

Type of the Paper (Article)

Evaluation and Development of Pedotransfer Functions for Predicting Saturated Hydraulic Conductivity for Mexican Soils

Josué Trejo-Alonso ¹, Antonio Quevedo ², Carlos Fuentes ² and Carlos Chávez ^{1,*}

¹ Water Research Center, Centro de Investigaciones del Agua-Queretaro (CIAQ), Autonomous University of Queretaro, Cerro de las Campanas SN, Col. Las Campanas, 76010, Queretaro, Mexico; josue.trejo@uaq.mx, chagcarlos@uaq.mx

² Mexican Institute of Water Technology, Jiutepec, Morelos, Mexico; jose_quevedo@tlaloc.imta.mx, cbfuentesr@gmail.com

* Correspondence: chagcarlos@uaq.mx; Tel.: +52-442-192-1200 (Ext. 6036)

Abstract: In the present work, we evaluate the prediction capability of six Pedotransfer functions (PTFs), reported in the literature, for the saturated hydraulic conductivity estimations (K_s). We used a database with 900 measured samples obtained from the Irrigation District 023, in San Juan del Rio, Queretaro, Mexico. Additionally, six new PTFs were constructed for K_s from clay percentage, bulk density and saturation water content data. The results show, for the evaluated models, that one model presents an overestimation for $K_s > 0.5 \text{ cm h}^{-1}$ values, three models have an underestimation for $K_s > 1.0 \text{ cm h}^{-1}$ and two models have a good correlation ($R^2 > 0.98$) but are necessary more than three parameters. Nevertheless, the last two models require from three to four parameters in order to get the optimization. By other hand, the models proposed in this work have a similar correlation with a less number of parameters: the fit is seen to be much better than using the existing ones, achieving a correlation of $R^2 = 0.9822$ with only one variable and a $R^2 = 0.9901$ when we use two.

Keywords: pedotransfer functions; inverse methods; gravity irrigation; model validation; experimental data

1. Introduction

The calculation of saturated hydraulic conductivity is a very crucial factor to optimize the flow rate applied to the border or furrow in the gravity irrigation [1-5]. Although this property is easily measurable in a laboratory, or in the field, for its application in small scales, most of the time it is required to be used on a large scale [4,5]. This inconvenient brings us numerous tests for soil measurements, which is time consuming, costly and impractical [6-7].

In recently years, there are a lot of studies about Pedotransfer Functions (PTFs) [6-9]. These mathematical models allow to estimate the saturated hydraulic conductivity (K_s) from some soil characteristics as: texture, field capacity, the permanent wilting point, bulk density, porosity, organic matter, among other [6-9]. The robustness of the model is linked to the number of physical parameters used to calculate the saturated hydraulic conductivity; the more parameters, the more accurate the prediction. But, as it was mentioned before, the amount of measurements make the PTFs difficult to get, because the economical resources and time, which it is a limitation in this kind of functions.

Diverse pedotransfer models can be found in the existing literature [6-11], several of them validated with already known databases (UNSODA, ROSETTA, among others). However, the predictive

capacity they have has been questioned, because the soils in which they want to apply are different from the soils used for their development [12]. Those results indicate that, for its use at the local level, it is necessary to take in-situ samples to validate its application [13], or where necessary, make the appropriate corrections or adjustments [7]. Thus, the objectives of the present work are: 1) Use some PTFs from the literature to obtain the hydraulic conductivity of the sampled soils with a great diversity of textural classes, 2) Develop new PTFs based on the principal component analysis technique to reproduce the observed data and validate them.

2. Materials and Methods

2.1. Pedotransfer functions in literature

For the estimation of K_s , some of the models shown in [9] were tested, however, due to in the samples analyzed in the laboratory in this work the content of organic matter was not quantified, some models were used using the measured variables (Table 1). Statistical analysis of predictive capacity was performed with the R stats package [14] using non-linear least squares estimation [15].

Table 1. Basic information about Pedotransfer functions evaluated in this study.

