizamir (2018), where ANN models with wavelet transform inputs had no considerable difference from the models developed using raw inputs.

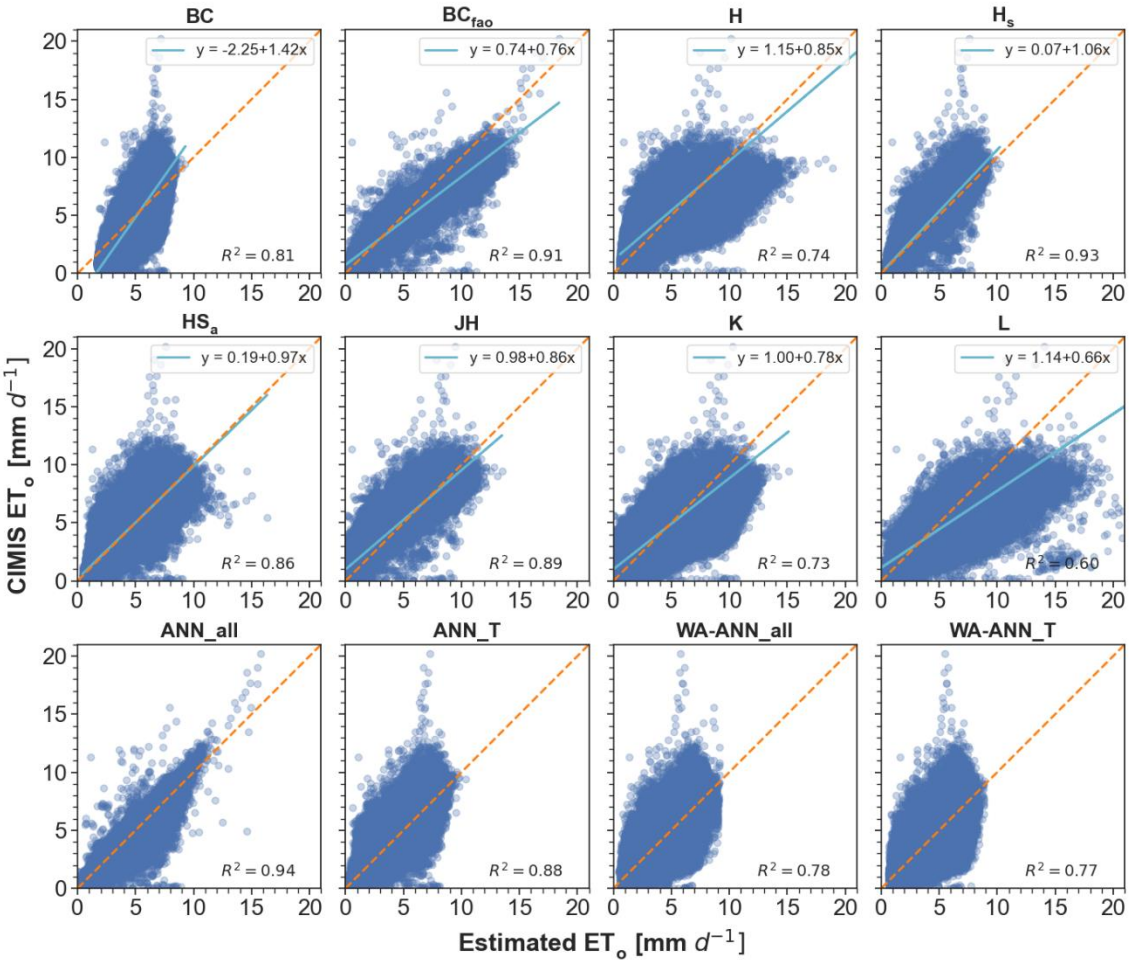


Figure 4.4. Scatterplots of the CIMIS ET_o versus estimated daily ET_o by the 12 temperature-based empirical and ANN models evaluated in this study. The dashed orange line is the 1:1 line.

Table 4.1. Overall performance of the temperature-based ET_o models compared to CIMIS ET_o.

Models	Inputs	RMSE	MAE	MBE	R^2	Linear eq
BC	T_a	1.13	0.91	0.43	0.81	y = 1.42x-2.25
BC_{fao}	T_a, RH, Ud	0.91	0.67	0.27	0.91	y = 0.76x+0.74
H	T_a	1.29	1.02	-0.68	0.74	y = 0.85x+1.15
H_s	T_a, R_s	0.61	0.41	-0.28	0.93	y = 1.06x+0.07
HS_a	T_{min}, $T_{max,}$ T_a, R_a	0.78	0.56	-0.07	0.86	y = 0.97x+0.19
JH	T_a, R_s	0.90	0.70	-0.49	0.89	y = 0.86x+0.98
K	T_a	1.18	0.91	-0.21	0.73	y = 0.78x+1.0
L	T_a,T_d	1.55	1.17	0.27	0.60	y = 0.66x+1.14
ANN_all	T_{min}, $T_{max,}$ T_a, R_a, RH, U	0.51	0.35	0.003	0.94	-
ANN_T	T_{min}, $T_{max,}$ T_a, R_a	0.69	0.49	0.000	0.88	-
WA-ANN_all	T_{min}, $T_{max,}$ T_a, R_a, RH, U	0.96	0.70	-0.009	0.78	-
WA-ANN_T	T_{min}, $T_{max,}$ T_a, R_a	0.99	0.73	0.001	0.77	-

BC: Blaney and Criddle, BC_{fao}: Blaney & Criddle FAO, H: Hamon, H_s: Hargreaves, HS_a: Hargreaves & Samani, JH: Jensen & Haise, K: Kharrufa, L: Linacre. ANN_all: ANN model with T_{min}, T_{max}, T_a, R_a, RH, and U as inputs, WA-ANN_all: Wavelet transform ANN model with T_{min}, T_{max}, T_a, R_a, RH, and U as inputs, ANN_T: ANN model with T_{min}, T_{max}, T_a, and R_a as inputs, WA-ANN_T: wavelet transform ANN model with T_{min}, T_{max}, T_a, and R_a as inputs.T_{min}, T_{max}, and T_a are in [°C], Relative Humidity, RH [%], daily average windspeed, U [m/s], Extraterrestrial solar radiation, R_a [mm/d].

The empirical and ANN-based models can be ranked as ANN_all >H_s >ANN_T >HS_a >BC_{fao} > JH >WA-ANN_all > WA-ANN_T > BC >K > H > L in terms of their overall performance based on the indices evaluated in this study. The ANN_all was the best performing model overall, followed by the H_s model. The ANN-T was the third-best overall model, outperforming the HS_a model with similar input variables. Better performance in estimating ET_o by ANN and WA-ANN models compared to HS_a was reported in the literature (Adamala, 2018; Traore et al., 2010). Linear calibration of the

empirical models using coefficients given in Table 4.1 resulted in improved performance, as calibrated models resulted in 1%(HS_a) to 33% (BC_{fao}) lower RMSE.

4.3.2 Importance of the Input Parameters

Overall, empirical models using solar radiation resulted in superior performance, as three out of the four best performing models (i.e., H_s, HS_a, BC_{fao}, and JH) required air temperature and solar radiation measurements as inputs. Only HS_a can be regarded as a truly temperature-based model since extraterrestrial solar radiation can be estimated based on the location and time of the year. Global solar radiation was a more critical parameter than relative humidity and wind speed in estimating daily ET_o, based on a machine learning study conducted in China (Fan et al., 2018). This can explain the H_s model's slightly better performance over the HS_a as using extraterrestrial solar radiation may ignore the considerable variability in ET_o due to cloud cover. However, daily ET_o anomalies in California strongly correlate with net radiation anomalies, relative humidity (RH), and cloud cover, and less with average daily air temperature (Hidalgo et al., 2005). The temperature difference used in the HS_a model is an indirect measure of cloud cover and RH at a location (Hargreaves and Samani, 1985). Clear sky conditions result in high maximum day temperatures and low minimum night temperatures, whereas cloudy days result in relatively lower maximum day temperatures and higher minimum night temperatures.

In our study, the ANN_all model using relative humidity, wind speed, air temperature, and extraterrestrial solar radiation resulted in better performance (29% reduction in MAE) than ANN_T, based on temperature and extraterrestrial solar radiation. However, ANN_all

requires more data typically unavailable for ET_o estimation by smart controllers in residential areas. Partal (2009) reported no improvement in performance by the ANN and WA-ANN models over the HS_a model. In contrast, the ANN_T model in our study resulted in a 12.5% reduction in MAE compared to the HS_a empirical model derived from the same inputs. Therefore, implementing calibrated HS_a or ANN_T models are suggested for obtaining accurate estimations of ET_o for landscape irrigation management by smart weather-based irrigation controllers in data-scarce conditions in urban settings.

4.3.3 Temporal Analysis of the ET_o models

Figure 4.5 shows the daily mean and standard deviation of estimated ET_o values throughout the year by the models evaluated in this study versus the CIMIS ET_o. The monthly performance statistics for each model to estimate monthly average ET_o are shown in Tables 4.2 to 4.4. The empirical temperature-based models can be ranked as HS_a > H_s >BC_{fao} >K > JH > BC > L > H, based on the monthly performance. The HS_a (MAE: 0.10 to 0.31 mm d^{-1}) and H_s (MAE: 0.13 to 0.51 mm d^{-1}) models consistently performed better than other models throughout the year; however, they had a high error in the spring months relative to other months. The BC (MAE: 0.11 to 1.20 mm d^{-1}) model had a low error in the spring and summer months but performed worse in the winter months. On the other hand, the BC_{fao} (MAE: 0.08 to 0.76 mm d^{-1}) model had comparable performance to the H_s and HS_a models in the winter and spring months but performed poorly in the summer months, also evident in Figure 4.5. All empirical models, except the BC_{fao} model, had relatively better performance in July and August, also suggested by the low magnitude of MBE in these months (Table 4.4).

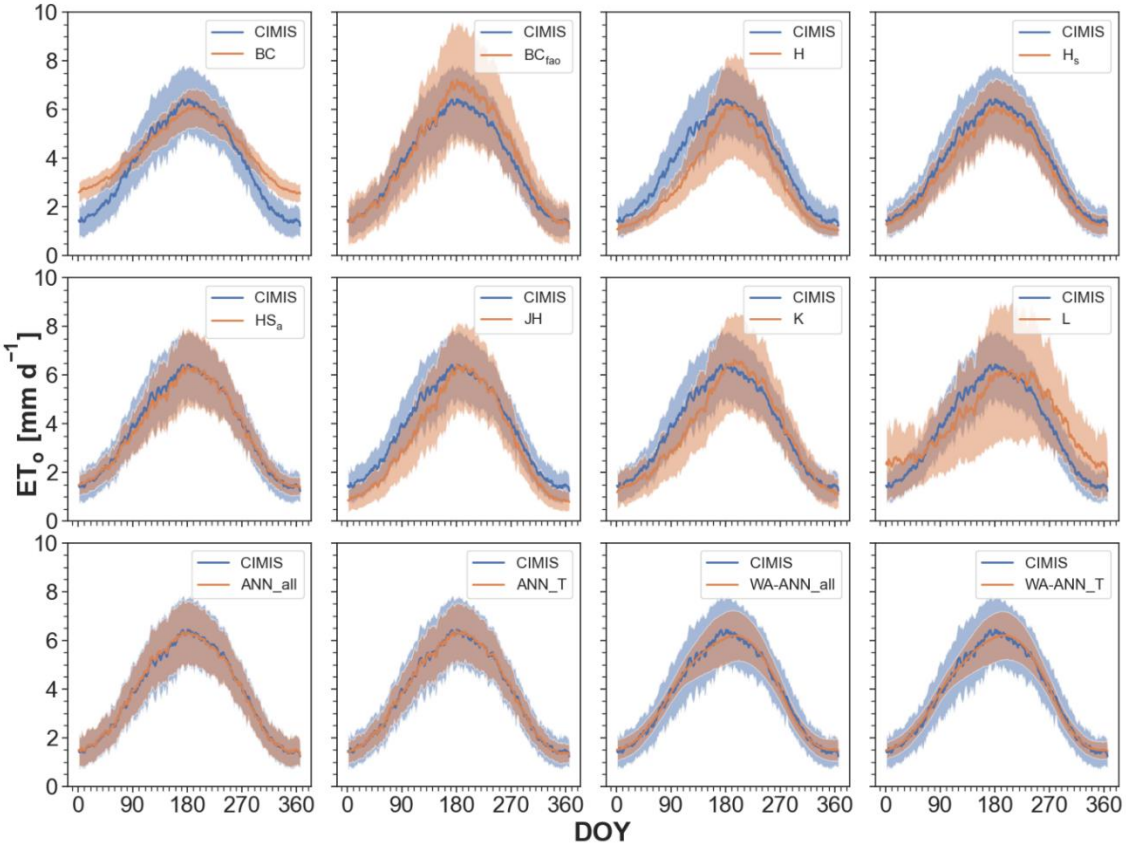


Figure 4.5. Comparison between long-term year-round ET_o obtained from CIMIS against the 12 temperature-based empirical and ANN ET_o models evaluated in this study. The solid lines show mean ET_o and the shaded bands depict the standard deviation of ET_o across all CIMIS stations. DOY: day of the year.

Table 4.2. Monthly root mean square error (RMSE) values for the temperature-based ET_o equations evaluated in this study against CIMIS ET_o.

Model	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
BC	1.21	0.98	0.55	0.25	0.39	0.44	0.29	0.21	0.39	0.81	1.16	1.21
BC_{fao}	0.15	0.15	0.24	0.23	0.22	0.61	0.79	0.70	0.69	0.53	0.25	0.11
H	0.41	0.70	1.12	1.50	1.31	0.78	0.38	0.54	0.72	0.77	0.50	0.38
H_s	0.18	0.24	0.35	0.53	0.52	0.42	0.33	0.34	0.31	0.32	0.22	0.20
HS_a	0.13	0.13	0.20	0.34	0.33	0.24	0.16	0.12	0.13	0.13	0.15	0.12
JH	0.60	0.76	0.93	1.06	0.82	0.36	0.17	0.16	0.28	0.52	0.57	0.60
K	0.34	0.51	0.83	1.16	0.97	0.47	0.40	0.35	0.40	0.33	0.28	0.31
L	0.83	0.45	0.25	0.68	0.84	0.68	0.41	0.39	0.91	1.19	1.11	0.94
ANN_all	0.08	0.08	0.09	0.14	0.16	0.13	0.13	0.16	0.09	0.10	0.07	0.08
ANN_T	0.11	0.11	0.15	0.16	0.16	0.16	0.13	0.13	0.12	0.15	0.14	0.14
WA-ANN_all	0.16	0.21	0.21	0.26	0.20	0.34	0.23	0.24	0.20	0.31	0.20	0.21
WA-ANN_T	0.16	0.22	0.22	0.26	0.23	0.33	0.21	0.25	0.23	0.31	0.21	0.21

Table 4.3. Monthly mean absolute error (MAE) values for the temperature-based ET_o equations evaluated in this study against CIMIS ET_o.

Model	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
BC	1.19	0.95	0.51	0.19	0.33	0.40	0.22	0.11	0.36	0.78	1.15	1.20
BC_{fao}	0.11	0.14	0.20	0.20	0.17	0.57	0.76	0.69	0.67	0.51	0.22	0.08
H	0.36	0.65	1.10	1.49	1.28	0.75	0.30	0.49	0.69	0.74	0.46	0.33
H_s	0.13	0.20	0.32	0.51	0.50	0.40	0.30	0.30	0.27	0.28	0.17	0.15
HS_a	0.11	0.10	0.16	0.31	0.28	0.21	0.13	0.10	0.11	0.11	0.11	0.10
JH	0.58	0.74	0.91	1.05	0.80	0.33	0.12	0.12	0.21	0.48	0.54	0.57
K	0.28	0.43	0.77	1.14	0.92	0.43	0.33	0.31	0.35	0.26	0.22	0.25
L	0.79	0.41	0.21	0.63	0.81	0.61	0.32	0.32	0.87	1.16	1.08	0.91
ANN_all	0.06	0.06	0.07	0.10	0.14	0.10	0.10	0.12	0.08	0.08	0.06	0.05
ANN_T	0.08	0.09	0.11	0.12	0.12	0.13	0.10	0.09	0.09	0.11	0.10	0.10
WA-ANN_all	0.13	0.17	0.16	0.19	0.16	0.29	0.18	0.19	0.16	0.24	0.15	0.15
WA-ANN_T	0.14	0.18	0.16	0.20	0.18	0.27	0.17	0.20	0.18	0.23	0.15	0.15

Table 4.4. Monthly mean bias error (MBE) values for the temperature-based ET_o equations evaluated in this study against CIMIS ET_o.

