Modeling in Soil Science:

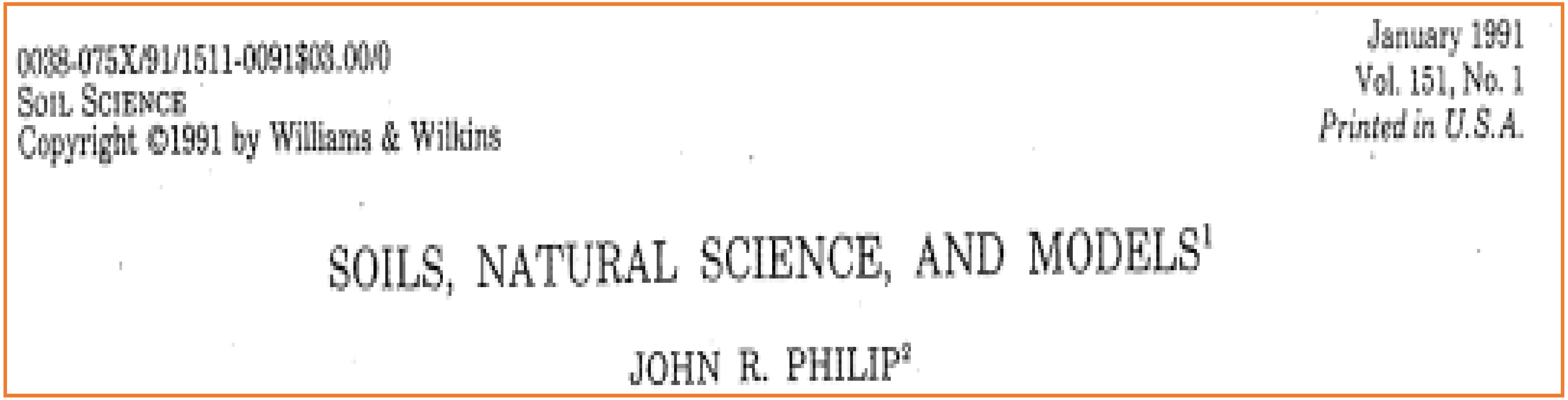
Opportunities and Threats

Habib Khodaverdiloo, Urmia University, Urmia, Iran

Email: h.khodaverdiloo@urmia.ac.ir

The Scientific Community and Scientific Standards

- **Universalism**
- **Communality**
- **Disinterestedness**
- **Organized Skepticism**



0038-075X/91/1511-0091$03.00/0
SOIL SCIENCE
Copyright ©1991 by Williams & Wilkins

January 1991
Vol. 151, No. 1
Printed in U.S.A.

SOILS, NATURAL SCIENCE, AND MODELS[1]

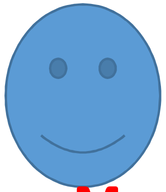
JOHN R. PHILIP[2]

Model and Modeling

- **What is a Model?**

Albert Einstein:
"Everything should be made as simple as possible, but not simpler."