PTF	Formula	Reference
PTF-1	$\ln\left(\frac{K_s}{0.2778}\right) = \begin{cases} a + bSa + cCl + dSa^2 + eCl^2 + f\theta_s^2 + gSa\theta_s + hSa^2\theta_s^2 \\ + iCl^2\theta_s^2 + jSa^2Cl + kCl^2\theta_s + lSa^2\theta_s + mClSa \end{cases}$	[4]
PTF-2	$\log_{10}\left(\frac{K_s}{0.070556}\right) = a + bSa + cCl$	[12]
PTF-3	$\ln\left(\frac{K_s}{0.2778}\right) = a + bSa + \frac{c + dSa + eCl + fCl^2}{\theta_s}$	[13]
PTF-4	$\ln\left(\frac{K_s}{0.2957}\right) = a + b \ln Cl + c \ln Sa + d\rho_a$	[6]
PTF-5	$K_s = a(e^{bSa}) + 0.3$	[7]
PTF-6	$K_s = a\phi^m$	[14]

Abbreviations are as follows. Sa: Sand (%); Cl: Clay (%); θ_s : saturation water content ($\text{cm}^3\text{cm}^{-3}$); ρ_a : bulk density (g cm^{-3}) and the coefficients from a to m are obtained by fitting the model to the experimental data.

2.2. Soil data base

The database used in this study was developed from samplings in 900 plots in the Irrigation District 023 San Juan del Río Querétaro (Figure 1). These samples were sent to the laboratory to obtain the following parameters: soil texture by the Bouyucos hydrometer, bulk density by the cylinder method of known volume, moisture content at saturation, field capacity and permanent wilting point by the method of the pressure membrane pot and the saturated hydraulic conductivity by the variable head permeameter method. The measurement of the variables and the hydrodynamic characterization of soils are widely discussed in [19].