Model	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
BC	1.19	0.95	0.49	-0.08	-0.28	-0.38	-0.20	-0.07	0.36	0.78	1.15	1.20
BC_{fao}	-0.06	-0.10	-0.19	-0.17	0.11	0.57	0.76	0.69	0.67	0.51	0.20	-0.04
H	-0.36	-0.65	-1.10	-1.49	-1.28	-0.75	-0.19	-0.48	-0.69	-0.74	-0.46	-0.33
H_s	-0.13	-0.20	-0.32	-0.51	-0.50	-0.40	-0.30	-0.30	-0.27	-0.28	-0.17	-0.15
HS_a	0.05	-0.02	-0.14	-0.30	-0.28	-0.18	-0.01	-0.02	0.04	-0.02	0.03	0.02
JH	-0.58	-0.74	-0.91	-1.05	-0.80	-0.33	0.04	-0.02	-0.21	-0.48	-0.54	-0.57
K	-0.18	-0.41	-0.77	-1.14	-0.92	-0.40	0.25	0.27	0.31	0.16	0.02	-0.17
L	0.79	0.41	-0.16	-0.63	-0.81	-0.60	-0.09	0.29	0.87	1.16	1.08	0.91
ANN_all	0.03	0.01	0.02	0.08	0.12	0.00	-0.05	-0.11	-0.03	-0.06	-0.02	0.02
ANN_T	-0.01	-0.02	0.01	-0.02	0.00	-0.05	0.02	-0.01	0.04	-0.05	-0.02	-0.05
WA-ANN_all	0.06	0.02	0.01	0.01	-0.04	-0.23	-0.06	0.04	0.03	-0.20	-0.04	0.06
WA-ANN_T	0.05	0.03	0.02	-0.04	-0.08	-0.19	-0.01	0.07	0.06	-0.16	-0.03	0.04

The ANN-based models can be ranked as ANN_all > ANN_T > WA-ANN_all >WA-ANN_T. The ANN models showed negligible bias throughout the year with MBE ranging from -0.23 to 0.13 mm d^{-1}, whereas MBE ranged from -0.51 to 0.05 mm d^{-1} for the best performing empirical models, i.e., H_s and HS_a (Table 4.2 to 4.4). The MAE values varied from 0.05 to 0.14 mm d^{-1} for the ANN models and 0.13 to 0.29 mm d^{-1} for the WA-ANN models. Better performance of ANN models is also apparent in Figure 4.5, where the mean and standard deviation of the estimated ET_o matches very closely to that of CIMIS ET_o throughout the year. Overall, the empirical and ANN-based models can be ranked as ANN_all >ANN_T >HS_a > WA-ANN_all > WA-ANN_T >H_s > BC_{fao} > K >JH > BC > L > H in terms of their monthly performance.

The relationship between MAE and MBE and the annual average of the meteorological parameters, including wind speed, relative humidity, vapor pressure deficit, and global solar radiation, are shown in Figures 4.6 and 4.7. All models tended to have higher MAE with increasing wind speed, vapor pressure deficit, and global solar radiation, except BC, BC_{fao}, and L models. Similarly, models tend to have lower MAE with increasing relative humidity except for BC, BC_{fao}, and L models. In the same way, models tend to underestimate ET_o with increasing wind speed, vapor pressure deficit, and global solar radiation, as indicated by the decreasing MBE. With increasing relative humidity, all models overestimate ET_o except L and BC_{fao} models, where MBE decreased or changed negligibly. The HS_a model tended to underestimate ET_o under high wind conditions (> 3 m s^{-1}) and to overestimate ET_o under conditions of high relative humidity (Allen et al., 1998), which is also supported by our results. Reasons may include the continual mixing

of warm, dry air from overhead into the equilibrium boundary layer at night, reducing the difference between the maximum and minimum air temperature values (Temesgen et al., 1999).

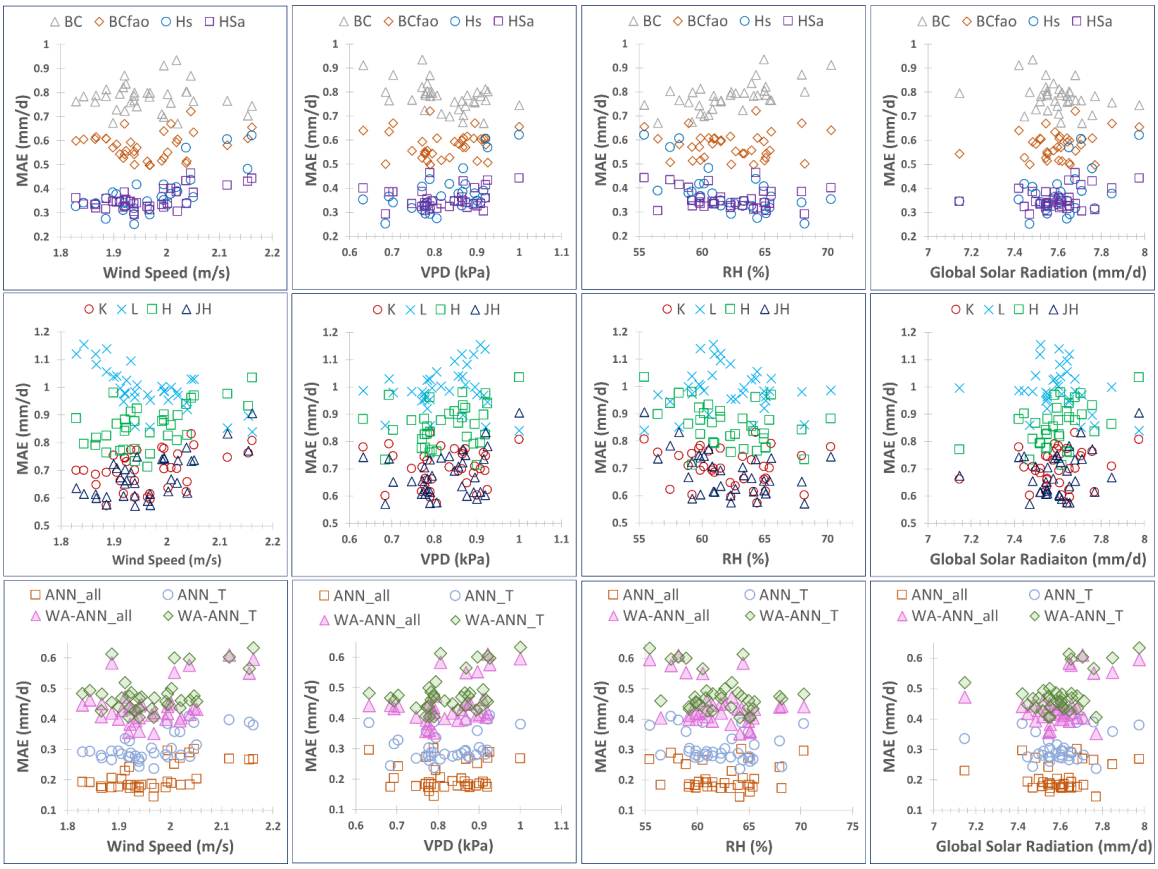


Figure 4.6. Variation in annual mean absolute error, MAE [mm d^{-1}] values against meteorological variables - wind speed, relative humidity (RH), vapor pressure deficit (VPD), and global solar radiation (R_s).

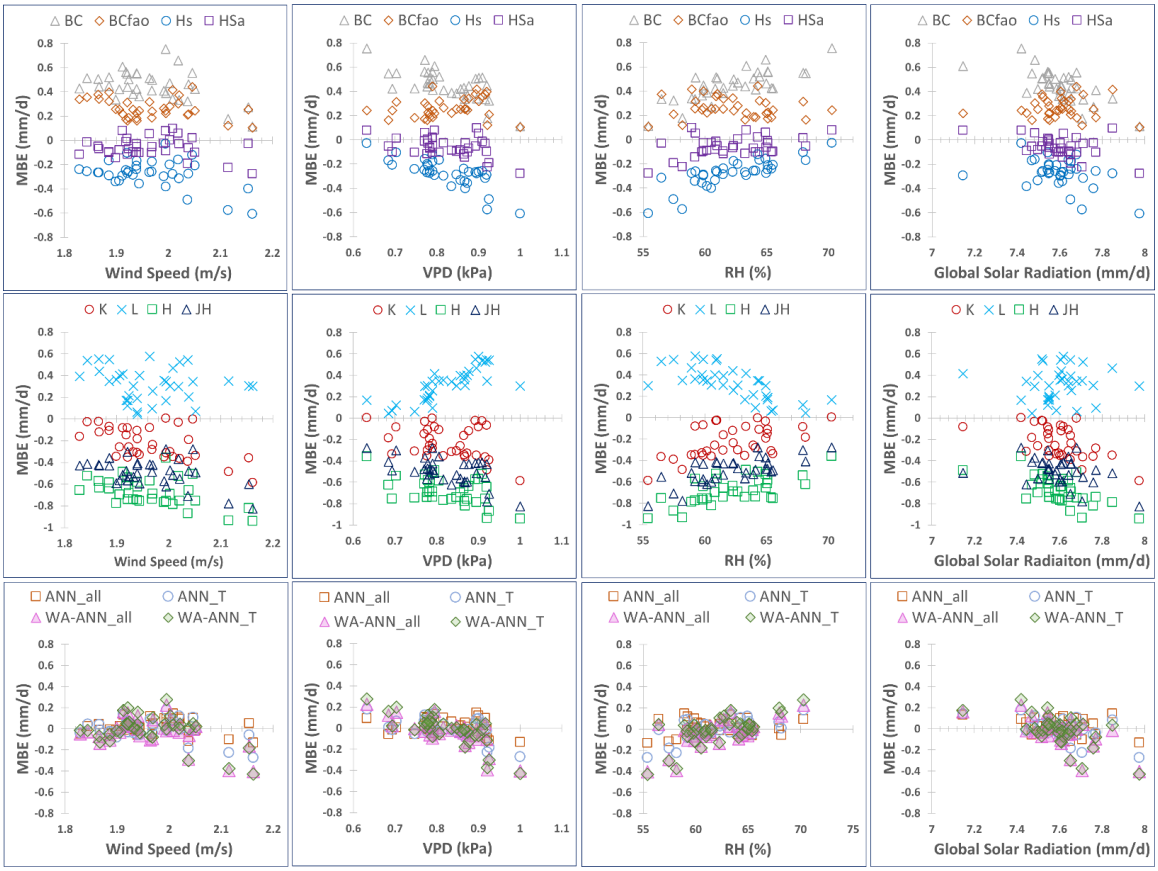


Figure 4.7. Variation in annual mean bias error, MBE [mm d^{-1}] values against meteorological variables - wind speed, relative humidity (RH), vapor pressure deficit (VPD), and global solar radiation (R_s).

4.3.4 Spatial Analysis of the ET_o models

Table 4.5 summarizes the performance of the models evaluated in this study for each climate division. Also shown in Table 4.5 are the coefficients (i.e., A and B) of BC_{fao} averaged for each climate division. The average values of these coefficients across all CIMIS stations were A = -1.96 and B = 1.37. The empirical models can be ranked as H_s >HS_a > BC_{fao} > JH > BC > K > H > L, based on their performance in the climate divisions of California. Based on the effects of weather variables on annual ET_o (Figure 4.6 and 4.7), it can be expected that the performance of temperature-based ET_o equations will vary based

on the climate, i.e., more arid regions will tend to have higher error. The southeast desert basin had the highest error among all climate divisions as it is the aridest region of the state (Table 4.5). The H_s was the best performing model with the lowest MAE (0.31 to 0.81 mm d^{-1}) in all climate divisions, except northeast interior basins where HS_a (MAE: 0.47 to 0.73 mm d^{-1}) performed best. As pointed out in previous sections, the H_s model also resulted in a high magnitude of negative bias (indicating underestimation) in all climate divisions compared to the HS_a model. Gabriela Arellano and Irmak (2016) reported an underestimation of ET_o and RMSE of 1.0 mm d^{-1} for the HS_a model at CIMIS station 6 in the Sacramento drainage region. This closely agrees with the results of our study with RMSE of 0.87 mm d^{-1} and underestimation (MBE = -20 mm d^{-1}) of ET_o, as is also evident in Figure 4.8. The BC_{fao} model (MAE: 0.47 to 1.15 mm d^{-1}) was a better performer than the BC model, particularly in the coastal regions, i.e., north coast drainage, central coast drainage, and south coast drainage, where its performance was similar to the HS_a model. This can also be noted in Figure 4.8, where the mean ET_o estimated by the BC_{fao} model is similar to the CIMIS ET_o in the coastal regions. The BC (MAE: 0.83 to 0.93 mm d^{-1}) model performed better than BC_{fao} only in the southeast desert basin, although the error was relatively high for all the empirical models in this region (Table 4.5). The JH (MAE: 0.55 to 1.36 mm d^{-1}) model also had satisfactory performance in San Joaquin and south coast drainage climate divisions. The mean long-term CIMIS ET_o ranged from 2.58 mm d^{-1} (at CIMIS #193 in central coast drainage) to 5.29 mm d^{-1} (CIMIS # 200 in the southeast desert basin) (Figure 4.8). The mean ET_o ranged from 1.93 to 8.03 mm d^{-1} for the temperature-

based models, with the lowest (at CIMIS # 19 in central coast drainage) and highest (CIMIS # 200 in the southeast desert basin) observed for the L model (Figure 4.8).

Table 4.5. Climate-division-specific calibration equations and performance statistics for the temperature-based ETo equations evaluated in this study against CIMIS ET_o.