Reality



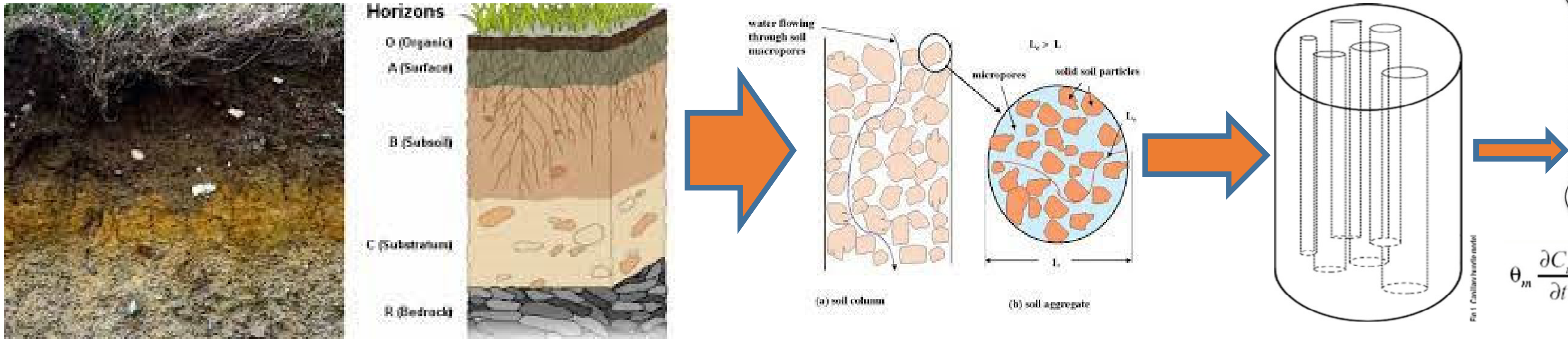
Model#1

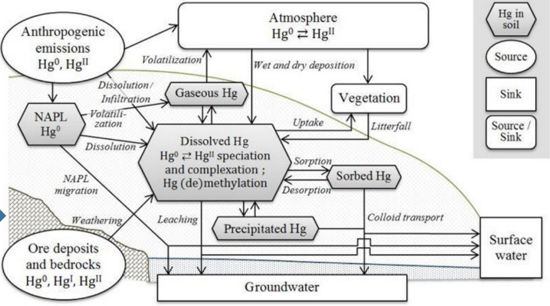
Model#2

- **What does modeling seek to do? To find truth!**

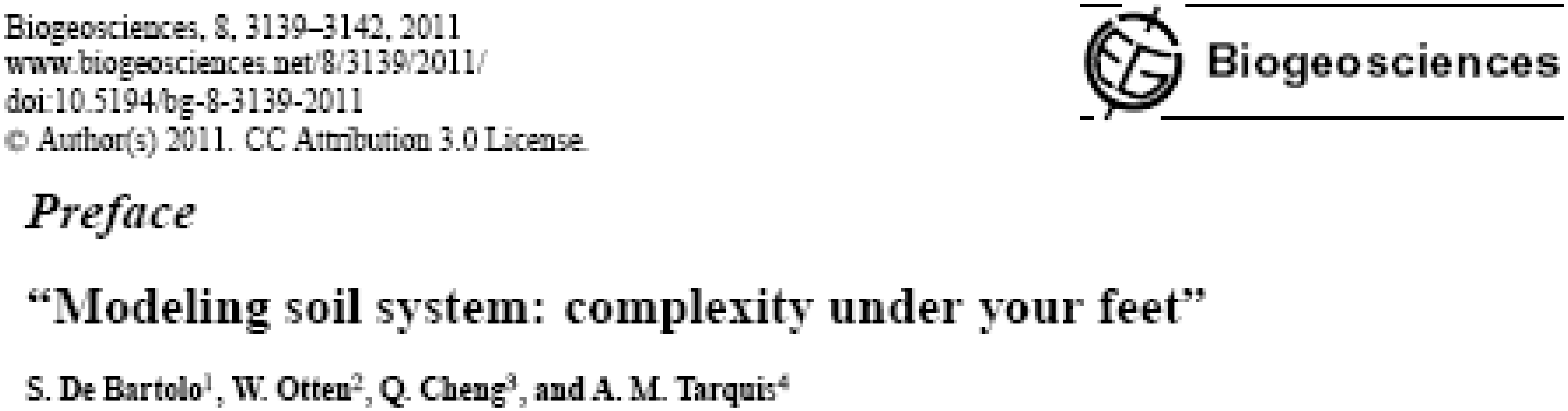
- **Models and the Real World**

George Box: "All models are wrong but some are useful".





$$\theta_m \frac{\partial C_m}{\partial t} + \theta_{im} \frac{\partial C_{im}}{\partial t} = \frac{\partial}{\partial x_i}\left(\theta_m D_{ij} \frac{\partial C_m}{\partial x_j}\right) - \frac{\partial}{\partial x_i}(\theta_m v_i C_m)$$
$$+ q_s C_s - q'_s C_m - \lambda_{1,m} \theta_m C_m - \lambda_{1,im} \theta_{im} C_{im}$$



Biogeosciences, 8, 3139–3142, 2011
www.biogeosciences.net/8/3139/2011/
doi:10.5194/bg-8-3139-2011

Biogeosciences

Preface

"Modeling soil system: complexity under your feet"