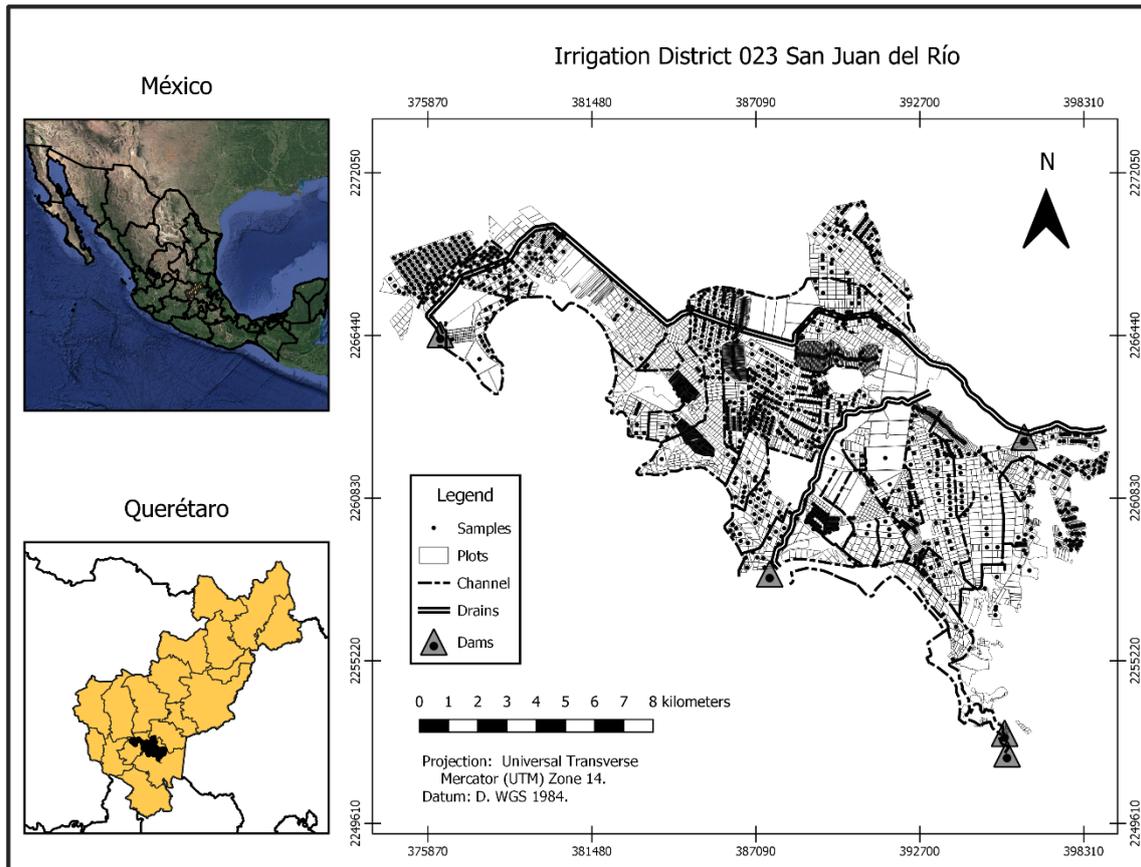


Figure 1. Location map of the sampling points in the Irrigation District 023 San Juan del Río Querétaro.

2.3. Statistical analysis

The accuracy of PTFs in predicting hydraulic conductivity was evaluated by calculating three statistical measures between the predictions and observations; root mean square error (RMSE) between the measured and predicted hydraulic conductivity, the modeling efficiency (EF) and the mean error (ME):

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (K_{smi} - K_{spi})^2}{N}} \quad (1)$$

$$EF = 1 - \frac{\sum_{i=1}^N (K_{smi} - K_{spi})^2}{\sum_{i=1}^N (K_{smi} - \overline{K_{smi}})^2} \quad (2)$$

$$ME = \frac{\sum_{i=1}^N (K_{smi} - K_{spi})}{N} \quad (3)$$

where K_{sm} is the measurement hydraulic conductivity (cm h^{-1}), K_{sp} the predicted hydraulic conductivity (cm h^{-1}), $\overline{K_{sm}}$ the mean of the measured values, and N the total number of observations.

3. Results and Discussions

3.1. Laboratory results

Table 2 shows a summary of the texture results obtained in the laboratory. It can be seen that the samples collected cover 11 of the 12 textural classes and, for this set, the predominant soils in the study area are those of Loamy silty clay and Silty clay texture. Furthermore, Table 3 presents a summary of the statistical properties of each of the variables used in this study.

Table 2. Classification of soil samples by textural class.

Texture	No. of samples	(%)
Clay	22	2.4
Silty clay	17	1.9
Sandy clay	10	1.1
Silt	136	15.1
Clay loam	64	7.1
Silty clay loam	163	18.1
Clay sandy loam	139	15.5
Silty loam	145	16.1
Loam	94	10.5
Sandy loam	110	12.2

Table 3. Statistical properties of data measured in the laboratory.

Variable	Min	Max	Median	Mean	SD	Q1	Q3
Sand (%)	0.07	77.83	28.35	31.14	20.22	13.75	52.00
Clay (%)	2.12	59.46	21.74	21.95	12.06	13.44	30.00
Silt (%)	0.80	92.00	45.27	46.91	23.48	27.30	59.79
Bulk density (ρ_a) (g cm^{-3})	1.18	1.70	1.40	1.41	0.11	1.32	1.47
Volumetric water content (θ_s) ($\text{cm}^3 \text{cm}^{-3}$)	0.35	0.56	0.47	0.47	0.04	0.45	0.50
Field Capacity (FC) ($\text{cm}^3 \text{cm}^{-3}$)	0.17	0.47	0.29	0.30	0.06	0.25	0.32
Permanent wilting point (PWP) ($\text{cm}^3 \text{cm}^{-3}$)	0.07	0.35	0.13	0.15	0.05	0.10	0.17
Saturated hydraulic conductivity (K_s) (cm h^{-1})	0.05	5.15	0.78	1.42	1.42	0.40	1.80

3.2. Saturated hydraulic conductivity obtained with existing models

Comparison of laboratory measured data on saturated hydraulic conductivity (K_s) with those estimated (K_{SE}) with models PTF-1 - PTF-6 are shown in Figure 2. The dotted diagonal line corresponds to the 1:1 ratio, while the horizontal indicates a value of zero in the estimation of the residuals. Pearson's correlation coefficient and standard deviations of the models used are shown in Table 4.