Climate Division	ET Model	linear model	R^2	RMSE	MAE	MBE
401 (North Coast Drainage) n = 10 A = -1.91 B = 1.29	**BC**	y = 1.48x -2.42	0.82	1.06	0.86	0.50
	BC_{fao}	y = 0.86x+0.60	0.92	0.61	0.47	-0.13
	H	y = 1.21x+0.35	0.81	1.23	1.03	-0.89
	H_s	y = 1.08x+0.01	0.96	0.47	0.34	-0.28
	HS_a	y = 0.95x+0.05	0.88	0.64	0.47	0.11
	JH	y = 0.99x+0.78	0.91	0.93	0.77	-0.75
	K	y = 0.98x+0.68	0.74	1.12	0.90	-0.63
	L	y = 0.88x+0.81	0.53	1.34	1.03	-0.45
	ANN_all		0.94	0.45	0.31	0.00
	ANN_T		0.90	0.58	0.40	0.15
	WA-ANN_all		0.77	0.91	0.67	0.22
	WA-ANN_T		0.77	0.94	0.68	0.30
402 (Sacramento Drng.) n = 11 A = -1.97 B = 1.42	**BC**	y = 1.44x -2.22	0.84	1.12	0.91	0.32
	BC_{fao}	y = 0.77x+0.59	0.94	0.90	0.67	0.38
	H	y = 0.97x+0.88	0.79	1.29	1.02	-0.80
	H_s	y = 1.07x+0.15	0.93	0.71	0.49	-0.39
	HS_a	y = 0.99x+0.22	0.85	0.87	0.62	-0.19
	JH	y = 0.88x+1.02	0.88	1.02	0.79	-0.64
	K	y = 0.85x+0.96	0.77	1.18	0.93	-0.42
	L	y = 0.79x+0.55	0.68	1.39	1.06	0.37
	ANN_all		0.95	0.47	0.32	0.03
	ANN_T		0.88	0.75	0.53	-0.11
	WA-ANN_all		0.80	0.98	0.71	-0.01
	WA-ANN_T		0.78	1.02	0.73	0.02
403 (Northeast Inter. Basins) n = 1 A = -1.99 B = 1.49	**BC**	y = 1.40x -1.43	0.85	0.99	0.83	-0.12
	BC_{fao}	y = 0.76x+0.97	0.94	0.83	0.67	0.01
	H	y = 1.06x+1.17	0.80	1.65	1.41	-1.34
	H_s	y = 1.12x+0.43	0.95	0.97	0.81	-0.81
	HS_a	y = 0.94x+0.49	0.88	0.78	0.58	-0.28
	JH	y = 0.92x+1.56	0.89	1.53	1.36	-1.35
	K	y = 0.86x+1.70	0.78	1.68	1.43	-1.32
	L	y = 0.75x+0.36	0.76	1.49	1.14	0.87
	ANN_all		0.93	0.57	0.39	0.15
	ANN_T		0.90	0.70	0.51	-0.19
	WA-ANN_all		0.77	1.01	0.75	-0.08

	WA-ANN_T		0.75	1.09	0.78	0.02
404 (Central Coast Drng.) n = 20 A = -1.86 B = 1.23	**BC**	y = 1.50x -2.64	0.75	1.13	0.92	0.63
	BC_{fao}	y = 0.88x+0.62	0.87	0.66	0.51	-0.24
	H	y = 1.21x+0.28	0.75	1.19	0.99	-0.82
	H_s	y = 1.08x -0.02	0.95	0.44	0.31	-0.21
	HS_a	y = 1.01x+0.16	0.84	0.68	0.52	-0.18
	JH	y = 1.02x+0.61	0.91	0.82	0.69	-0.67
	K	y = 1.01x+0.42	0.67	1.04	0.84	-0.46
	L	y = 0.81x+1.03	0.41	1.37	1.11	-0.49
	ANN_all		0.90	0.53	0.40	-0.13
	ANN_T		0.87	0.59	0.45	-0.10
	WA-ANN_all		0.70	0.92	0.69	0.22
	WA-ANN_T		0.69	0.94	0.70	0.23
405 (San Joaquin Drng.) n = 23 A = -1.98 B = 1.42	**BC**	y = 1.50x - 2.57	0.88	1.13	0.93	0.35
	BC_{fao}	y = 0.77x+0.54	0.96	0.95	0.70	0.50
	H	y = 0.93x+0.88	0.82	1.16	0.91	-0.62
	H_s	y = 1.08x - 0.01	0.95	0.60	0.40	-0.30
	HS_a	y = 1.01x - 0.05	0.90	0.73	0.52	0.02
	JH	y = 0.88x+0.84	0.92	0.79	0.62	-0.38
	K	y = 0.85x+0.70	0.82	1.04	0.80	-0.09
	L	y = 0.81x+0.50	0.78	1.23	0.93	0.37
	ANN_all		0.96	0.45	0.31	0.02
	ANN_T		0.92	0.65	0.46	0.05
	WA-ANN_all		0.86	0.85	0.61	-0.01
	WA-ANN_T		0.85	0.88	0.64	-0.06
406 (South Coast Drng.) n = 21 A = -1.93 B = 1.32	**BC**	y = 1.35x-2.17	0.70	1.17	0.93	0.64
	BC_{fao}	y = 0.82x+0.57	0.85	0.72	0.55	0.09
	H	y = 0.95x+0.82	0.69	1.14	0.92	-0.66
	H_s	y = 1.01x+0.08	0.91	0.52	0.36	-0.12
	HS_a	y = 0.94x+0.31	0.80	0.76	0.56	-0.08
	JH	y = 0.87x+0.80	0.87	0.74	0.55	-0.37
	K	y = 0.80x+0.76	0.63	1.05	0.82	-0.04
	L	y = 0.51x+1.68	0.33	1.64	1.26	0.22
	ANN_all		0.88	0.56	0.40	-0.01
	ANN_T		0.83	0.70	0.50	0.05
	WA-ANN_all		0.64	1.00	0.76	-0.04
	WA-ANN_T		0.63	1.01	0.76	0.01
407 (Southeast Desert Basin) n = 15 A = -2.10 B = 1.60	**BC**	y = 1.24x-1.26	0.78	1.14	0.90	0.08
	BC_{fao}	y = 0.72x+0.64	0.92	1.47	1.15	0.99
	H	y = 0.61x+2.09	0.69	1.81	1.43	-0.34
	H_s	y = 1.00x+0.42	0.89	0.86	0.60	-0.43
	HS_a	y = 0.90x+0.55	0.81	1.03	0.73	-0.06
	JH	y = 0.73x+1.46	0.85	1.21	0.97	-0.25

K	y = 0.62x+1.67	0.70	1.72	1.35	0.23
L	y = 0.64x+0.68	0.71	2.31	1.83	1.62
ANN_all		0.95	0.54	0.33	0.13
ANN_T		0.85	0.89	0.61	-0.07
WA-ANN_all		0.77	1.17	0.86	-0.44
WA-ANN_T		0.75	1.22	0.91	-0.46

n: the number of CIMIS stations in each climate division. A and B: coefficients of the BC_{fao} model.

The ANN models can be ranked as ANN_all >ANN_T > WA-ANN_all > WA-ANN_T, based on their performance in the climate divisions. For the ANN models, MAE ranged from 0.31 mm d^{-1} (ANN_all in north coast drainage) to 0.91 mm d^{-1} (WA-ANN_T in the southeast desert basin). The MBE values ranged from -0.46 mm d^{-1} (WA-ANN_T in the southeast desert basin) to 0.30 mm d^{-1} (WA-ANN_T in north coast drainage). The R^2 values varied from 0.63 (WA-ANN_T in south coast drainage) to 0.96 (ANN_all in San Joaquin drainage). The mean ET_o values ranged from 2.24 mm d^{-1} (at CIMIS #193 in central coast drainage) to 5.42 mm d^{-1} (CIMIS # 200 in the southeast desert basin) for models ANN_T and ANN_all, respectively (Figure 4.8). Overall, the empirical and ANN models can be ranked as ANN_all > ANN_T > H_s >HS_a > BC_{fao} WA-ANN_all > WA-ANN_T > JH > BC > K > H > L, based on their performance in different climate divisions. ANN_T model, which used the same inputs as the HS_a model, resulted in a reduction of MAE ranging from 11% to 17% in all climate divisions, thus offering a promising alternative for accurate estimations of ET_o.

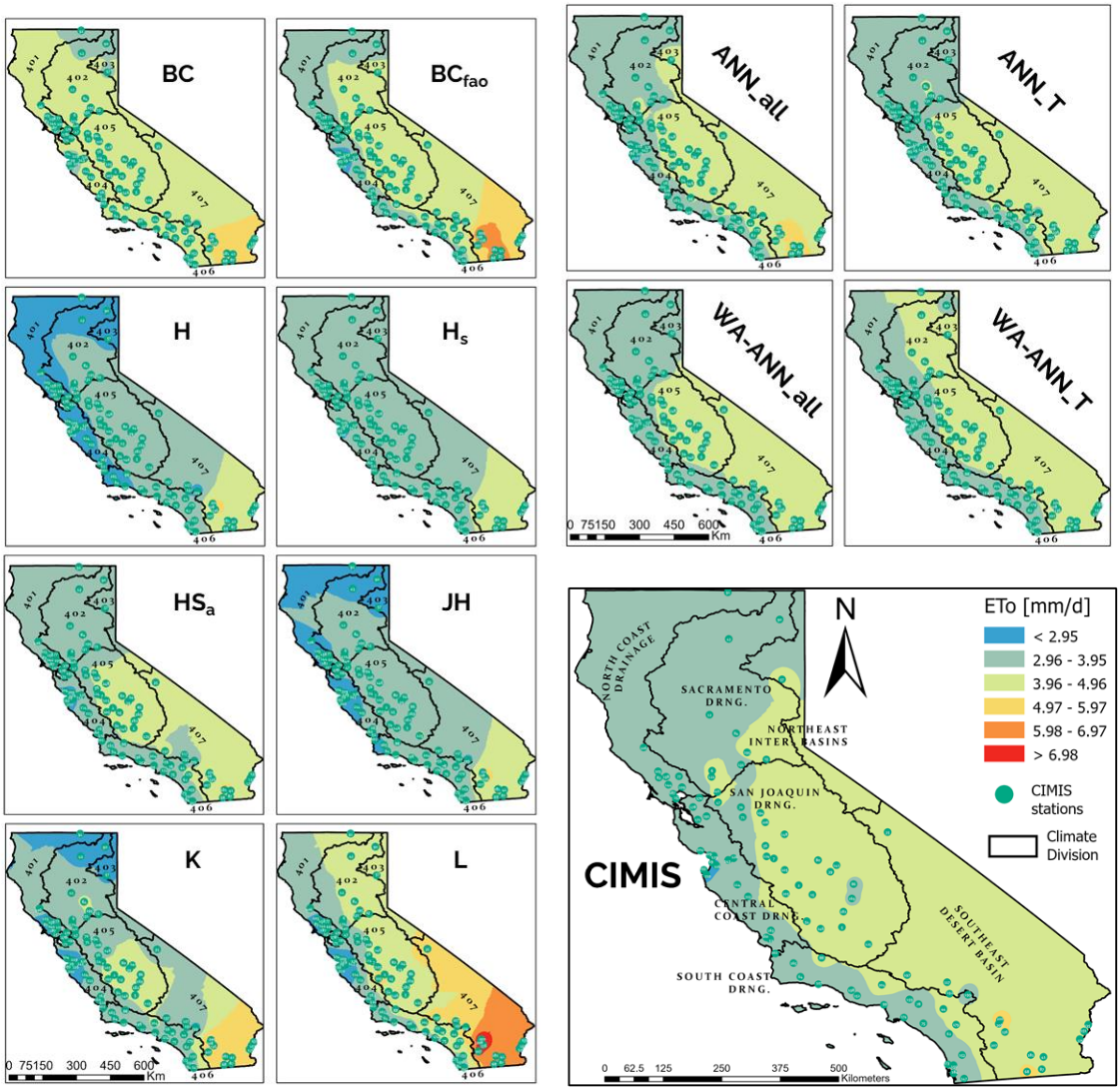


Figure 4.8. Comparison between the long-term ET_o maps obtained from CIMIS against the estimated maps by the 12 temperature-based empirical and ANN models evaluated in this study.

4.5. Conclusion

This study evaluated the performance of 8 empirical temperature-based and 4 Artificial neural network models (ANN) ET_o models at 101 active California Irrigation Management Information System (CIMIS) weather stations in California using more than 725,000 observations from 1985 to 2019. The ANN models outperformed the empirical

equations and showed high generalization ability. Our results suggest that using raw input data for neural network models is better than using the reconstructed signal obtained from wavelet transform. The Hargreaves and Hargreaves & Samani were the best performing empirical models across the seasons, years, and climate divisions. We recommend the application of the calibrated Hargreaves & Samani and the ANN model we developed in this study for accurate estimations of ET_o in data-scarce conditions in urban settings across California climate divisions. Both these models only require on-site air temperature measurements, which are typically collected by smart weather-based landscape irrigation controllers.

4.6. References

Adamala, S., 2018. Temperature based generalized wavelet-neural network models to estimate evapotranspiration in India. Inf. Process. Agric. 5, 149–155. https://doi.org/10.1016/j.inpa.2017.09.004

Adamowski, J., Sun, K., 2010. Development of a coupled wavelet transform and neural network method for flow forecasting of non-perennial rivers in semi-arid watersheds. J. Hydrol. 390, 85–91. https://doi.org/10.1016/j.jhydrol.2010.06.033

Alexandridis, A.K., Zapranis, A.D., 2013. Wavelet neural networks: A practical guide. Neural Networks 42, 1–27. https://doi.org/10.1016/j.neunet.2013.01.008

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration. FAO Irrig. Drain. Pap. No. 56.

Allen, R.G., Walter, I.A., Elliott, R.L., Howell, T.A., Itenfisu, D., Jensen, M.E., Snyder, R.L. (Eds.), 2005. The ASCE Standardized Reference Evapotranspiration Equation. Technical Committee on Standardization of Reference Evapotranspiration, Books. American Society of Civil Engineers, Reston, VA. https://doi.org/10.1061/9780784408056

Blaney, H.F., Criddle, W.D., 1950. Determining water requirements in irrigated areas from climatological and irrigation data. Soil Conserv. Serv. Tech. Pap. 96, Soil Conserv. Serv. US Dep. Agric. Washingt.

Cardenas, B., Dukes, M.D., Breder, E., Torbert, J.W., 2021. Long-term performance of smart irrigation controllers on single-family homes with excess irrigation . AWWA Water Sci. 3, 1–14. https://doi.org/10.1002/aws2.1218

Chen, Z., Zhu, Z., Jiang, H., Sun, S., 2020. Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. J. Hydrol. 591, 125286. https://doi.org/10.1016/j.jhydrol.2020.125286

Cobaner, M., 2013. Reference evapotranspiration based on Class A pan evaporation via wavelet regression technique. Irrig. Sci. 31, 119–134. https://doi.org/10.1007/s00271-011-0297-x

Davis, S.L., Dukes, M.D., 2010. Irrigation scheduling performance by evapotranspiration-based controllers. Agric. Water Manag. 98, 19–28. https://doi.org/10.1016/j.agwat.2010.07.006

Djaman, K., Balde, A.B., Sow, A., Muller, B., Irmak, S., N'Diaye, M.K., Manneh, B., Moukoumbi, Y.D., Futakuchi, K., Saito, K., 2015. Evaluation of sixteen reference evapotranspiration methods under sahelian conditions in the Senegal River Valley. J. Hydrol. Reg. Stud. 3, 139–159. https://doi.org/10.1016/j.ejrh.2015.02.002

Doorenbos, J., Pruitt, W.O., 1977. Guidelines for predicting crop water requirements. FAO Irrig. Drain. Pap. 24, 144.

Evrendilek, F., 2014. Assessing neural networks with wavelet denoising and regression models in predicting diel dynamics of eddy covariance-measured latent and sensible heat fluxes and evapotranspiration. Neural Comput. Appl. 24, 327–337. https://doi.org/10.1007/s00521-012-1240-7

Falamarzi, Y., Palizdan, N., Huang, Y.F., Lee, T.S., 2014. Estimating evapotranspiration from temperature and wind speed data using artificial and wavelet neural networks (WNNs). Agric. Water Manag. 140, 26–36. https://doi.org/10.1016/j.agwat.2014.03.014

Fan, J., Yue, W., Wu, L., Zhang, F., Cai, H., Wang, X., Lu, X., Xiang, Y., 2018. Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China. Agric. For. Meteorol. 263, 225–241. https://doi.org/10.1016/j.agrformet.2018.08.019

Gabriela Arellano, M., Irmak, S., 2016. Reference (Potential) Evapotranspiration. I: Comparison of Temperature, Radiation, and Combination-Based Energy Balance Equations in Humid, Subhumid, Arid, Semiarid, and Mediterranean-Type Climates. J. Irrig. Drain. Eng. 142, 04015065. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000978

Haghverdi, A., Reiter, M., Sapkota, A., Singh, A., 2021a. Hybrid Bermudagrass and Tall Fescue Turfgrass Irrigation in Central California: I. Assessment of Visual Quality, Soil Moisture and Performance of an ET-Based Smart Controller. Agronomy 11, 1666. https://doi.org/10.3390/agronomy11081666

Haghverdi, A., Singh, A., Sapkota, A., Reiter, M., Ghodsi, S., 2021b. Developing irrigation water conservation strategies for hybrid bermudagrass using an evapotranspiration-based smart irrigation controller in inland southern California. Agric. Water Manag.

245, 106586. https://doi.org/10.1016/j.agwat.2020.106586

Hamon, W.R., 1961. Estimating Potential Evapotranspiration. J. Hydraul. Div. Am. Soc. Civ. Eng. 871, 107–120. https://doi.org/10.1061/jyceaj.0000749

Hargreaves, G.H., 1975. Moisture availability and crop production. Trans. ASAE 18, 980–984.