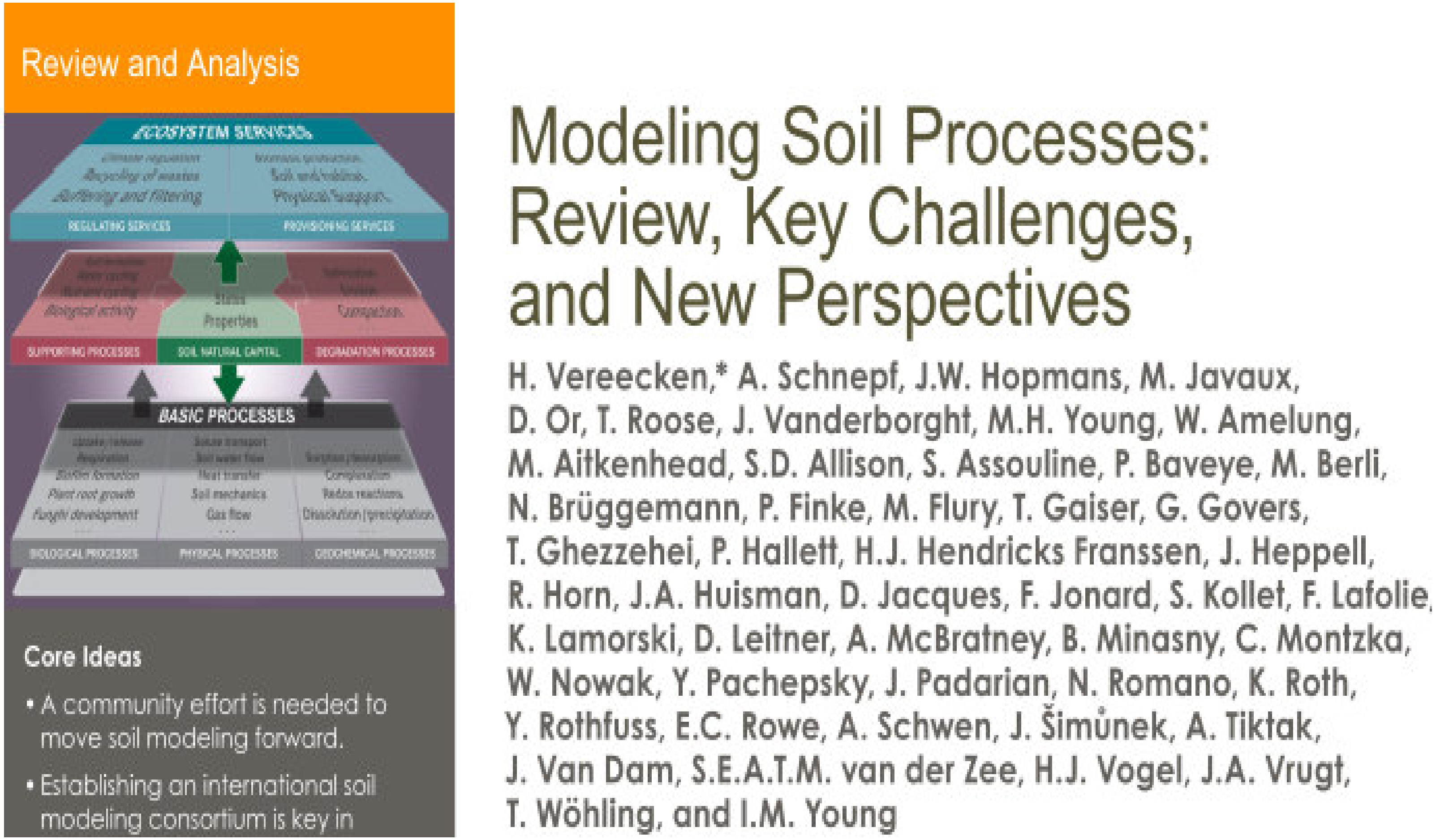
S. De Bartolo[1], W. Otten[2], Q. Cheng[3], and A. M. Tarquis[4]

- **Soil as a complex 3-phase SYSTEM**
- **Soil from natural to human-natural body**
- **Soil and Modeling**

Leonardo da Vinci:

"We know more about the celestial bodies than about the soil underfoot!"