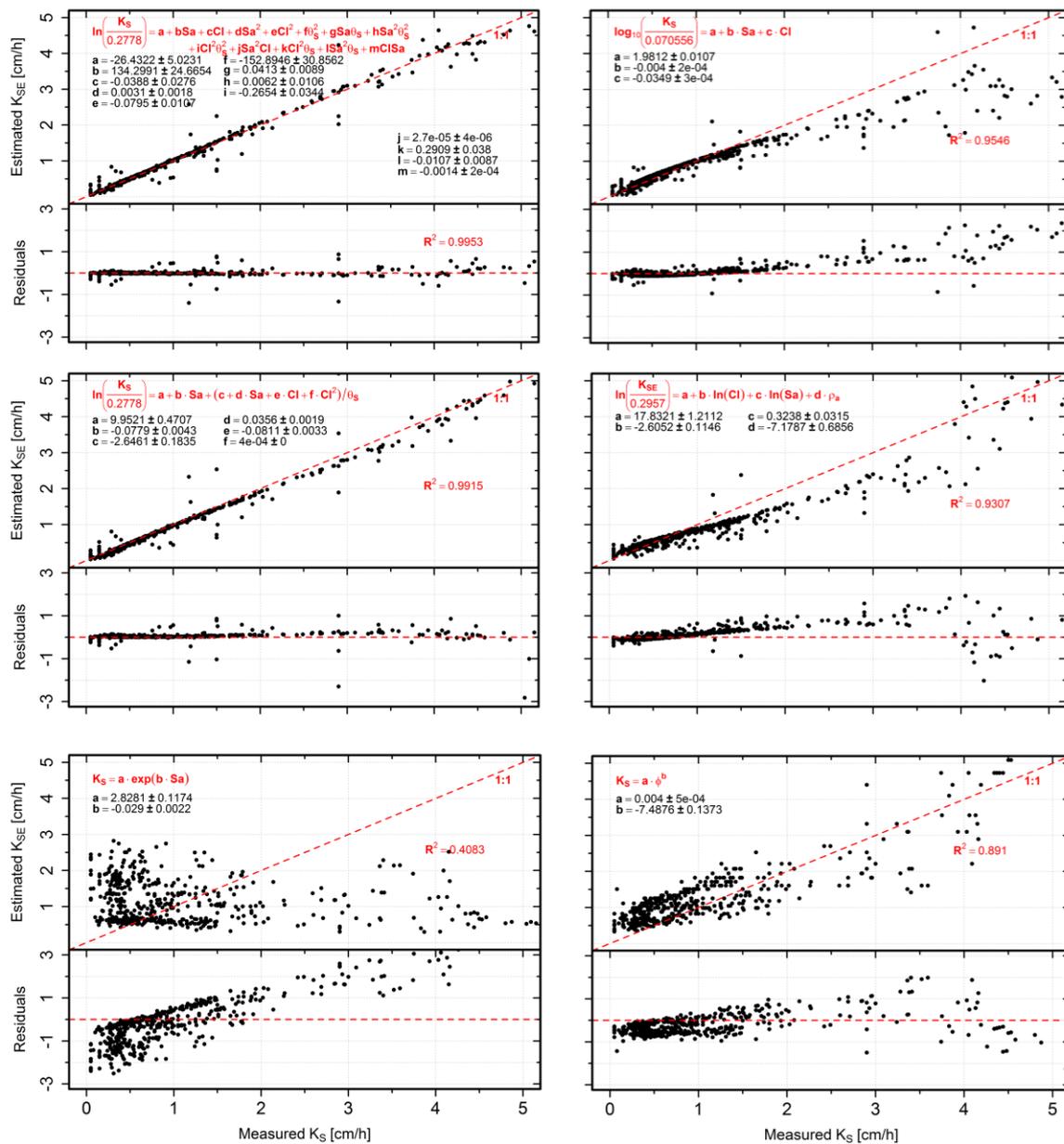


Figure 2. Comparison of the estimated saturated hydraulic conductivity (K_{SE}) with the PTF1-PTF6 and those measured in the laboratory (K_S).