Hargreaves, G.H., Allen, R.G., 2003. History and Evaluation of Hargreaves Evapotranspiration Equation. J. Irrig. Drain. Eng. 129, 53–63. https://doi.org/10.1061/(ASCE)0733-9437(2003)129:1(53)

Hargreaves, G.H., Samani, Z.A., 1985. Reference Crop Evapotranspiration from Temperature. Appl. Eng. Agric. 1, 96–99. https://doi.org/10.13031/2013.26773

Hidalgo, H.G., Cayan, D.R., Dettinger, M.D., 2005. Sources of variability of evapotranspiration in California. J. Hydrometeorol. 6, 3–19. https://doi.org/10.1175/JHM-398.1

Hope, A.S., Evans, S.M., 1993. Estimating Reference Evaporation in the Central Valley of California using the Linacre Model. Water Resour. Bull. Am. Water Resour. Assoc. 28, 695–702.

Jensen, M.E., Haise, H.R., 1963. Estimating Evapotranspiration from Solar Radiation. J. iIrigation Drain. Div. 89, 15–41.

Kharrufa, N.S., 1985. Simplified equation for evapotranspiration in arid regions. Beiträge zur Hydrol. 5, 39–47.

Kişi, Ö., 2010. Evapotranspiration modeling using a wavelet regression model. Irrig. Sci. https://doi.org/10.1007/s00271-010-0232-6

Kisi, O., Alizamir, M., 2018. Modelling reference evapotranspiration using a new wavelet conjunction heuristic method: Wavelet extreme learning machine vs wavelet neural networks. Agric. For. Meteorol. 263, 41–48. https://doi.org/10.1016/j.agrformet.2018.08.007

Kukal, M., Irmak, S., Walia, H., Odhiambo, L., 2020. Spatio-temporal Calibration of Hargreaves-Samani Model to Estimate Reference Evapotranspiration across U.S. High Plains. Agron. J. 1–17. https://doi.org/10.1002/agj2.20325

Lee, G., Gommers, R., Waselewski, F., Wohlfahrt, K., O'Leary, A., 2019. PyWavelets: A Python package for wavelet analysis. J. Open Source Softw. 4, 1237. https://doi.org/10.21105/joss.01237

Linacre, E.T., 1977. A simple formula for estimating evaporation rates in various climates, using temperature data alone. Agric. Meteorol. 18, 409–424. https://doi.org/10.1016/0002-1571(77)90007-3

Liu, X., Xu, C., Zhong, X., Li, Y., Yuan, X., Cao, J., 2017. Comparison of 16 models for reference crop evapotranspiration against weighing lysimeter measurement. Agric. Water Manag. 184, 145–155. https://doi.org/10.1016/j.agwat.2017.01.017

Long, H., Shuai, X., Lei, Q., Zhang, R., 2013. Spatiotemporal Distribution of Calibration Coefficients of Hargreaves Equation for Estimating Potential Evapotranspiration in Mainland China. J. Irrig. Drain. Eng. 139, 293–299. https://doi.org/10.1061/(asce)ir.1943-4774.0000534

Muhammad, M.K.I., Nashwan, M.S., Shahid, S., Ismail, T. bin, Song, Y.H., Chung, E.S., 2019. Evaluation of empirical reference evapotranspiration models using compromise programming: A case study of Peninsular Malaysia. Sustain. 11. https://doi.org/10.3390/su11164267

Muniandy, J.M., Yusop, Z., Askari, M., 2016. Evaluation of reference evapotranspiration models and determination of crop coefficient for Momordica charantia and Capsicum annuum. Agric. Water Manag. 169, 77–89. https://doi.org/10.1016/j.agwat.2016.02.019

Partal, T., 2009. Modelling evapotranspiration using discrete wavelet transform and neural networks. Hydrol. Process. 23, 3545–3555. https://doi.org/10.1002/hyp.7448

Pruitt, W.O., Doorenbos, J., 1977. Empirical calibration, a requisite for evapotranspiration formulae based on daily or longer mean climate data., in: Int. Comm. on Irrig. and Drain., Budapest, Hungary. p. 20.

Serena, M., Velasco-Cruz, C., Friell, J., Schiavon, M., Sevostianova, E., Beck, L., Sallenave, R., Leinauer, B., 2020. Irrigation scheduling technologies reduce water use and maintain turfgrass quality. Agron. J. 112, 3456–3469. https://doi.org/10.1002/agj2.20246

Tabari, H., Grismer, M.E., Trajkovic, S., 2013. Comparative analysis of 31 reference evapotranspiration methods under humid conditions. Irrig. Sci. 31, 107–117. https://doi.org/10.1007/s00271-011-0295-z

Temesgen, B., Allen, R.G., Jensen, D.T., 1999. Adjusting Temperature Parameters to Reflect Well-Watered Conditions. J. Irrig. Drain. Eng. 125, 26–33. https://doi.org/10.1061/(ASCE)0733-9437(1999)125:1(26)

Temesgen, B., Eching, S., Davidoff, B., Frame, K., 2005. Comparison of Some Reference Evapotranspiration Equations for California. J. Irrig. Drain. Eng. 131, 73–84. https://doi.org/10.1061/(ASCE)0733-9437(2005)131:1(73)

TensorFlow Developers, 2021. TensorFlow. https://doi.org/10.5281/ZENODO.5189249

Trabucco, A., Zomer, R.J., 2018. Global Aridity Index and Potential Evapo-Transpiration (ET0) Climate Database v2. CGIAR Consortium for Spatial Information (CGIAR-CSI). Published online, available from the CGIAR-CSI GeoPortal at https://cgiarcsi.community.

Traore, S., Luo, Y., Fipps, G., 2016. Deployment of artificial neural network for short-term forecasting of evapotranspiration using public weather forecast restricted messages. Agric. Water Manag. 163, 363–379. https://doi.org/10.1016/j.agwat.2015.10.009

Traore, S., Wang, Y.M., Kerh, T., 2010. Artificial neural network for modeling reference

evapotranspiration complex process in Sudano-Sahelian zone. Agric. Water Manag. 97, 707–714. https://doi.org/10.1016/j.agwat.2010.01.002

Xu, C.Y., Singh, V.P., 2001. Evaluation and generalization of temperature-based methods for calculating evaporation. Hydrol. Process. 15, 305–319. https://doi.org/10.1002/hyp.119

Zhang, Q., Benveniste, A., 1992. Wavelet networks. IEEE Trans. Neural Networks 3, 889–898. https://doi.org/10.1109/72.165591

Chapter 5. Autonomous Hybrid Bermudagrass Recycled Water Irrigation Management Using a Smart Soil Moisture Sensor-Based Controller

Abstract

A three-year (2019–2021) irrigation research trial was conducted to evaluate the response of bermudagrass to soil moisture sensor (SMS) based deficit irrigation treatments and assess the efficacy of Acclima smart controller for autonomous irrigation scheduling in southern California. The irrigation levels applied were based on upper and lower soil moisture thresholds and ranged between 29% and 109% of reference ET (ET_o). The controller was also programmed to implement two irrigation frequencies, i.e., restricted (3d/week) and on-demand (7 d/week) irrigation. Normalized difference vegetation index (NDVI) and turf temperature data were collected weekly during the summer irrigation season to evaluate the response of bermudagrass to irrigation treatments. We also measured soil salinity, sodium adsorption ratio (SAR), and infiltration rate before and after the summer irrigation season. There was a significant effect of irrigation levels and irrigation frequency restrictions on NDVI for all three years, while turf temperature was significantly affected in 2020 and 2021. On-demand irrigation resulted in 3.4% lower turf temperatures over the study period compared to restricted irrigation while maintaining slightly better visual quality. Although soil salinity oscillated with the seasonal cycles because of leaching observed due to winter rainfall, both salinity and SAR increased as the study progressed. By the end of the research period, turf quality was below the acceptable NDVI of 0.5, suggesting that bermudagrass generally do not perform well when deficit-irrigated with recycled water on a long-term basis in a semi-arid climate. Further investigation is needed

to substantiate the use of SMS based autonomous deficit irrigation scheduling when recycled water is used.

5.1. Introduction

The availability of recycled water plays an essential role in fulfilling landscape irrigation demand in southern California, considering the treatment cost compared to other alternatives like seawater desalination and imported surface waters. A state-wide survey in California in 2015 indicated, approximately 31%, 18%, and 8% of the recycled water, was utilized for agriculture, landscape, and golf course irrigation, respectively (SWRCB, 2015). Opportunities for increasing the use of recycled water to irrigate urban landscapes exist as the proportion of total recycled urban water use in southern California varied from 26% in Santa Ana, 38% for Los Angeles, to 94% for the San Diego region in 2015.

High concentrations of salt present in recycled water could negatively impact plant growth and soil health (Gonçalves et al., 2007; Qian and Mecham, 2005). This problem is critical in arid and semi-arid climates of Southern California, where low precipitation may not adequately leach soluble salts from the root zone in some years. In addition, a higher concentration of ions such Na^{+} vs. $Ca^{2+}$and Mg^{2+} can increase the sodium adsorption ratio (SAR) in soil, which can cause permeability issues (Gao et al., 2021; Zalacáin et al., 2019).

Advancements in smart irrigation technologies have led to affordable smart irrigation controllers. Homeowners prefer smart irrigation controllers to conventional automated systems because of the potential savings on annual water bills associated with these products (Khachatryan et al., 2019). Soil moisture sensor (SMS) based smart irrigation controllers have shown significant water saving potential (Blonquist et al., 2006; Cardenas-

Lailhacar and Dukes, 2012; Cardenas et al., 2021; Qualls et al., 2001). For example, SMS-based irrigation controllers reduced irrigation by 11-53 % under relatively dry conditions compared to a time-based schedule (McCready et al., 2009). Cardenas and Dukes (2016) showed the potential to save water using SMSs for recycled water irrigation with a salinity of around 0.75dS/m. However, knowledge on using smart irrigation controllers to implement deficit irrigation strategies using recycled water in turfgrass is lacking. Much of the scientific research on the application of SMS-based landscape irrigation scheduling has been done in humid regions, wherein the main focus has been to avoid over-irrigation when rainfall is abundant. Currently, information is lacking about applying smart irrigation technologies leading to deficit recycled water irrigation strategies in arid regions such as Southern California. Consequently, this study was conducted to: (1) determine the impact on hybrid bermudagrass quality under different deficit irrigation strategies with recycled water in Orange County, CA, (2) Determine the effect of recycled water irrigation on soil salinity (ECe), sodium adsorption ratio (SAR), and soil infiltration rate, and (3) Evaluate the performance of an SMS-based smart irrigation controller for efficient automatic recycled water irrigation management.

5.2. Material and methods

5.2.1 Study site

A three-year (2019–2021) hybrid bermudagrass (*Cynodon dactylon x C. transvaalensis* Tilfgreen 328) irrigation research study was conducted at UC ANR South Coast Research and Extension Center (SCREC) in Irvine, California (33°41'20.9"N 117°43'23.3"W). Hybrid bermudagrass made up around 34% of all maintained turfgrass

acreage in 2015, being the most heavily used turfgrass species in the nation (GCSAA, 2017). Warm-season grasses such as hybrid bermudagrass are also considered the best choice for salinity and drought tolerance (Dean et al., 1996; Marcum, 2006).

The soil at the research site is classified as a San Emigdio fine sandy loam with a soil volumetric water content of 18% at 33 kPa (websoilsurvey.sc.egov.usda.gov). The daily reference evapotranspiration (ET_o) and precipitation data from CIMIS (California Irrigation Management Information System) weather station#75 situated about 65 meters from experimental site are shown in Figure 5.1.

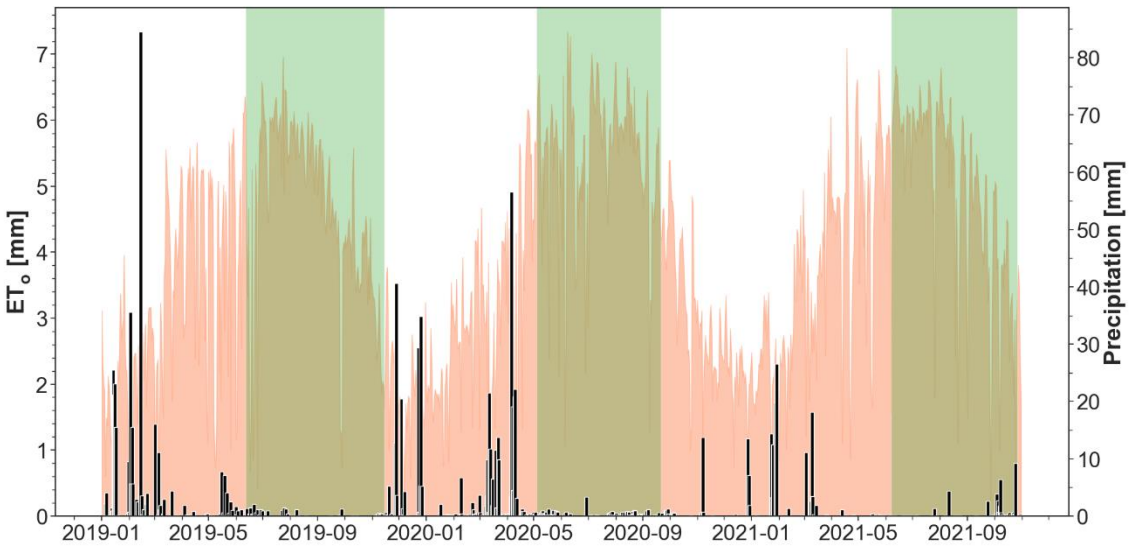


Figure 5.1. Daily reference evapotranspiration (ET_o) and precipitation during the experimental period obtained from the CIMIS station #75. Shaded green region represents the period when NDVI and turf temperature measurements were taken during each irrigation season.

5.2.2. Irrigation system and instrumentation

A research site consisting of 48 research plots sized 3.7 m × 3.7 m with approximately 60–90 cm border between the plots was built in July 2018 accommodate 12 treatments

replicated four times in a randomized block design. Each plot was irrigated by four quarter-circle pop-up Hunter Pro-Spray PRO12Q sprinkler heads (Hunter Industries, Inc., San Marcos, CA) with an operating pressure range and flow rate of 138–275 kPa and 2.04–2.95 lm^{-1}, respectively. A flow meter was installed to precisely record irrigation runtimes and water application on each zone.

An Acclima CS3500 smart irrigation controller (Acclima, Inc. Idaho, USA) was installed, and all solenoid valves were wired to the controller. One block was instrumented with soil moisture sensors (Acclima TDT sensors) and were used for irrigation scheduling by connecting the sensors to the controller. Each treatment received a separate soil moisture sensor, and the controller was wired to irrigate all four replications of the treatment at the same time. These sensors were found to provide reliable moisture measurements under saline water irrigation (Cardenas-Lailhacar and Dukes, 2014).

The controller was programmed to irrigate overnight (between 10 pm to 6 am) to minimize evaporative losses and wind drift. Total daily irrigation run time was divided into multiple irrigation applications of 5 minutes each to avoid runoff and allow a 1-hour percolation time between irrigation applications. The same irrigation programming was maintained throughout the trial to mimic the real-world irrigation application, i.e., the set and forget mentality of the users (Bremer et al., 2013). An electrical conductivity sensor ES-2 (Meter group Inc.) was installed at a common irrigation inlet to continuously monitor recycled water salinity, and data was recorded every 30 minutes using a ZL6 data logger (Meter group Inc.).

5.2.3 Irrigation trial and treatment design

A total of 48 research plots sized 3.7 m × 3.7 m with approximately 60–90 cm border between the plots were planted with hybrid bermudagrass sod in August 2018 and were under non-limiting irrigation for several months until the trial started in June 2019 and was terminated in October 2021. Each plot was irrigated by four quarter-circle pop-up sprinkler heads controlled by a solenoid valve for independent irrigation control. A catch-can irrigation uniformity test was performed on the plots in February 2019 following the ANSI/ASABE S626 standard method (ANSI/ASABE, 2016). The catch can test performed at the start of the experiment resulted in DU_{LH} and CU values of 0.85 and 85%, respectively. We followed standard cultural practices to maintain the plots throughout the experimental periods.