Published May 13, 2016



Review and Analysis

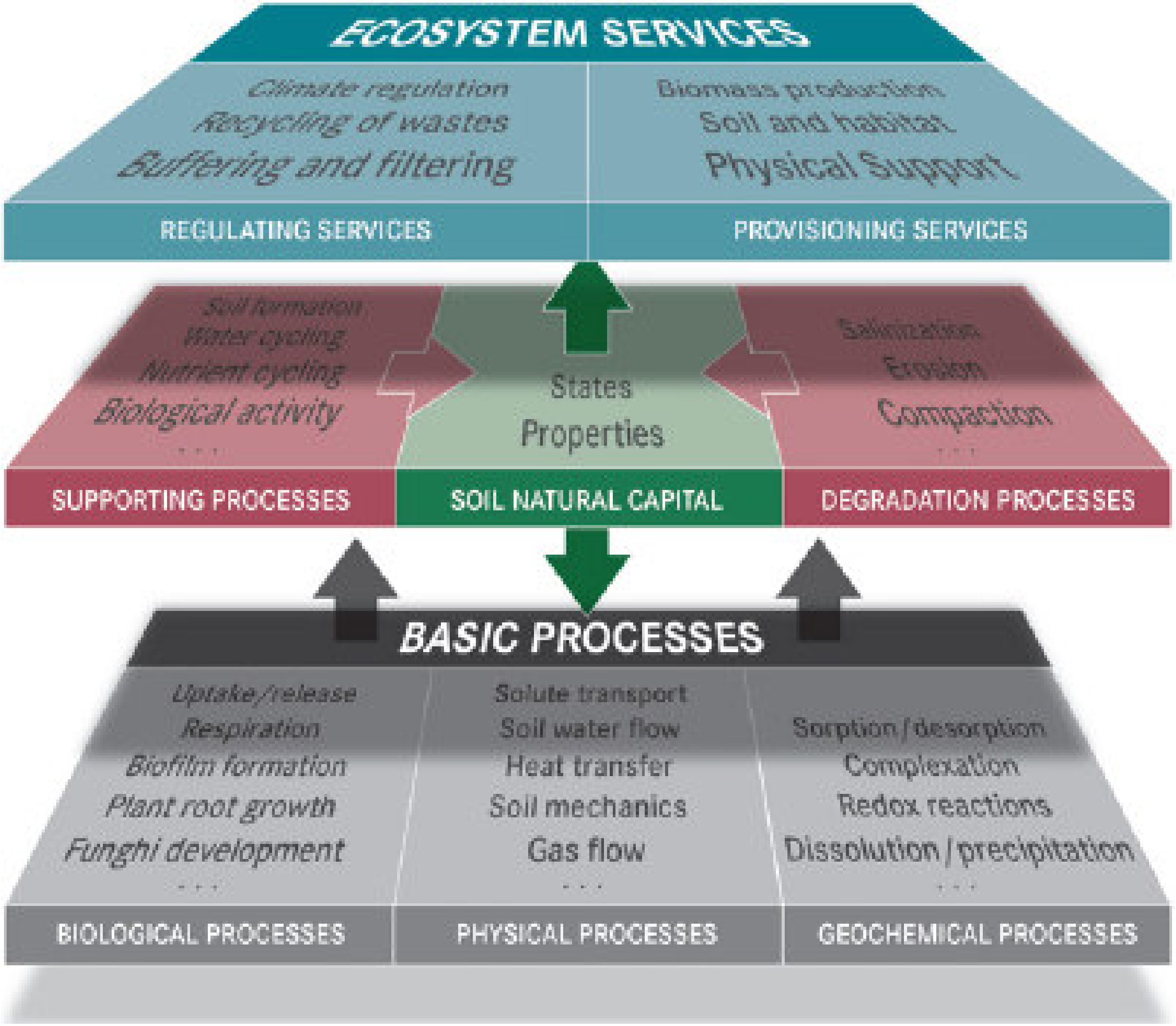
Core Ideas

- A community effort is needed to move soil modeling forward.
- Establishing an international soil modeling consortium is key in

Modeling Soil Processes: Review, Key Challenges, and New Perspectives

H. Vereecken,* A. Schnepf, J.W. Hopmans, M. Javaux, D. Or, T. Roose, J. Vanderborght, M.H. Young, W. Amelung, M. Aitkenhead, S.D. Allison, S. Assouline, P. Baveye, M. Berli, N. Brüggemann, P. Finke, M. Flury, T. Gaiser, G. Govers, T. Ghezzehei, P. Hallett, H.J. Hendricks Franssen, J. Heppell, R. Horn, J.A. Huisman, D. Jacques, F. Jonard, S. Kollet, F. Lafolie, K. Lamorski, D. Leitner, A. McBratney, B. Minasny, C. Montzka, W. Nowak, Y. Pachepsky, J. Padarian, N. Romano, K. Roth, Y. Rothfuss, E.C. Rowe, A. Schwen, J. Šimůnek, A. Tiktak, J. Van Dam, S.E.A.T.M. van der Zee, H.J. Vogel, J.A. Vrugt, T. Wöhling, and I.M. Young

Fundamental soil processes and their interactions remain lacking and deficient



Sources of Error in Modeling

- **Oversimplification and ignoring the vital processes**

Einstein: 'Everything should be as simple as possible, but not simpler ...