Table 4. Statistical analysis for tested models.

PTFs	RMSE	EF	ME	σ	R^2
PTF-1	0.1370	0.9907	0.0574	0.1370	0.9953
PTF-2	0.4325	0.9104	0.2449	0.4325	0.9546
PTF-3	0.1895	0.9823	0.0823	0.1895	0.9915
PTF-4	0.7143	0.8434	0.3837	0.7143	0.9307
PTF-5	1.3018	-2.2687	1.0160	1.3018	0.4083
PTF-6	0.6498	0.6874	0.5125	0.6498	0.8910

According to the residual analysis [20,23] any value greater than 2 corresponds to anormal values in the distribution of errors, commonly called "outlier". In the Figure 2 (evaluated models) they are also shown the saturated hydraulic conductivity value at this characteristic appear in each of the models used. Thus, it can be seen that the first four models begin to show this characteristic in

the range of 3.6 to 4.25 cm h⁻¹ (PTF-1 - PTF-4), however, in the PTF-5 model it is presented from 2.00 cm h⁻¹, in addition to which is the model that has more variation in the estimation of K_s. This indicates that the variance of the errors is not constant with respect to the value of the saturated hydraulic conductivity, which is reflected in the range of application in the estimation of this parameter [22]. In addition, the models where they present a better prediction, there is a greater presence of outliers, that is, the existence of anormal values in the distribution of errors. However, this can be attributed to the low predictive capacity of the models used, land use, characteristics of the plot, formation factors, among others [22-24].

3.3. Principal component analysis

Despite the predictive capacity of some models, the number of variables necessary to be able to apply them is high. Also, when taking too many variables on a set of objects we will have to consider an excessive number of possible correlations. Another fact to consider is that there is self-correlation between the variables already measured, for example, we know that there is a close relationship between the bulk density and the clay percentage. This makes it possible to reduce the number of variables, in this way a principal component analysis was carried out [25] using the packages "factoextra" [26] and "FactoMineR" [27]. The results of the correlations between the variables that we have helped us to propose new pedotransference functions.

3.4. Development of new models

Given the low predictive capacity of saturated hydraulic conductivity of the models used in the literature to predict the saturated hydraulic conductivity of our database, we decided to develop new models: linear, exponential and potential using the parameters that are most correlated with K_s: Cl, θ_s and ρ_a . For optimization, it was decided to use the NLS method, which uses a relative displacement convergence criterion comparing the numerical imprecision in the estimates of the current parameters with the residual sum of squares, executed on the data, using the constraint to avoid failures in the method: $y = f(x, \theta) + \text{eps}$.

The optimization of the constants in the models was performed with 50% of the data obtained randomly in each textural class. After trying to fit several models, it was decided to choose six that are shown in Table 5, where a, b, c and d are constant.

Table 5. New pedotransfer models (NPTF) and statistical measurements of the entire database.

NPTF	MODEL	RMSE	EF	ME	σ	R ²
NF-1	$K_s = a + bCl + c\theta_s + d\rho_a$	0.6888	0.6865	0.6144	0.6845	0.8709
NF-2	$K_s = a \times \exp(bCl)$	0.2636	0.9623	0.1958	0.2636	0.9822
NF-3	$K_s = a \times \exp[b(Cl/\theta_s)]$	0.2717	0.9582	0.1871	0.2651	0.9813
NF-4	$K_s = a - b \times \exp[c(Cl/\theta_s)^d]$	0.1983	0.9788	0.1186	0.1985	0.9901
NF-5	$K_s = a \times \exp[(-Cl/\theta_s)/b]$	0.2738	0.9582	0.1877	0.2717	0.9813
NF-6	$K_s = a / \{1 + \exp[b - c(Cl/\theta_s)]\}$	0.2110	0.9783	0.1767	0.2473	0.9844

The predictive capacity of the models, once calibrated in the previous section, was evaluated with the remaining 50% of the samples. The results of the NPTF adjustments are shown in Figure 3. The dotted lines at the top of each graph correspond to the fit, while the horizontal dotted line indicates a value of zero in the estimate of the residuals. In addition, the values of the constants and the Pearson correlation coefficient are shown.