Table 5.1 shows the 12 irrigation treatments implemented in this study consisting of 6 soil moisture thresholds ×2 irrigation frequencies. The soil moisture thresholds were centered around the field capacity of soil which was estimated for each soil treatment separately as the soil moisture recorded by sensors approximately 18 hours after receiving heavy rainfall in mid-February, 2019. The Acclima CS3500 smart irrigation controller required specific lower and upper thresholds for each treatment for irrigation scheduling. A lower limit is the soil moisture level where the smart controller will trigger irrigation if the moisture level in the root zone drops below that level, and an upper level is the soil moisture level where the irrigation is shut off if the moisture level exceeds this limit (Grabow et al., 2013). Idea is based on the management allowed depletion (MAD), which is the fraction of plant available water that is allowed to be depleted from the soil before

plant begin to stress. Plant available water is the water is the difference between field capacity and permanent wilting point. Permanent wilting point refers to the amount of water in soil that is unavailable to the plant. The recommended level is usually field capacity set as the upper limit and the 75% of field capacity soil moisture set as the lower limit, which also adheres to the guidance by the manufacturer. A total of 12 treatments were implemented with different lower and upper soil moisture limits and two irrigation frequencies, i.e., restricted irrigation (3-day/week) and on-demand irrigation (7-day/week), as shown in table 5.1.

Table 5.1. Treatments imposed with the Soil moisture sensors in the instrumented plots.

Treatment	LL (soil moisture)	UL (soil moisture)	Irrigation Frequency
T1	75%FC (24.1)	FC (32.1)	Restricted
T2	65%FC (18.8)	FC (28.9)	“
T3	65%FC (20.6)	FC-10% (28.6)	“
T4	55%FC (16.1)	FC (29.2)	“
T5	55%FC (15.0)	FC-20% (21.8)	“
T6	75%FC (23.9)	FC+10% (35.0)	“
T7	75%FC (19.8)	FC (26.4)	On-demand
T8	65%FC (20.0)	FC (30.8)	“
T9	65%FC (19.6)	FC-10% (27.1)	“
T10	55%FC (17.8)	FC (32.3)	“
T11	55%FC (16.6)	FC-20% (24.2)	“
T12	75%FC (22.4)	FC+10% (32.9)	“

FC: Field Capacity; LL: Lower limit; UL: upper limit

5.2.4 Handheld remote sensing data

The handheld data were collected about once a week close to the solar noon on non-cloudy days, 18 times in 2019 (from June 12 to November 15), 20 times in 2020 (from May 4 to September 21), and 21 times in 2021(from June 7 to October 27). Data was collected by hovering the handheld sensors over the plots by avoiding the edges and

average for each plot was recorded. We focused on the summer months since water demand is high due to high evapotranspiration demand (Balling et al., 2008), and because bermudagrass go dormant over the winter. Fluke 64 max Infrared thermometer (Fluke Corporation, Everett, WA) was used to measure turf temperature and GreenSeeker handheld crop sensor (Trimble Inc., Sunnyvale, CA) for NDVI measurements. NDVI sensor emits brief bursts of red and infrared light and then measures the amount of each type of light that is reflected from the plant. NDVI is sensitive to sparse vegetation, which makes it ideal for use in turfgrass. The data were used to assess the effect of different irrigation scenarios on turfgrass health. NDVI equal to 0.5 was considered the minimum threshold to maintain an acceptable quality of bermudagrass for residential areas based on our recent hybrid bermudagrass fields research trials in southern and central California (Haghverdi et al., 2021a, 2021b).

NDVI collected by the spectral sensors is given as:

$$NDVI = \frac{NIR - R}{NIR + R} \quad (1)$$

where *NIR* and *R* correspond to the spectral measurements from the near-infrared and red regions of the electromagnetic spectrum, respectively.

5.2.5 Soil sampling and Infiltration data collection

Soil samples were collected biannually, once before the irrigation season and after the irrigation season from 0-60 cm depth with 15-cm increments. Samples were collected on May 1 and October 14, in 2019, on May 6 and October 13, in 2020, and on April 23 and October 11, in 2021. We collected samples from all the plots, and repetitions from the same

treatment were mixed, before analyzing in the lab. Soil samples were air-dried in the oven at 55^{o}C and were ground to pass through a 2 mm sieve. Saturated soil pastes were prepared with deionized water using ~200 g of air-dried soil and were allowed to stand overnight prior to vacuum filtration (Rhoades, 1982). Soil salinity (EC_e) was measured from the paste extract using Orion Star™ A212 conductivity meter with Orion™ DuraProbe™ 4-Electrode Conductivity Cell (Thermo Fisher Scientific Inc.). Sodium Adsorption Ratio (SAR) was also measured, giving information on the comparative concentration of Na^{+}, Ca^{2+}, and Mg^{2+} in the soil solution. It is defined as follows:

$$SAR = \frac{[Na]}{\sqrt{\frac{([Ca] + [Mg])}{2}}} \quad (2)$$

where concentration is expressed in milliequivalents per liter (meq/l).

Infiltration data using SATURO infiltrometer (METER Group, Inc. USA) was also collected before and after the irrigation season on the same dates when soil samples were collected. Infiltration measurements were taken from one block of plots, i.e., one replication of treatments. SATURO is an automated instrument that allows quick and reliable measurements of saturated hydraulic conductivity (K_{fs}) in the field. It utilizes the two-ponding head approach (Reynolds and Elrick, 1990), with some simplifications and modifications for three-dimensional flow from a single-ring infiltrometer (Nimmo et al., 2009).

Saturated Hydraulic conductivity (K_{fs}) is calculated by SATURO as follows:

$$\mathrm{Kfs} = \frac{\Delta(\mathrm{i1} - \mathrm{i2})}{\mathrm{D1} - \mathrm{D2}} \quad (3)$$

where D1 (cm) is the actual high-pressure head, D2 (cm) is the actual low-pressure head, Δ is 0.993d + 0.578b (cm), i1 is the infiltration rate at the high-pressure head, and i2 is infiltration rate at the low-pressure head. For Δ, d is infiltrometer insertion depth, and b is the infiltrometer radius.

5.2.6 Statistical analysis

Statistical analysis of turf temperature and NDVI measurements was performed using repeated-measures ANOVA with spatial power error structure in statistical software 'R'(R Core Team, 2021). Normality was assessed by the Shapiro-Wilk normality test, and we used F-test to test for homogeneity in variances. For statistical analysis of soil samples, the differences between the means of irrigation frequency were established using a t-test, and between different depths using the one-way ANOVA/one-sample Wilcoxon rank test. Tukey multiple comparison tests at a 0.05 significance level for post-hoc analysis.

5.3. Results

5.3.1. Irrigation applications

The cumulative ET_o in 2017, 2018, and 2019 were 1319, 1427, 1289 mm, respectively. The cumulative ET_o values in the irrigation season were 765, 768, and 699 mm in 2019, 2020, and 2021, respectively. Throughout the irrigation season, the average daily ET_o was 4.61, 5.44, and 4.89 mm day^{-1} in 2019, 2020, and 2021, respectively. The cumulative precipitation was negligible during the irrigation season, with only 22 mm, 33 mm, and 40

mm in 2019, 2020, and 2021, respectively. In comparison, the cumulative precipitation was 423 mm and 164 mm, respectively, in the non-irrigation seasons. Irrigation efficiency for each irrigation season was obtained from the flow measurements divided by the actual flow readings measured in a flow test performed in late 2020. Based on the calibrated flow meter readings, the actual applied irrigation varied from 66% to 109% ET_o in 2019, 51% to 104% ET_o in 2020, and 29% to 88% ET_o in 2021 (Table 5.2). Treatment 2, i.e., 65-90 FC irrigated 3d/week, resulted in higher irrigation application in 2020 and 2021. This could be attributed to the lower soil moisture very close to the threshold for this treatment, which triggered more irrigations by the controller. (Figures 5.2 – 5.4).

Table 5.2. Percent of ET_o applied to each treatment during the experimental period.

	2019		**2020**		**2021**	
Treatment	Restricted	On-demand	Restricted	On-demand	Restricted	On-demand
55-80FC	69 %	66 %	51 %	71 %	29 %	46 %
55-100FC	77 %	76 %	70 %	82 %	36 %	51 %
65-90FC	82 %	81 %	103 %	82 %	75 %	76 %
65-100FC	82 %	81 %	74 %	78 %	na	50 %
75-100FC	108 %	109 %	94 %	104 %	68 %	88 %
75-110FC	89 %	85 %	87 %	76 %	79 %	69 %

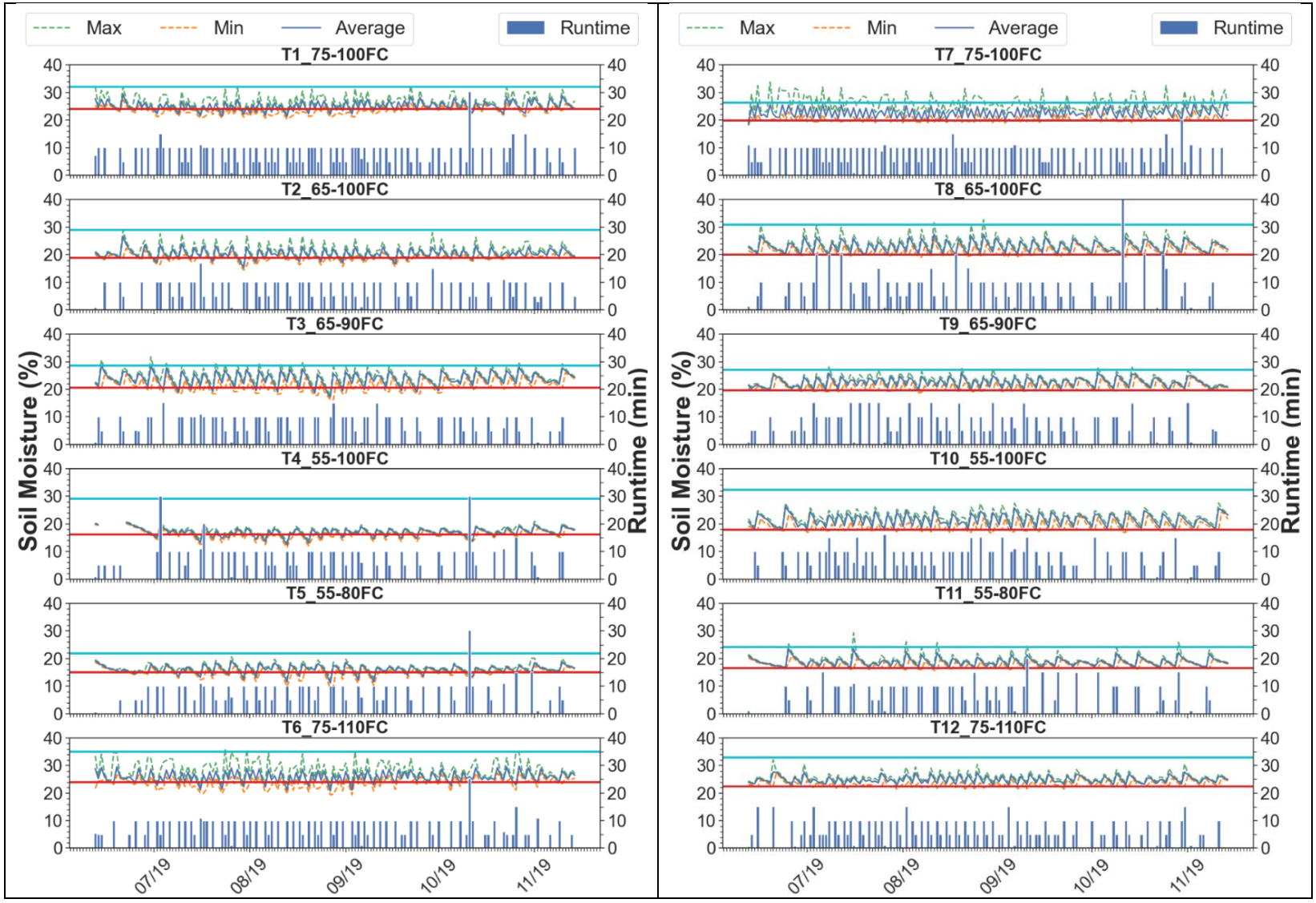

Max
Min
Average
Runtime
T1_75-100FC
T2_65-100FC
T3_65-90FC
T4_55-100FC
T5_55-80FC
T6_75-110FC
T7_75-100FC
T8_65-100FC
T9_65-90FC
T10_55-100FC
T11_55-80FC
T12_75-110FC
Soil Moisture (%)
Runtime (min)
0
10
20
30
40
07/19
08/19
09/19
10/19
11/19

Figure 5.2. Soil moisture and irrigation runtime for restricted (3d/week) and on-demand (7 d/week) irrigation treatments implemented for 2019. Solid red and dark blue lines are the lower (LL) and upper (UL) soil moisture thresholds levels, respectively.

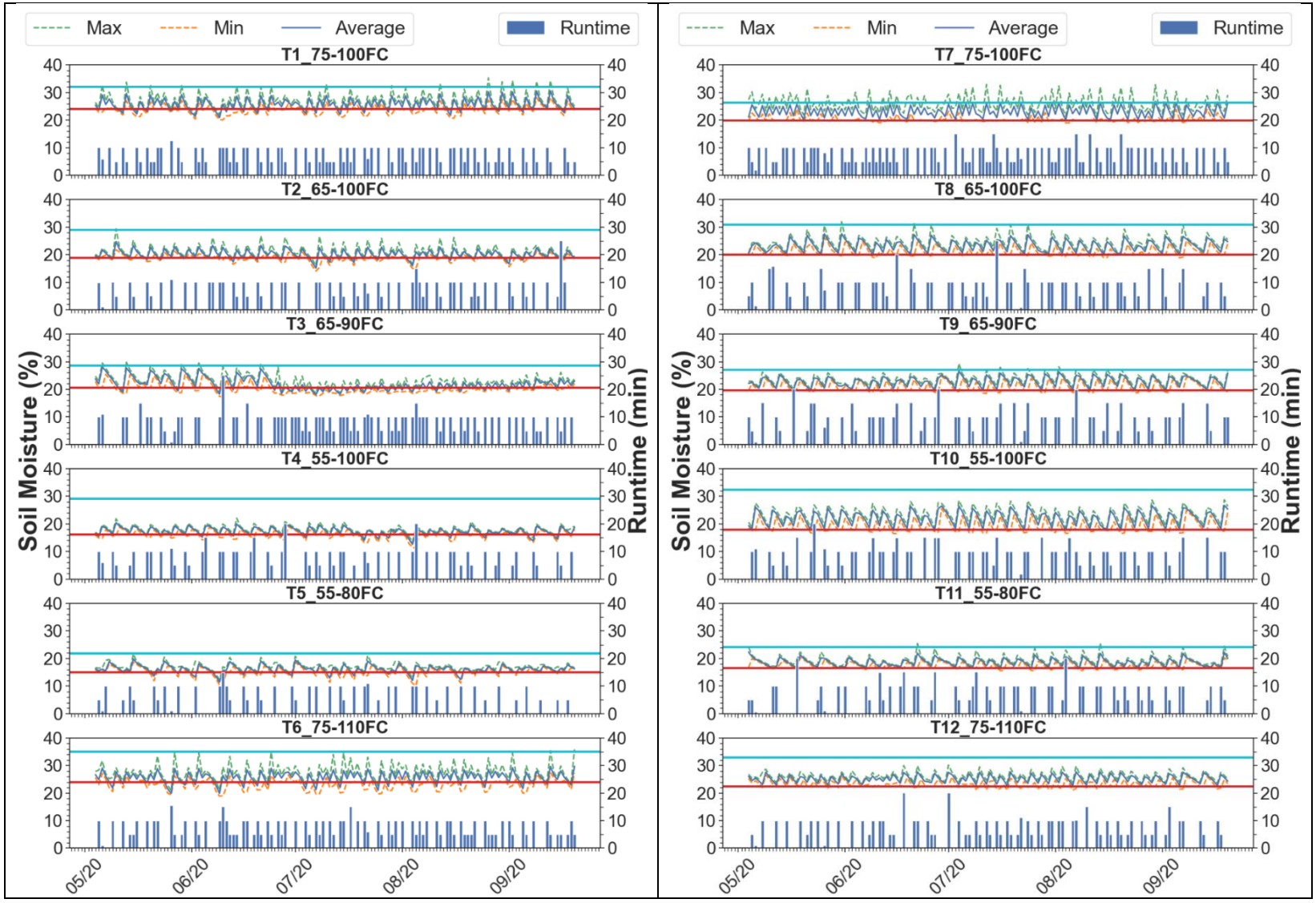

Max
Min
Average
Runtime
T1_75-100FC
T2_65-100FC
T3_65-90FC
T4_55-100FC
T5_55-80FC
T6_75-110FC
T7_75-100FC
T8_65-100FC
T9_65-90FC
T10_55-100FC
T11_55-80FC
T12_75-110FC
Soil Moisture (%)
Runtime (min)
0
10
20
30
40
05/20
06/20
07/20
08/20
09/20

Figure 5.3. Soil moisture and irrigation runtime for restricted (3d/week) and on-demand (7 d/week) irrigation treatments implemented for 2020. Solid red and dark blue lines are the lower (LL) and upper (UL) soil moisture thresholds levels, respectively.