- **Applying an inefficient and incorrect mathematical model for the desired process**

H. L. Mencken: "For every complex problem there is an answer that is clear, simple, and wrong...

- **Inaccurate observations, measurements and model fit**

Einstein: "The only source of knowledge is experience"

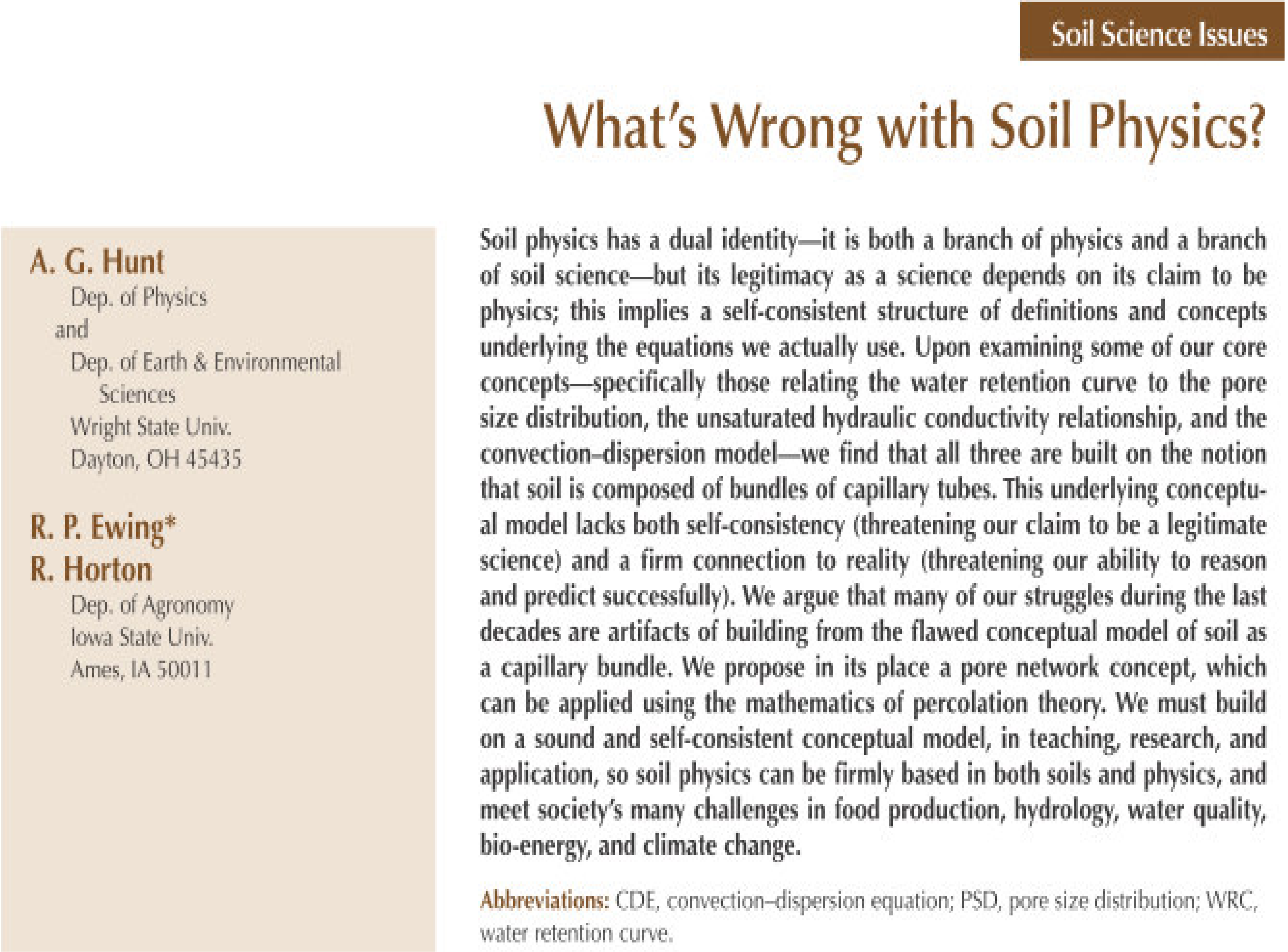
- **Incorrect evaluation of model accuracy**

- **Misinterpretation of results**

Example#1

Soil Physics:
Inappropriate Underlying Conceptual Model

Oversimplification
incorrect mathematical model



Soil Science Issues

What's Wrong with Soil Physics?