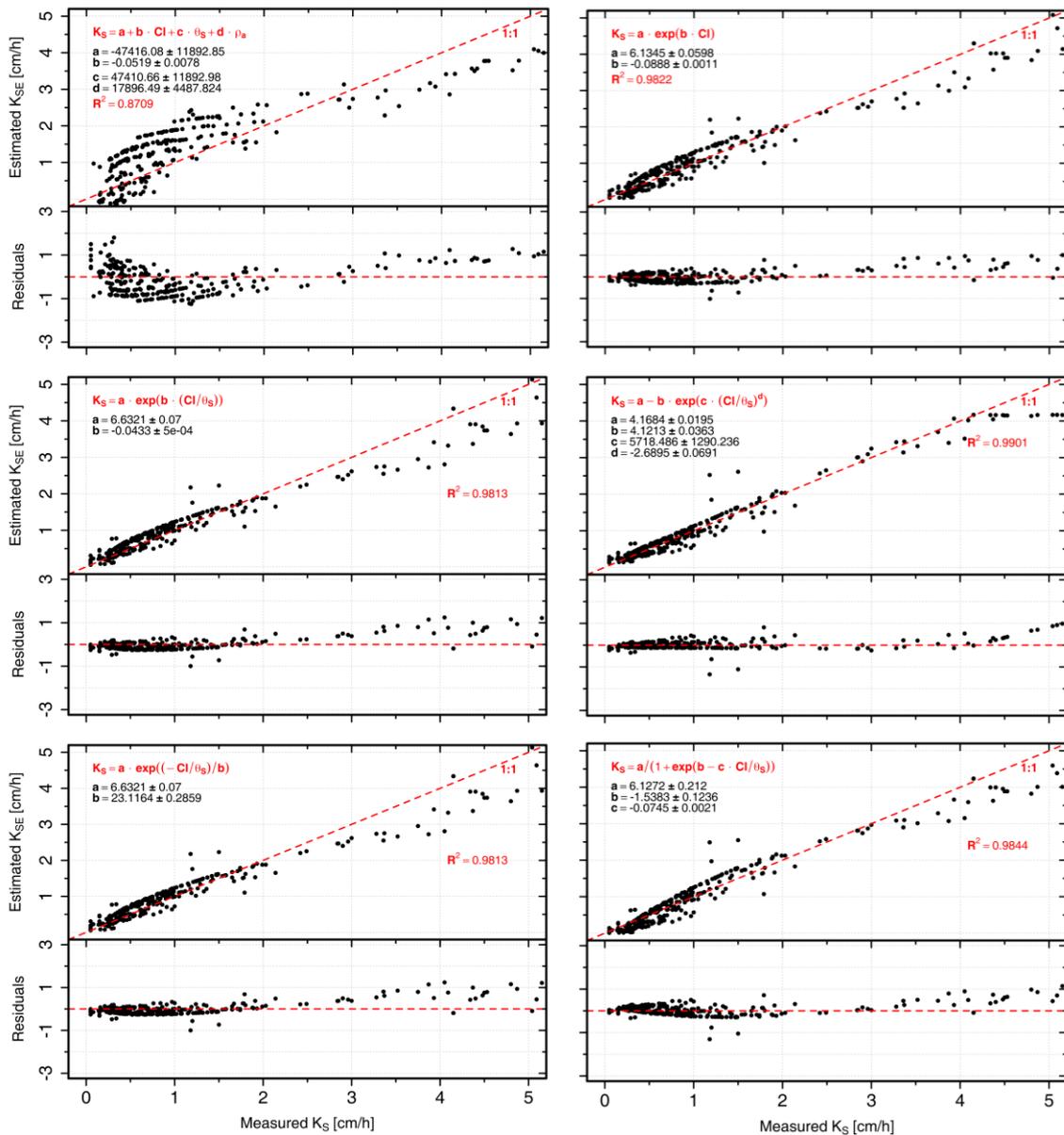


Figure 3. Comparison of the estimated saturated hydraulic conductivity (K_{SE}) with the NF-1 - NF6 and those measured in the laboratory (K_S).