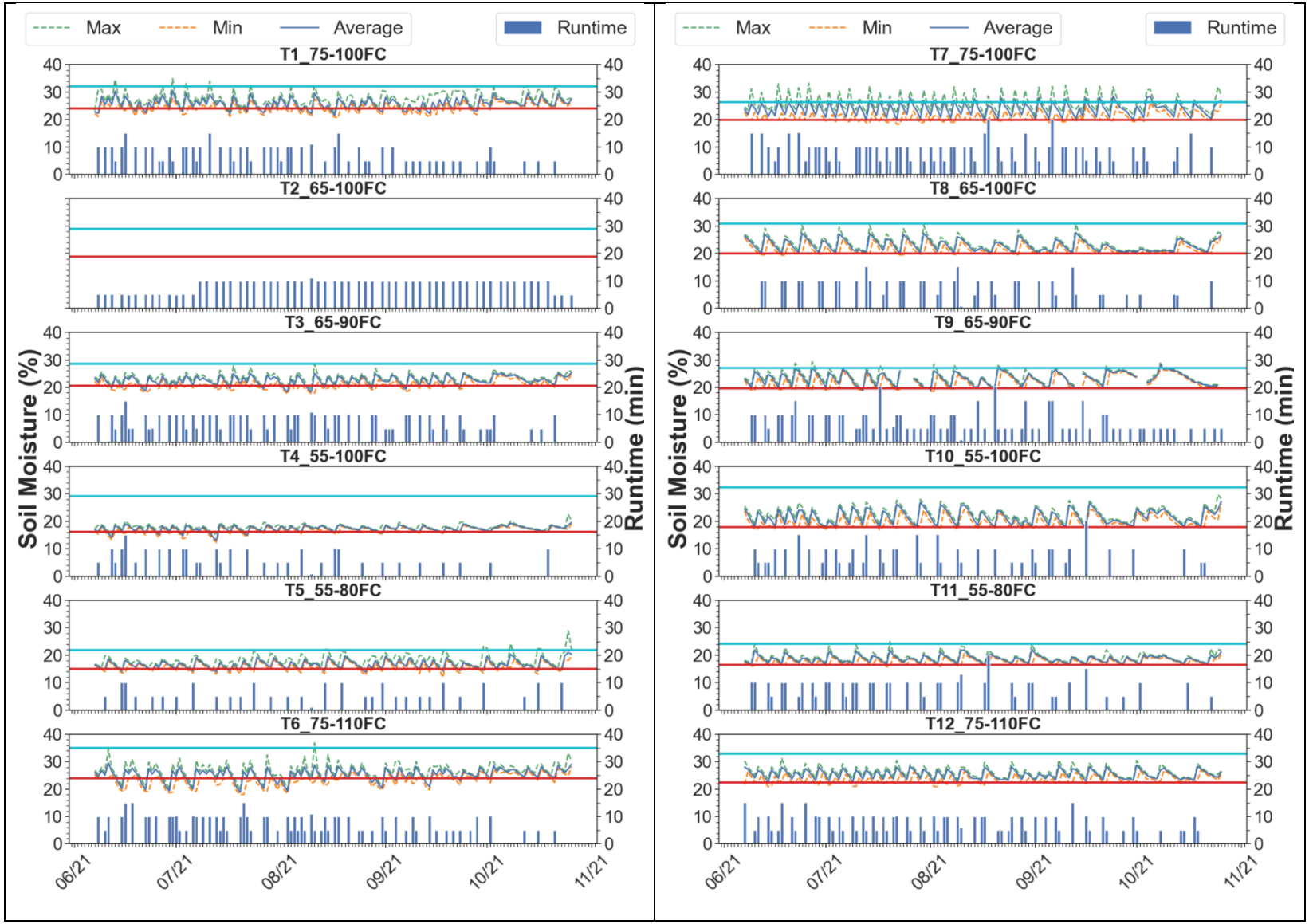

Max
Min
Average
Runtime
T1_75-100FC
T2_65-100FC
T3_65-90FC
T4_55-100FC
T5_55-80FC
T6_75-110FC
T7_75-100FC
T8_65-100FC
T9_65-90FC
T10_55-100FC
T11_55-80FC
T12_75-110FC
Soil Moisture (%)
Runtime (min)
0
10
20
30
40
06/21
07/21
08/21
09/21
10/21
11/21

Figure 5.4. Soil moisture and irrigation runtime for restricted (3d/week) and on-demand (7 d/week) irrigation treatments implemented for 2021. Solid red and dark blue lines are the lower (LL) and upper (UL) soil moisture thresholds levels, respectively.

We observed a good response from the smart controller to schedule irrigation based on the implemented thresholds (Figures 5.2-5.4). The analysis of irrigation runtime data showed that the actual irrigation frequencies for the restricted (3 d/week) irrigation treatments varied from 2.6 to 4.0 d/week in 2019, 2.4 to 4.4 d/week in 2020, and 1.7 to 3.4 d/week in 2021. For the on-demand (7 d/week) irrigation treatments, actual irrigation frequencies varied from 2.6 to 4.0 d/week in 2019, 3.0 to 4.7 d/week in 2020, and from 2.2 to 3.4 d/week in 2021. The reason for greater than 3d/week averages for restricted irrigation treatments is likely controller crossing midnight to complete the irrigation cycles, since it was programmed to irrigate late at night. As illustrated in Figures 5.2-5.4, the average soil moisture during the irrigation of 2019 varied from 16.0 to 25.8% for the restricted, and 18.7 to 24.8% for the on-demand irrigation treatments. In 2020, the average soil moisture varied from 16.1 to 25.7% for the restricted, and 18.6 to 25.2% for the on-demand irrigation treatments. In 2021, the soil moisture ranged from 16.6 to 25.7% for restricted, and 18.6 to 25.4% for the on-demand irrigation treatments.

5.3.2. NDVI and turf temperatures

Table 5.3 shows the results from the statistical analysis for the NDVI and turf temperature. Treatment T2 and T8 are not shown in Table 5.3 for year 2021, because a sensor malfunction impacted T2, but T7 is shown in Figure 5.5. The NDVI values ranged from 0.25 to 0.76 in 2019, 0.19 to 0.69 in 2020, and 0.15 to 0.62 in 2021. The irrigation levels significantly affected NDVI values in all three years ($p < 0.001$). The irrigation frequency restrictions also showed a significant impact on NDVI in 2019 ($p < 0.05$), 2020 ($p < 0.01$), and 2021 ($p < 0.001$). The interaction of irrigation levels and frequency had no

significant effect on NDVI in 2019, while the effect was significant in 2020 ($p < 0.001$) and 2021 (0.001). Figure 5.5 shows the dynamics of NDVI values over time across the irrigation treatments for 2019, 2020, and 2021. All treatments showed a steady decline in the NDVI during the data collection period in 2019. Most treatments stayed above the acceptable quality range, i.e., NDVI of 0.50 for an extended time. For 2020, NDVI stayed above the acceptable quality range majority of the time for the highest irrigated treatment 75-100FC for both irrigation frequencies. Other irrigation treatments saw a steady decline in the NDVI as the summer progressed, and the decline was steeper for the 3-day irrigation frequency treatments (Figure 5.5). For the year 2021, NDVI for most of the treatments stayed below the acceptable quality range except for 75-100FC 7-day irrigation frequency treatment, which had NDVI≥0.5 for half of the data collection period. The decline in the NDVI values was more apparent in the 3-day irrigation frequency treatments as the summer progressed.

Table 5.3. Statistical analysis of the bermudagrass response in terms of normalized difference vegetation index (NDVI) and turf temperature to irrigation treatments imposed in years 2019, 2020 and 2021 (each year was analyzed separately).

	2019		2020		2021	
Soil moisture treatments	**NDVI**	**Turf temperature**	**NDVI**	**Turf temperature**	**NDVI**	**Turf temperature**
55-80FC	0.553 a	36.3 a	0.41 a	40.5 c	0.279 a	47.2 b
55-100FC	0.563 ab	35.8 a	0.454 b	38.5 bc	0.266 a	45.7 b
65-90FC	0.616 cd	35.4 a	0.504 c	38 ab	0.416 c	38.6 a
65-100FC	0.595 bc	35.1 a	0.475 bc	37.8 ab	na	na
75-100FC	0.638 d	33.6 a	0.551 d	36.2 a	0.408 c	38.3 a
75-110FC	0.617 cd	34.4 a	0.488 c	37 ab	0.374 b	39.8 a
Irrigation Frequency						
Restricted	0.59 a	35.2 a	0.471 a	38.5 b	0.319 a	43.4 b
On-demand	0.604 b	35.1 a	0.49 b	37.5 a	0.378 b	40.4 a
p-value						
I	***	NS	***	***	***	***
F	*	NS	**	*	***	***
I x F	NS	NS	***	**	***	*
T	***	**	***	***	***	NS
I x T	NS	NS	***	NS	**	NS
F x T	NS	NS	NS	NS	NS	NS
I x F x T	NS	NS	*	NS	**	NS

NS, ***, **, and * are non-significant or significant at $p \leq 0.001$, 0.01, and 0.05, respectively. Means sharing a similar letter are not significantly different, based on Turkey's test at significance level (α) = 0.05. I, F, and T in the table refer to irrigation levels, frequency, and time (i.e., repeated measures of visual rating each year over time), respectively.

The turf temperature values ranged from 22 to 66.6 °C in 2019, from 21.7 to 64.4 °C in 2020, and from 18.5 to 69.8 °C in 2021. The irrigation levels showed no significant effect on the temperature in 2019, while there was a significant effect ($p < 0.001$) in 2020 and 2021. The irrigation frequency restrictions also had no significant effect on turf temperature in 2019, while there was a significant effect in 2020 ($p < 0.05$) and 2021 ($p < 0.001$), as the temperature was lower for the on-demand irrigation (7-days/week). On-demand irrigation resulted in a 0.3, 3, and 7% decrease in canopy temperatures compared

to the restricted irrigation treatments in 2019, 2020, 2021, respectively. The interaction of irrigation levels and frequency had no significant effect on turf temperature in 2019, while there was a significant effect in 2020 ($p < 0.01$) and 2021 (0.05). Figure 5.6 shows the dynamics of turfgrass temperature across the irrigation treatments for the years 2019, 2020, and 2021. All treatments showed similar trends during the data collection period in all three years, while the fluctuations were more pronounced as we progressed from 2019 to 2021 and were more perceptible in the restricted irrigation treatments (3day/week).

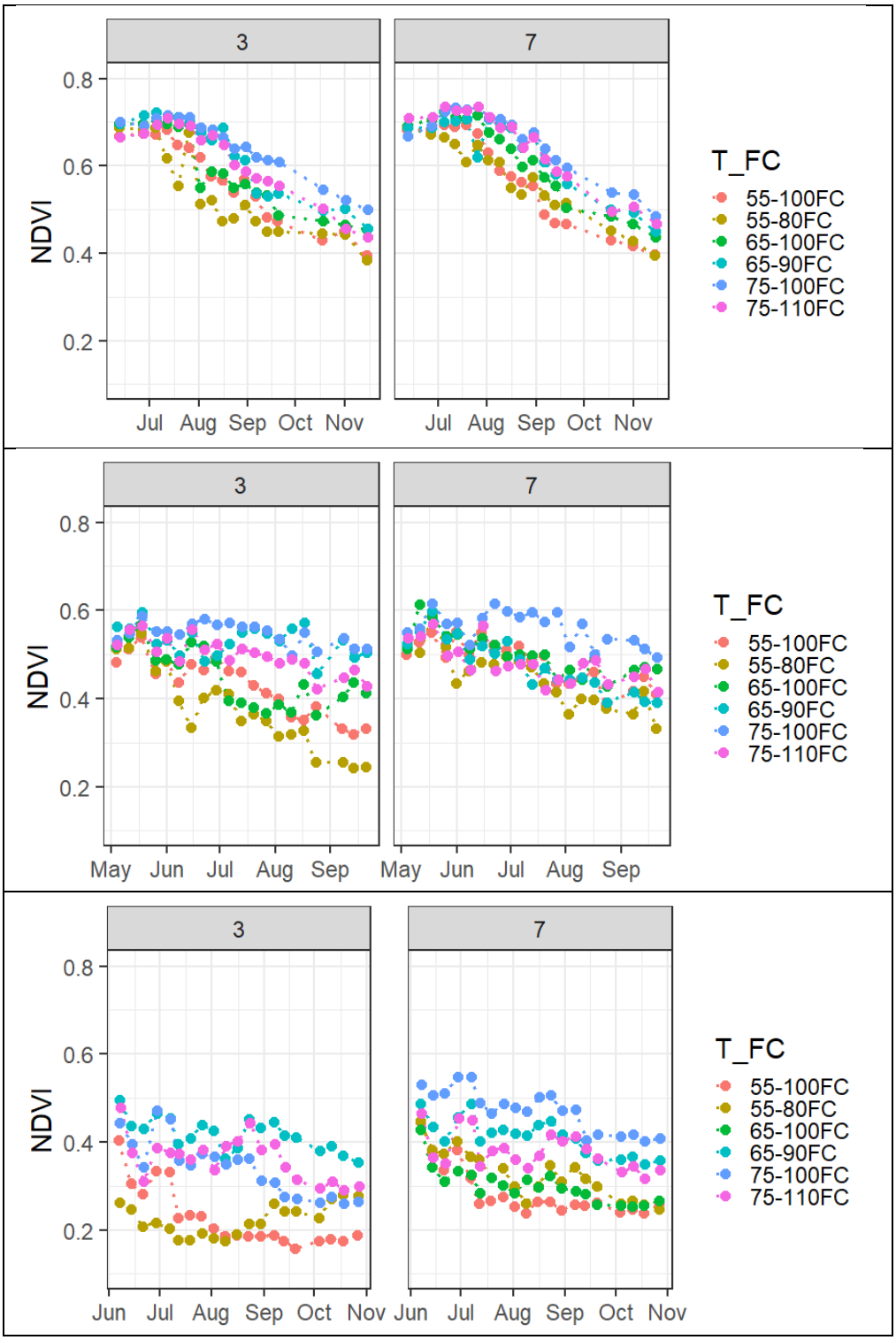


Figure 5.5. Changes in normalized difference vegetation index (NDVI) values over time across the irrigation treatments for the restricted (3 d/week) and on-demand (7 day/week) irrigation treatments imposed in 2019 (top), 2020 (middle), and 2021 (bottom).