A. G. Hunt
Dep. of Physics
and
Dep. of Earth & Environmental Sciences
Wright State Univ.
Dayton, OH 45435

R. P. Ewing*
R. Horton
Dep. of Agronomy
Iowa State Univ.
Ames, IA 50011

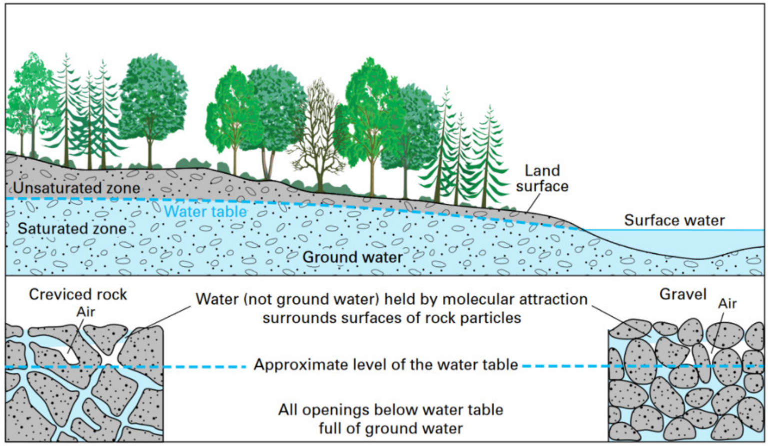
Soil physics has a dual identity—it is both a branch of physics and a branch of soil science—but its legitimacy as a science depends on its claim to be physics; this implies a self-consistent structure of definitions and concepts underlying the equations we actually use. Upon examining some of our core concepts—specifically those relating the water retention curve to the pore size distribution, the unsaturated hydraulic conductivity relationship, and the convection–dispersion model—we find that all three are built on the notion that soil is composed of bundles of capillary tubes. This underlying conceptual model lacks both self-consistency (threatening our claim to be a legitimate science) and a firm connection to reality (threatening our ability to reason and predict successfully). We argue that many of our struggles during the last decades are artifacts of building from the flawed conceptual model of soil as a capillary bundle. We propose in its place a pore network concept, which can be applied using the mathematics of percolation theory. We must build on a sound and self-consistent conceptual model, in teaching, research, and application, so soil physics can be firmly based in both soils and physics, and meet society's many challenges in food production, hydrology, water quality, bio-energy, and climate change.

Abbreviations: CDE, convection–dispersion equation; PSD, pore size distribution; WRC, water retention curve.

The Covert Capillary Concept

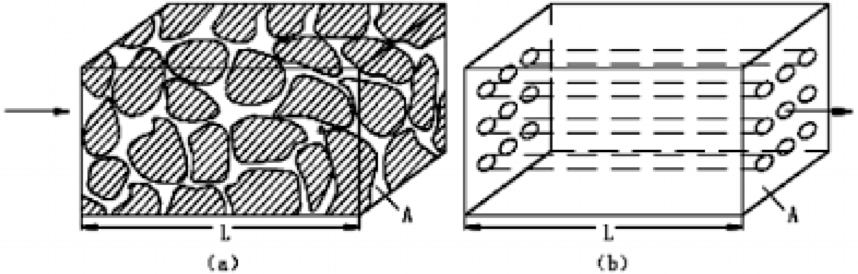
- **Engelmann and Huntoon (2011):**

Students—and often those who teach them—come to class with preconceptions and misconceptions that hinder their learning. For instance, students and their teachers believe groundwater exists in the ground in actual rivers or lakes. Such misconceptions need to be addressed before students can learn scientific concepts correctly.

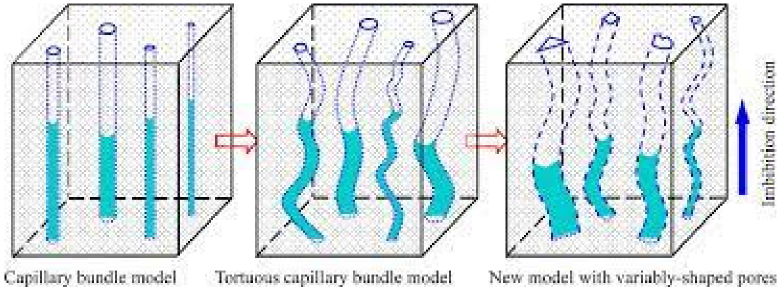


How ground water occurs in rocks.