The results obtained with the proposed models show that the NF-4 model better predicts the data obtained experimentally by showing less variance and higher Pearson correlation coefficient ($R^2 = 0.9901$), on the contrary, with the linear model (NF-1) when using 3 variables, it has more bias in the residuals than the exponential model ($R^2 = 0.8747$). However, the NF-2 model only uses the clay content as the input variable, which makes it a simple model with good predictability to obtain saturated hydraulic conductivity ($R^2 = 0.9822$). Furthermore, in the six proposed models there are no outliers. These results are particularly relevant for understanding regional soil-water dynamics for irrigation and drainage studies in plots or Irrigations Districts with similar soil properties.

5. Conclusions

This work reported laboratory measurements of 8 physical parameters for 900 plots in the Irrigation District 023 located in the state of Queretaro, Mexico. With those samples, six Pedotransfer models found in the literature were tested. As a result, it can be seen that when applying these models to the database we have, the prediction of saturated hydraulic conductivity is deficient for values greater than 2.00 cm h^{-1} .

The importance of clay content and moisture content at saturation as predictors of PTFs for saturated hydraulic conductivity has been analyzed in previous studies. The above characteristics, and the results obtained with existing and own models give us tools to conclude that there are no universal Pedotransfer models, so the use of any of them should be taken with caution, or failing that, develop regional equations that are able to obtain saturated hydraulic conductivity, given an error criterion, so that the results obtained are in accordance with the reality of the phenomenon and the place being studied.

Author Contributions: Conceptualization, C.C. and C.F.; methodology, C.C.; software, J.T.A. and C.C.; validation, J.T.A., C.F. and A.Q.; formal analysis, C.F.; investigation, J.T.A. and C.C.; resources, C.C.; data curation, C.C., A.Q. and C.F.; writing—original draft preparation, J.T.A. and C.C.; writing—review and editing, A.Q. and C.F.; visualization, C.F.; supervision, C.C.; project administration, C.C.; funding acquisition, C.C. and C.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported as part of a collaboration between the National Water Commission (CONAGUA, according to its Spanish acronym); the Irrigation District 023, San Juan del Río, Querétaro; and the Autonomous University of Querétaro, under the program RIGRAT 2015-2019.

Conflicts of Interest: The authors declare that they have no conflict of interests regarding the publication of this paper.

References

1. Chávez, C.; Fuentes, C. Design and evaluation of surface irrigation systems applying an analytical formula in the irrigation district 085, La Begoña, Mexico. *Agric. Water Manage.* **2019**, *221*, 279-285.
2. Di, W.; Xue, J.; Bo, X.; Meng, W.; Wu, Y.; Du, T. Simulation of irrigation uniformity and optimization of irrigation technical parameters based on the SIRMOD model under alternate furrow irrigation. *Irrig. Drain.* **2017**, *66*, 478-491.
3. Gillies, M.H.; Smith, R.J. SISCO: Surface irrigation simulation, calibration and optimization. *Irrig. Sci.* **2015**, *33*, 339-355.
4. Saucedo, H.; Zavala, M.; Fuentes, C. Complete hydrodynamic model for border irrigation. *Water Technol. Sci.* **2011**, *2* (2), 23-38.
5. Weibo, N.; Ma, X.; Fei, L. Evaluation of infiltration models and variability of soil infiltration properties at multiple scales. *Irrig. Drain.* **2017**, *66*, 589-599.
6. Zhang, Y.; Schaap, M.G. Estimation of saturated hydraulic conductivity with pedotransfer functions: A review. *J. Hydrol.* **2019**, *575*, 1011-1030.
7. Abdelbaki, A.M. Evaluation of pedotransfer functions for predicting soil bulk density for U.S. soils. *Ain Shams Eng. J.* **2018**, *9*, 1611-1619.
8. Brakensiek, D.; Rawls, W.J.; Stephenson, G.R. Modifying SCS hydrologic soil groups and curve numbers for rangeland soils. *ASAE* **1984**, 84-203.
9. Rasoulzadeh A. Estimating Hydraulic Conductivity Using Pedotransfer Functions, Hydraulic Conductivity-Issues; Determination and Applications. In: Elango L (Ed.), Croatia, 2011, pp. 145-164.
10. Vereecken, H.; Maes, J.; Feyen, J. Estimating unsaturated hydraulic conductivity from easily measured soil properties. *Soil Sci.* **1990**, *149*, 1-12.
11. Ferrer-Julà, M.; Estrela-Monreal, T.; Sánchez-del Corral-Jiménez, A.; García-Meléndez, E. Constructing a saturated hydraulic conductivity map of Spain using pedotransfer functions and spatial prediction. *Geoderma* **2004**, *123*, 275-277.
12. Moreira, L.; Righetto, A.M.; Medeiros V.M. Soil hydraulics properties estimation by using pedotransfer functions in a northeastern semiarid zone catchment, Brazil. International Environmental Modelling and Software Society, 2004, Osnabrueck. Complexity and Integrated Resources Management, Transactions of the 2nd Biennial Meeting of the International Environmental Modelling and Software Society, iEMSs 2004. Manno, Switzerland **2004**, *2*, 990-995.
13. Kaur, R.; Kumar, S.; Gurung, H. A pedo-transfer function (PTF) for estimating soil bulk density from basic soil data and its comparison with existing PTFs. *Aust J Soil Res.* **2002**, *40*, 847-857.
14. R Core Team (2019) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.