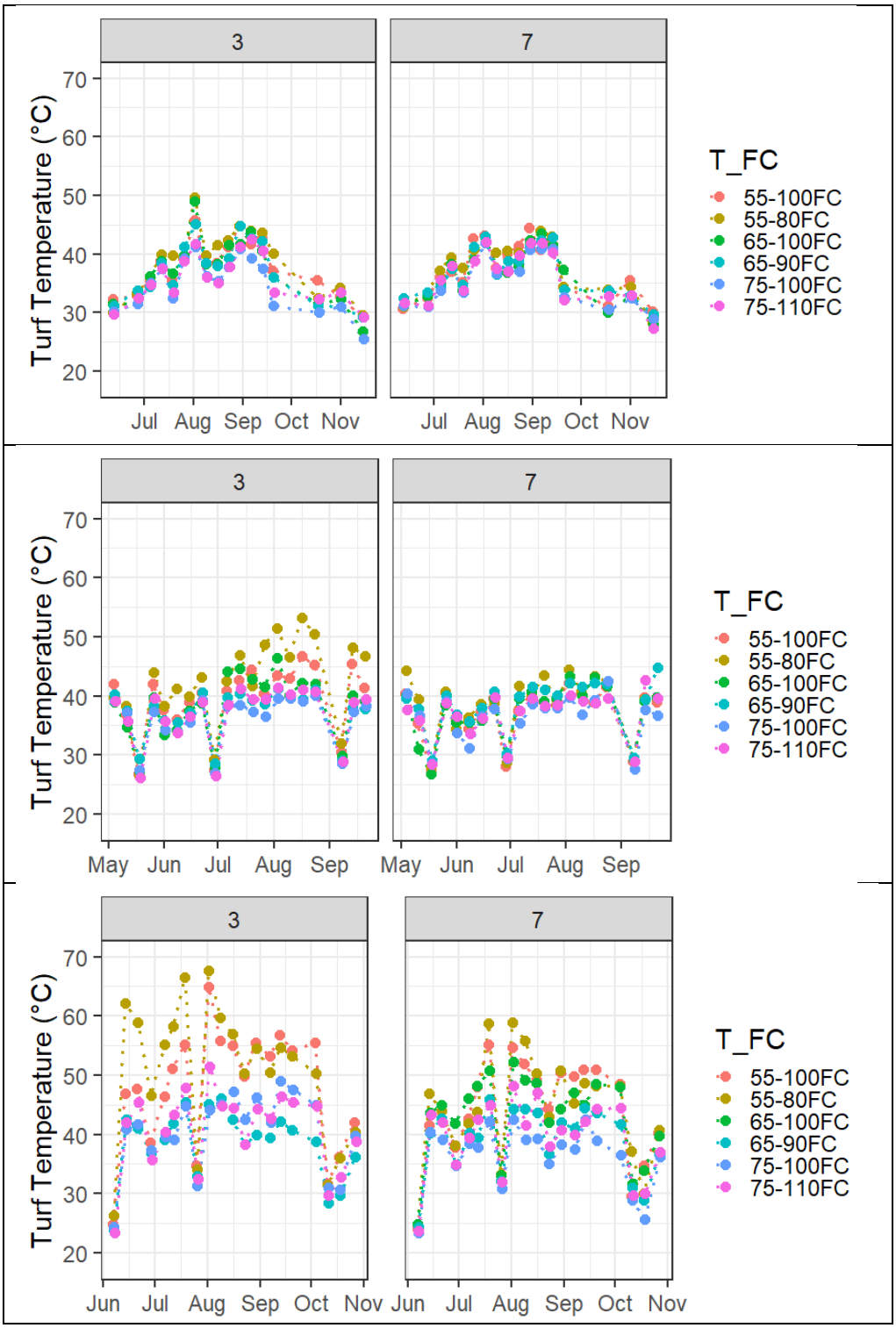


Figure 5.6. Changes in turfgrass temperature values over time across the irrigation treatments for the restricted (3 d/week) and on-demand (7 day/week) irrigation treatments imposed in 2019 (top), 2020 (middle), and 2021 (bottom).

5.3.3. Soil salinity and Infiltration

The daily average EC of the irrigation water (EC_w) was 1.18 dS/m and ranged from 0.92 and 1.65 dS/m during the experimental period. The soil salinity ranged from 0.79 ds/m to 1.66 ds/m in spring 2019 and 0.82 to 2.98 ds/m in fall 2019. In 2020, soil salinity ranged from 0.45 ds/m to 1.56 ds/m in spring and 0.52 to 3.90 ds/m in fall. In 2021, soil salinity ranged from 0.73 ds/m to 2.37 ds/m in spring and 1.13 to 4.72 ds/m in fall. As shown in Figure 5.7, the soil salinity increased in the fall season, especially in the shallow soil depths (0-30cm), indicating the accumulation of salts due to high evapotranspiration demand over summer. There was no significant difference between soil moisture threshold based treatments and irrigation frequencies. Spring collected samples were associated with lower soil salinity compared to fall samples. High rainfall before the spring 2020 sampling reduced salinity compared to the spring of 2019 and 2021(fig 1 & 6). Soil salinity distribution at different depths varied with the irrigation season as spring collected samples showed significantly higher accumulation at the 0-15 cm depth, while 15- 30 cm depth had substantially higher accumulation in the fall (Table 5.4).

The sodium adsorption ratio (SAR) ranged from 1.67 to 4.68 in spring 2020 and 1.68 to 5.51 in fall 2020. SAR ranged from 2.89 to 5.68 in spring and 3.26 to 7.76 in fall of 2020. Irrigation frequency had no significant difference on SAR, except in spring 2020, as shown in table 5.4. SAR distribution at different depths varied with the irrigation season as spring collected samples showed significantly higher values at the middle depths (30-45 cm), while shallow depths, particularly 0- 30 cm, had significantly higher values in the fall. Infiltration rate (IR) measurements were not as well defined as soil salinity and SAR for

each season. Mean IR values observed during the study period are shown in Table 5.5. Although there was no significant difference, the mean IR was higher for the on-demand irrigation treatments.

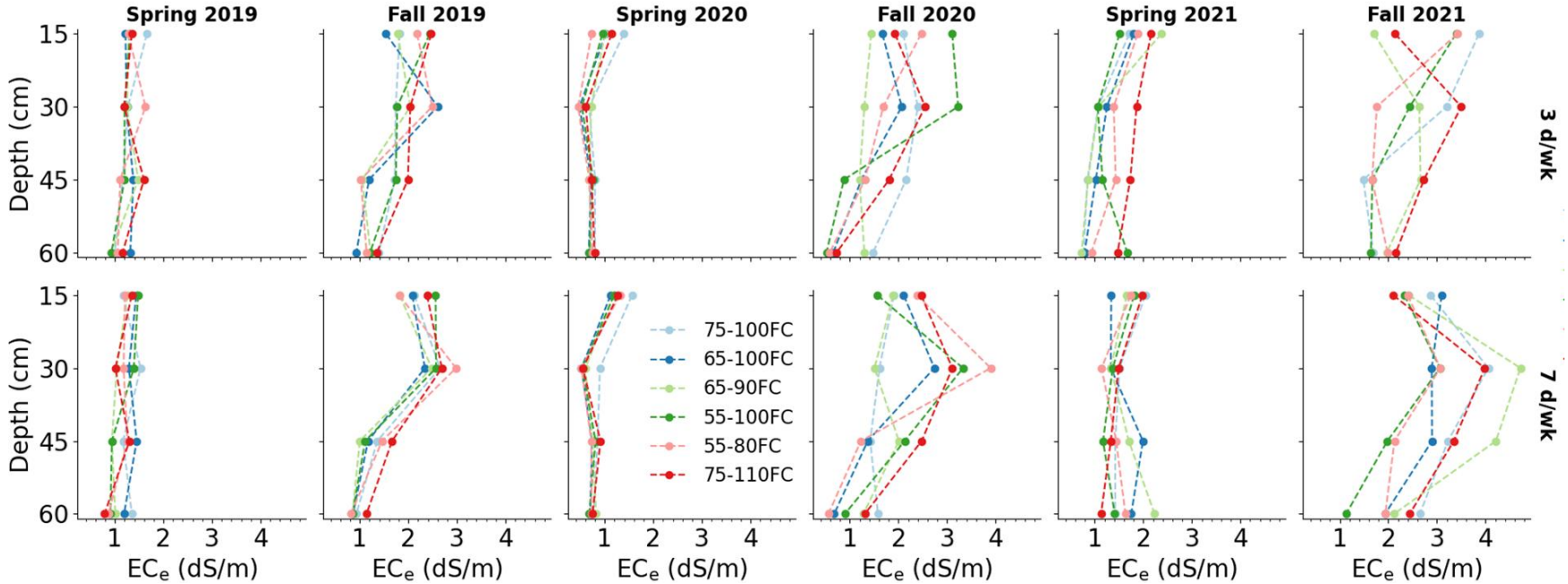


Figure 5.7. Soil salinity (EC_e) distribution in the soil profile from soil samples collected before (Spring) and after (Fall) of the summer irrigation seasons for 2019, 2020, and 2021.

Table 5.4. Statistical analysis of the soil samples collected before (Spring) and after (Fall) the irrigation season in 2019, 2020 and 2021.

		Depth (cm)				**Days**	
		0 - 15	15 - 30	30 - 45	45 - 60	3	7
5/1/2019	**ECe**	1.33 b	1.27 b	1.25 b	1.04 a	1.26 a	1.19 a
	SAR						
10/14/2019	**ECe**	2.08 b	2.37 b	1.37 a	1.06 a	1.70 a	1.74 a
	SAR						
5/6/2020	**ECe**	1.16 c	0.59 a	0.76 b	0.74 b	0.77 a	0.85 a
	SAR	2.19 a	4.06 c	4.09 c	3.20 b	3.28 a	3.49 a
10/13/2020	**ECe**	2.09 bc	2.45 c	1.60 b	0.97 a	1.66 a	1.89 a
	SAR	4.76 b	4.74 b	4.61 b	2.93 a	4.40 a	4.12 a
4/23/2021	**ECe**	1.83 b	1.32 a	1.34 a	1.33 a	1.36 a	1.56 a
	SAR	3.94 ab	4.8 c	4.35 bc	3.61 a	3.76 a	4.59 b
10/11/2021	**ECe**	2.7 ab	3.21 b	2.55 ab	1.97 a	2.39 a	2.79 a
	SAR	6.22 b	6.07 b	5.05 a	4.25 a	5.33 a	5.45 a

Means sharing a similar letter are not significantly different at significance level (α) = 0.05.

Table 5.5. Mean infiltration rate (cm/s) for the entire study period measured by SATURO before (Spring) and after (Fall) the irrigation season.

	Mean Infiltration (cm/s)	**Standard Error**
Bare soil	0.000394 a	2.29E-04
3 d/week	0.000521 a	7.37E-05
7 d/week	0.000687 a	7.64E-05

Means sharing a similar letter are not significantly different at significance level (α) = 0.05.

5.4. Discussion

5.4.1. Autonomous SMS based irrigation

Scheduling the controller to irrigate 7d/week with proper threshold settings can allow for complete irrigation automation. The lower irrigation application by the SMS based irrigation using on-demand irrigation is also reported in other studies(Cardenas-Lailhacar et al., 2008; Cardenas-Lailhacar and Dukes, 2012). However, we observed no considerable difference in the average irrigation applications between restricted and on-demand

irrigation treatments. Therefore, on-demand irrigation might be the optimum programming choice during dry months in southern California. Furthermore, allowing SMS to decide when to irrigate results in running the irrigation for short periods, thus efficiently maintaining soil moisture between the programmed thresholds. This is evident from Figures 5.2–5.4, where soil moisture dropped below the lower threshold limit several times for the restricted treatments (T1 – T6) during the experiment since the controller was programmed to only irrigate 3 days of the week. It is widely known that infrequent irrigation promotes deeper rooting and increased drought resistance in turfgrasses. However, this was not supported by recent studies (Haghverdi et al., 2021b; Serena et al., 2020), which is also concurrent with our research findings.

The importance of understanding the smart controller settings for irrigation scheduling has been emphasized in literature (Davis and Dukes, 2016). The actual water applied by the irrigation controller in three years study was substantially lower than the reference ET_o across all treatments. The applied irrigation water steadily decreased each year as also shown in Table 5.2, although the same soil moisture thresholds were maintained throughout three-year experimental period, which is unclear, and further understanding of the functioning of the controller is required. However, development of thatch over turfgrass as the experiment progressed might be the reason for reduced evapotranspirative demand and thus increasing the moisture holding capacity of the soil (Liang et al., 2017). Furthermore, relatively lower ET_o and higher precipitation in the 2021 irrigation season could also be attributed to the reduced irrigation applications.

5.4.2. Response of bermudagrass to deficit irrigation with recycled water

A minimum NDVI value of 0.5 for bermudagrass was considered acceptable to maintain the aesthetic quality in central California (Haghverdi et al., 2021a). Furthermore, 75%ET_o was suggested as the minimum irrigation application for inland southern California, and severe deficit irrigation levels are only feasible for shorter periods before the hybrid bermudagrass quality falls below the minimum acceptable quality (Haghverdi et al., 2021b). This was also supported in this study as the treatments irrigated with more than 75% of ET_o resulted in an acceptable quality of turf (NDVI $\geq$0.5) for a longer duration during the irrigation season (Figure 5.5, Table 5.3). Therefore, further investigations are required to reevaluate the commonly accepted 0.6 crop coefficient for the bermudagrass(Meyer and Gibeault, 1986). Sevostianova et al. (2011) noted the acceptable quality of bermudagrass under saline irrigation (110%ET_o) in New Mexico. On the contrary, we observed a reduction in the turf quality under deficit recycled water irrigation despite the seasonal leaching by rainfall, suggesting that bermudagrass does not perform well under autonomous deficit irrigation with recycled water on a long-term basis a semi-arid climate. Future studies are needed to explore whether the continued deficit irrigation with recycled water had any effect on the spring green-up of hybrid bermudagrass.

On-demand irrigation resulted in a 3.4% reduction in turf temperatures over the study period while maintaining relatively better turf quality compared to restricted irrigation. We attribute this to a more pronounced evaporative cooling associated with on-demand irrigation while minimizing runoff and deep percolation. In addition, the irrigated urban landscape can reduce daytime temperatures through evapotranspiration, thus moderating

the climate of urban areas (Shashua-Bar et al., 2009). A previous study in Los Angeles reported that irrigation-induced increases in latent heat flux could lead to land surface temperature reductions in urban parks (Vahmani and Hogue, 2014). Furthermore, Bonfils and Lobell (2007) demonstrated the significant cooling effect of irrigation expansion on the summertime average daily daytime temperatures in California. Therefore, more work is required to elucidate better the tradeoffs between water conservation and the vital role of the irrigated urban landscape to mitigate the urban heat island effect under deficit irrigation.

5.4.3. Soil salinity, SAR, and infiltration rate under deficit recycled water irrigation

Deficit irrigation with recycled water resulted in the increase of soil salinity which oscillated with the rainy season, as precipitation resulted in leaching of salts. Samples collected after the irrigation season were associated with the highest salinity measurements, whereas samples collected before the irrigation season had the lowest salinity measurements, which concurs with the findings studies utilizing recycled water or deficit irrigation strategies (Aragüés et al., 2014; Lockett et al., 2008). On average, there was a 51% increase in soil salinity in fall 2021 compared to fall 2019. Lockett et al. (2008) observed the greatest fluctuation in soil salinity at the shallowest depth of 15 cm in golf courses irrigated with recycled water, which agrees with the results observed in our study with most seasonal fluctuation in soil salinity observed at the 0-30 cm of depth (Figure 5.7). Domínguez et al. (2011) indicated based on the MOPECO- Salt model results that deficit irrigation strategies without LF are remediable if the off-season rainfall is sufficient to leach out the salts from the soil root zone, which also agrees with the results observed

from our study as lower irrigation applications resulted in higher salt accumulation. The lower irrigation water being applied by the controller as the experiment progressed left less remaining water available to leach salts from the root zone.