15. Bates, D.M.; Chambers, J.M. Statistical Models In S. In *Nonlinear Models*. Chambers, J.M., Hastie, T.J. Eds.; CRC Press, Boca Raton, Florida, USA, 1992; pp. 421–453
16. Cosby, B.; Hornberger, G.; Clapp, R.; Ginn, T. A statistical exploration of the relationship of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.* **1984**, *20*, 682-690.
17. Saxton, K.; Rawls, W.J.; Romberger, J.S.; Papendick, R.I. Estimating generalized soil water characteristics from texture. *Soil Sci. Soc. Am. J.* **1986**, *5*, 1301-1036.
18. Ahuja, L.R.; Naney, J.W.; Green, R.E.; Nielsen, D.R. Macroporosity to characterize spatial variability of hydraulic conductivity and effects of land management. *Soil Sci. Soc. Am. J.* **1984**, *48*, 699-702
19. Chávez, C.; Fuentes, C. Optimization of furrow irrigation by an analytical formula and its impact on reduction of the water applied. *Agrociencia* **2018**, *52*, 483-496.
20. Venables, W.N.; Ripley, B.D. *Modern Applied Statistics with S*. 4th ed.; New York: Springer, USA, 2002; 495 PP.
21. Montgomery, D.C.; Runger, G.C. *Applied Statistical and Probability for Engineers*, 6th Edition, John Wiley & Sons, Inc., USA, 2014; 811 pp.
22. Vereecken, H.; Schnepf, A.; Hopmans, J.W.; Javaux, M.; Or, D.; Roose, T.; Vanderborght, J.; Young, M.H.; Amelung, W.; Aitkenhead, M.; et al. Modeling Soil Processes: Review, Key Challenges, and New Perspectives. *Vadose Zone J.* **2016**, *15*, vzt2015.09.0131.
23. Hwang, H.T.; Jeon, S.W.; Suleiman, A.A.; Lee, K.K. Comparison of saturated hydraulic conductivity estimated by three different methods. *Water (Switzerland)* **2017**, *9*, 942.
24. Chapuis, R.P. Predicting the saturated hydraulic conductivity of soils: A review. *Bull. Eng. Geol. Environ.* **2012**, *71*, 401–434.
25. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ Psychol.* **1933**, *24(6)*, 417-441.
26. Kassambara, A.; Mundt, F. factextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.5. **2017**, Available from: <https://CRAN.R-project.org/package=factextra>
Le, S.; Josse, J.; Husson, F. 2008. FactoMineR: An R Package for Multivariate Analysis. *J. Stat. Soft.* **2008**, *25(1)*, 1-18.