Recycled water irrigation can increase the SAR of the soil, which can cause infiltration issues. High rainfall enhances the leaching of Na and K more than that of Ca and Mg since these are more soluble cations, which can explain the high SAR at deeper depths after the rainy season. There was an increase in SAR in our study, although no significant infiltration issues were observed. This agrees with the results from a study in 7 parks of Beijing, China, indicating an increase in soil salinity and SAR with recycled water irrigation compared to tap water (Chen et al., 2015). Aragüés et al. (2014) also highlighted the high transient salinity and sodicity risk under the combined effects of recycled and deficit drip irrigation in Mandarin trees. They concluded that soil water deficits should be avoided whenever saline reclaimed water is used for irrigation. Irrigation waters previously thought unsuitable for irrigation can often be used successfully without hazardous long-term consequences to crops or soils if proper management strategies are employed (Assouline et al., 2015; Rhoades et al., 1989). Although seasonal rainfall reduced the soil salinity (Figure 5.7) and no effect on infiltration was observed (Table 5.5) in our study, a steady increase in salinity and SAR over these cyclic occurrences need to be further investigated. Adding a leaching fraction to the required irrigation to manage salinity accumulation (Ayers and Westcot, 1985) might be necessary if seasonal rainfall is insufficient to leach down the salts.

5.5. Conclusion

Implementation of smart irrigation technologies offers to improve water use efficiency by maintaining the soil water status at the active root zone within a predefined desired range. This study focused on the implementation of soil moisture sensor-based deficit irrigation scheduling using recycled water. We observed that an irrigation application with soil moisture thresholds maintained between field capacity and 75% of field capacity results in the optimum visual quality of turfgrass with a minimum NDVI of 0.5 for more extended periods. This corresponded to at least 75%ET_o of irrigation application, which is significantly higher than recommended crop coefficient of 0.6 for bermudagrass. On-demand irrigation, i.e., programming the controller to irrigate 7 days/week, resulted in better visual quality and lower turf temperatures. Soil salinity oscillated with the seasonal cycles due to the rainfall in the non-irrigation season, although there was a steady increase in salinity and SAR as the experiment progressed. This was attributed to the combined effect of deficit irrigation and high evapotranspiration demand during the summer. No significant impact on the infiltration rate of the soil was observed. However, our results suggest that if irrigation management cannot minimize the salt accumulation, turfgrass's visual quality can be impacted over extended periods under deficit irrigation with recycled water. Furthermore, the use of smart irrigation controllers receiving feedback from the soil moisture sensors needs to be investigated for deficit irrigation scheduling.

5.6. References

ANSI/ASABE, 2016. Landscape Irrigation System Uniformity and Application Rate Testing 1–11.

Aragüés, R., Medina, E.T., Martínez-Cob, A., Faci, J., 2014. Effects of deficit irrigation strategies on soil salinization and sodification in a semiarid drip-irrigated peach orchard. Agric. Water Manag. 142, 1–9. https://doi.org/10.1016/j.agwat.2014.04.004

Assouline, S., Russo, D., Silber, A., Or, D., 2015. Balancing water scarcity and quality for sustainable irrigated agriculture. Water Resour. Res. 51, 3419–3436. https://doi.org/10.1002/2015WR017071

Ayers, R.S., Westcot, D.W., 1985. Water Quality for Agriculture. FAO UNITED NATIONS,Rome,italy 97.

Balling, R.C., Gober, P., Jones, N., 2008. Sensitivity of residential water consumption to variations in climate: An intraurban analysis of Phoenix, Arizona. Water Resour. Res. 44. https://doi.org/10.1029/2007WR006722

Blonquist, J.M., Jones, S.B., Robinson, D.A., 2006. Precise irrigation scheduling for turfgrass using a subsurface electromagnetic soil moisture sensor. Agric. Water Manag. 84, 153–165. https://doi.org/10.1016/j.agwat.2006.01.014

Bonfils, C., Lobell, D., 2007. Empirical evidence for a recent slowdown in irrigation-induced cooling. Proc. Natl. Acad. Sci. 104, 13582–13587. https://doi.org/10.1073/PNAS.0700144104

Bremer, D.J., Keeley, S.J., Jager, A., Fry, J.D., 2013. Lawn-watering perceptions and behaviors of residential homeowners in three Kansas (USA) cities: Implications for water quantity and quality. Int. Turfgrass Soc. Res. J. 12, 23–29.

Cardenas-Lailhacar, B., Dukes, M.D., 2014. Effect of temperature and salinity on the precision and accuracy of landscape irrigation soil moisture sensor systems. J. Irrig. Drain. Eng. 141, 4014076. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000847.

Cardenas-Lailhacar, B., Dukes, M.D., 2012. Soil Moisture Sensor Landscape Irrigation Controllers: A Review of Multi- Study Results and Future Implications. Am. Soc. Agric. Biol. Eng. 55, 581–590.

Cardenas-Lailhacar, B., Dukes, M.D., Miller, G.L., 2008. Sensor-Based Automation of Irrigation on Bermudagrass, during Wet Weather Conditions. J. Irrig. Drain. Eng. 134, 120–128. https://doi.org/10.1061/(ASCE)0733-9437(2008)134:2(120)

Cardenas, B., Dukes, M.D., 2016. Soil moisture sensor irrigation controllers and reclaimed water; Part I: Field-plot study. Appl. Eng. Agric. 32, 217–224. https://doi.org/10.13031/aea.32.11196

Cardenas, B., Dukes, M.D., Breder, E., Torbert, J.W., 2021. Long-term performance of smart irrigation controllers on single-family homes with excess irrigation . AWWA Water Sci. 3, 1–14. https://doi.org/10.1002/aws2.1218

Chen, W., Lu, S., Pan, N., Wang, Y., Wu, L., 2015. Impact of reclaimed water irrigation on soil health in urban green areas. Chemosphere 119, 654–661. https://doi.org/10.1016/j.chemosphere.2014.07.035

Davis, S.L., Dukes, M.D., 2016. Importance of et controller program settings on water conservation potential. Appl. Eng. Agric. 32, 251–262. https://doi.org/10.13031/aea.32.11182

Dean, D.E., Devitt, D.A., Verchick, L.S., Morris, R.L., 1996. Turfgrass quality, growth, and water use influenced by salinity and water stress. Agron. J. 88, 844–849. https://doi.org/10.2134/agronj1996.00021962008800050026x

Domínguez, A., Tarjuelo, J.M., de Juan, J.A., López-Mata, E., Breidy, J., Karam, F., 2011. Deficit irrigation under water stress and salinity conditions: The MOPECO-Salt Model. Agric. Water Manag. 98, 1451–1461. https://doi.org/10.1016/J.AGWAT.2011.04.015

Gao, Y., Shao, G., Wu, S., Xiaojun, W., Lu, J., Cui, J., 2021. Changes in soil salinity under treated wastewater irrigation: A meta-analysis. Agric. Water Manag. 255, 106986. https://doi.org/10.1016/j.agwat.2021.106986

GCSAA, G.C.S.A. of A., 2017. Land use characteristics and environmental stewardship programs on US golf courses. Golf Course Environ. Profile 4, 1–28.

Gonçalves, R.A.B., Folegatti, M. V., Gloaguen, T. V., Libardi, P.L., Montes, C.R., Lucas, Y., Dias, C.T.S., Melfi, A.J., 2007. Hydraulic conductivity of a soil irrigated with treated sewage effluent. Geoderma 139, 241–248. https://doi.org/10.1016/j.geoderma.2007.01.021

Grabow, G.L., Ghali, I.E., Huffman, R.L., Miller, G.L., Bowman, D., Vasanth, A., 2013. Water application efficiency and adequacy of ET-based and soil moisture-based irrigation controllers for turfgrass irrigation. J. Irrig. Drain. Eng. 139, 113–123. https://doi.org/10.1061/(ASCE)IR.1943-4774.0000528

Haghverdi, A., Reiter, M., Singh, A., Sapkota, A., 2021a. Hybrid Bermudagrass and Tall Fescue Turfgrass Irrigation in Central California: II. Assessment of NDVI, CWSI, and Canopy Temperature Dynamics. Agronomy 11, 1733. https://doi.org/10.3390/agronomy11091733

Haghverdi, A., Singh, A., Sapkota, A., Reiter, M., Ghodsi, S., 2021b. Developing irrigation water conservation strategies for hybrid bermudagrass using an evapotranspiration-based smart irrigation controller in inland southern California. Agric. Water Manag. 245, 106586. https://doi.org/10.1016/j.agwat.2020.106586

Khachatryan, H., Suh, D.H., Xu, W., Useche, P., Dukes, M.D., 2019. Towards sustainable water management: Preferences and willingness to pay for smart landscape irrigation technologies. Land use policy 85, 33–41. https://doi.org/10.1016/j.landusepol.2019.03.014

Liang, X., Su, D., Wang, Z., Qiao, X., 2017. Effects of Turfgrass Thatch on Water

Infiltration, Surface Runoff, and Evaporation. J. Water Resour. Prot. 09, 799–810. https://doi.org/10.4236/jwarp.2017.97053

Lockett, A.M., Devitt, D.A., Morris, R.L., 2008. Impact of reuse water on golf course soil and turfgrass parameters monitored over a 4.5-year period. HortScience 43, 2210–2218. https://doi.org/10.21273/hortsci.43.7.2210

Marcum, K.B., 2006. Use of saline and non-potable water in the turfgrass industry: Constraints and developments. Agric. Water Manag. 80, 132–146. https://doi.org/10.1016/J.AGWAT.2005.07.009

McCready, M.S., Dukes, M.D., Miller, G.L., 2009. Water conservation potential of smart irrigation controllers on St. Augustinegrass. Agric. Water Manag. 96, 1623–1632. https://doi.org/10.1016/j.agwat.2009.06.007

Meyer, J.L., Gibeault, V.A., 1986. Turfgrass performance under reduced irrigation. Calif. Agric. 19–20.

Nimmo, J.R., Schmidt, K.M., Perkins, K.S., Stock, J.D., 2009. Rapid measurement of field-saturated hydraulic conductivity for areal characterization. Vadose Zo. J. 8, 142–149. https://doi.org/10.2136/vzj2007.0159

Qian, Y.L., Mecham, B., 2005. Long-term effects of recycled wastewater irrigation on soil chemical properties on golf course fairways. Agron. J. 97, 717–721. https://doi.org/10.2134/agronj2004.0140

Qualls, R.J., Scott, J.M., DeOreo, W.B., 2001. Soil moisture sensors for urban landscape irrigation: Effectiveness and reliability. J. Am. Water Resour. Assoc. 37, 547–559. https://doi.org/10.1111/j.1752-1688.2001.tb05492.x

R Core Team, 2021. R: A Language and Environment for Statistical Computing.

Reynolds, W.D., Elrick, D.E., 1990. Ponded infiltration from a single ring: I. Analysis of steady flow. Soil Sci. Soc. Am. J. 54, 1233–1241. https://doi.org/10.2136/sssaj1990.03615995005400050006x

Rhoades, J.D., 1982. Soluble salts.--p. 167-179. En: Methods of soil analysis: part 2; chemical and microbiological properties.--Winsconsin, US: American Society of Agronomy, 1986.

Rhoades, J.D., Bingham, F.T., Letey, J., Hoffman, G.J., Dedrick, A.R., Pinter, P.J., Replogle, J.A., 1989. Use of saline drainage water for irrigation: Imperial Valley study. Agric. Water Manag. 16, 25–36. https://doi.org/10.1016/0378-3774(89)90038-3

Serena, M., Velasco-Cruz, C., Friell, J., Schiavon, M., Sevostianova, E., Beck, L., Sallenave, R., Leinauer, B., 2020. Irrigation scheduling technologies reduce water use and maintain turfgrass quality. Agron. J. 112, 3456–3469. https://doi.org/10.1002/agj2.20246

Sevostianova, E., Leinauer, B., Sallenave, R., Karcher, D., Maier, B., 2011. Soil salinity

and quality of sprinkler and drip irrigated warm-season turfgrasses. Agron. J. 103, 1773–1784. https://doi.org/10.2134/agronj2011.0163

Shashua-Bar, L., Pearlmutter, D., Erell, E., 2009. The cooling efficiency of urban landscape strategies in a hot dry climate. Landsc. Urban Plan. 92, 179–186. https://doi.org/10.1016/J.LANDURBPLAN.2009.04.005

SWRCB, California State Water Resources Control Board. [WWW Document], 2015. . Munic. wastewater Recycl. Surv. URL https://www.waterboards.ca.gov/water_issues/programs/grants_loans/water_recycling/munirec.shtml (accessed 10.31.18).

Vahmani, P., Hogue, T.S., 2014. Incorporating an Urban Irrigation Module into the Noah Land Surface Model Coupled with an Urban Canopy Model. J. Hydrometeorol. 15, 1440–1456. https://doi.org/10.1175/JHM-D-13-0121.1

Zalacáin, D., Bienes, R., Sastre-Merlín, A., Martínez-Pérez, S., García-Díaz, A., 2019. Influence of reclaimed water irrigation in soil physical properties of urban parks: A case study in Madrid (Spain). Catena 180, 333–340. https://doi.org/10.1016/j.catena.2019.05.012

Chapter 6. Summary and Conclusion

The overall goal of this doctoral dissertation was to development of models and techniques for advancement of efficient urban landscape irrigation scheduling. Accurate estimation of the soil hydraulic properties and reference evapotranspiration (ET_o) is critical for calculating soil water balance and are widely used for irrigation scheduling by smart irrigation controllers. Direct measurement of these properties is costly and time-consuming, thus accurate models are required for better estimation using readily available data. Furthermore, use of smart controllers to implement autonomous deficit irrigation using recycled water needs to be explored in semi-arid conditions of California.

In chapter 2 and 3, we developed and evaluated PC_{NN}-PTFs to estimate the SWRC and SHCC measured using the evaporation experiments, mainly via the HYPROP system. The PC_{NN}-PTF approach showed promising performance for continuous soil water retention hydraulic conductivity estimation over a wide range of soil tensions. The HYPROP system offers the advantage of producing high-resolution soil hydraulic conductivity data over a wide range of soil tensions (pF = 1.5 to 3.5), which is critical for developing robust PC_{NN}-PTF models since this approach learns the shape of the SWRC or SHCC directly from measured data. We recommend the PC_{NN}-PTF approach to derive the next generation of water retention and hydraulic conductivity models using high-resolution data measured via the HYPROP system.

In Chapter 4, we evaluated the performance of 8 empirical temperature-based and 4 Artificial neural network models (ANN) ET_o models at 101 active California Irrigation Management Information System (CIMIS) weather stations in California using more than

725,000 observations from 1985 to 2019. The ANN models outperformed the empirical equations and showed high generalization ability. Our results suggest that using raw input data for neural network models is better than using the reconstructed signal obtained from wavelet transform. The Hargreaves, and Hargreaves & Samani were the best performing empirical models across the seasons, years, and climate divisions. We recommend the application of the calibrated Hargreaves & Samani and the ANN model we developed in this study for accurate estimations of ET_o in data-scarce conditions in urban settings across California climate divisions. Both these models only require on-site air temperature measurements, which are typically collected by smart weather-based landscape irrigation controllers. Therefore, these models can serve as an assessment guide for researchers in the realm of automatic weather-based smart irrigation scheduling, thus advancing better water management in urban landscapes.

In Chapter 5, we focused on the implementation of soil moisture sensor-based deficit irrigation scheduling using recycled water. We observed that an irrigation application with soil moisture thresholds maintained between field capacity and 75% of field capacity results in the optimum visual quality of turfgrass with a minimum NDVI of 0.5 for more extended periods. On-demand irrigation, i.e., programming the controller to irrigate 7 days/week, resulted in better visual quality and lower turf temperatures. Soil salinity oscillated with the seasonal cycles due to the rainfall in the non-irrigation season, although there was a steady increase in salinity and SAR as the experiment progressed. This was attributed to the combined effect of deficit irrigation and high evapotranspiration demand during the summer. Our results suggest that if irrigation management cannot minimize the

salt accumulation, turfgrass's visual quality can be impacted over extended periods under deficit irrigation with recycled water. Furthermore, the use of smart irrigation controllers receiving feedback from the soil moisture sensors needs to be investigated for deficit irrigation scheduling.