- **Progress in soil physics is hindered by an analogous misconception held by many of the *scientists*, namely that capillary bundle models adequately describe water retention, flow, and transport properties of a porous medium.**



Dullien (1992):
The danger inherent in such models is that, owing to their simplicity, they become popular and some people may believe that they closely approximate reality, in this case, the actual pore structure. In fact, nothing could be further from the truth.



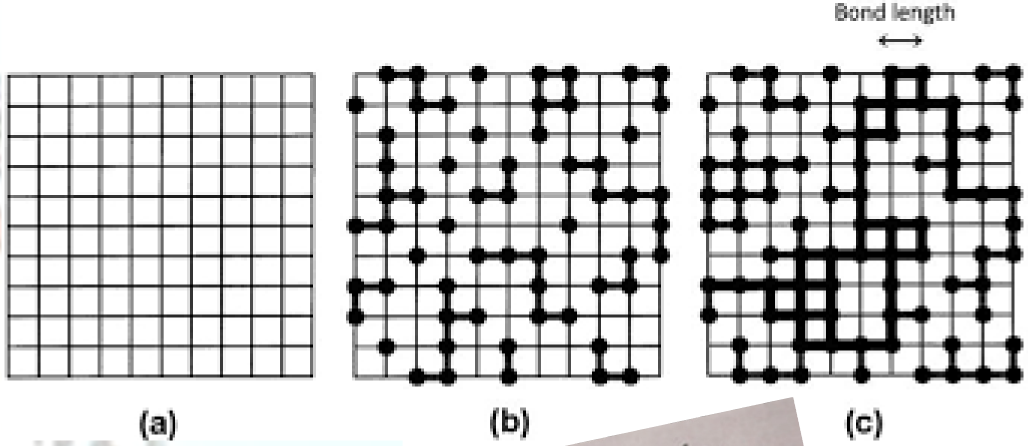
Capillary bundle model Tortuous capillary bundle model New model with variably-shaped pores

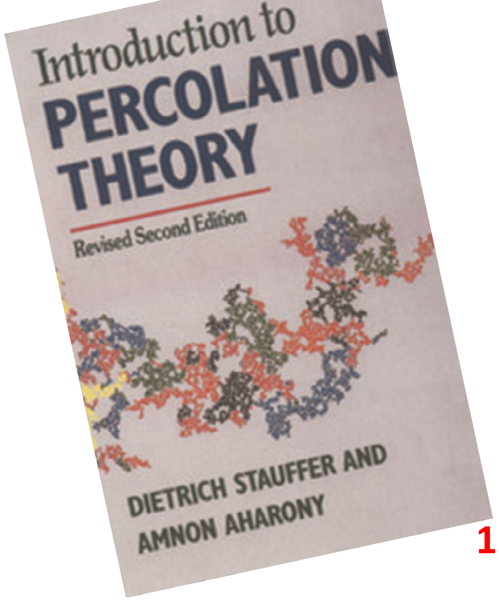
What’s Wrong with Soil Physics?

- **The underlying “wrongness” is the construction of soil physics on capillary bundle models, whether overt or covert.**
- **We must be prepared to think critically about even our most commonly used models and their associated mathematical methods.**
- **We must inquire whether our conceptual models are consistent across related phenomena; where they are not, we must consider that the inconsistency may be a flaw and an opportunity.**
- **The sooner we embrace these challenges, the sooner we will begin to enjoy the benefits of the resulting clearer perspective.**

Einstein, 1916:
Concepts that have proven useful in ordering things easily achieve such an authority over us that we forget their earthly origins and accept them as unalterable givens.

- **The “useful ordering of things” made possible by capillary bundle models has run its course, but fortunately the model is not an unalterable given.**
- **A more sensible ordering becomes possible using the connection-based perspective of percolation theory.**
- **We also expect that improving our foundational concepts will yield unexpected benefits, as some of our current difficulties will turn out to be simple artifacts of the flawed tools in current use.**
- **If we accept the challenge to rebuild our foundation, soil physics will have a bright future.**





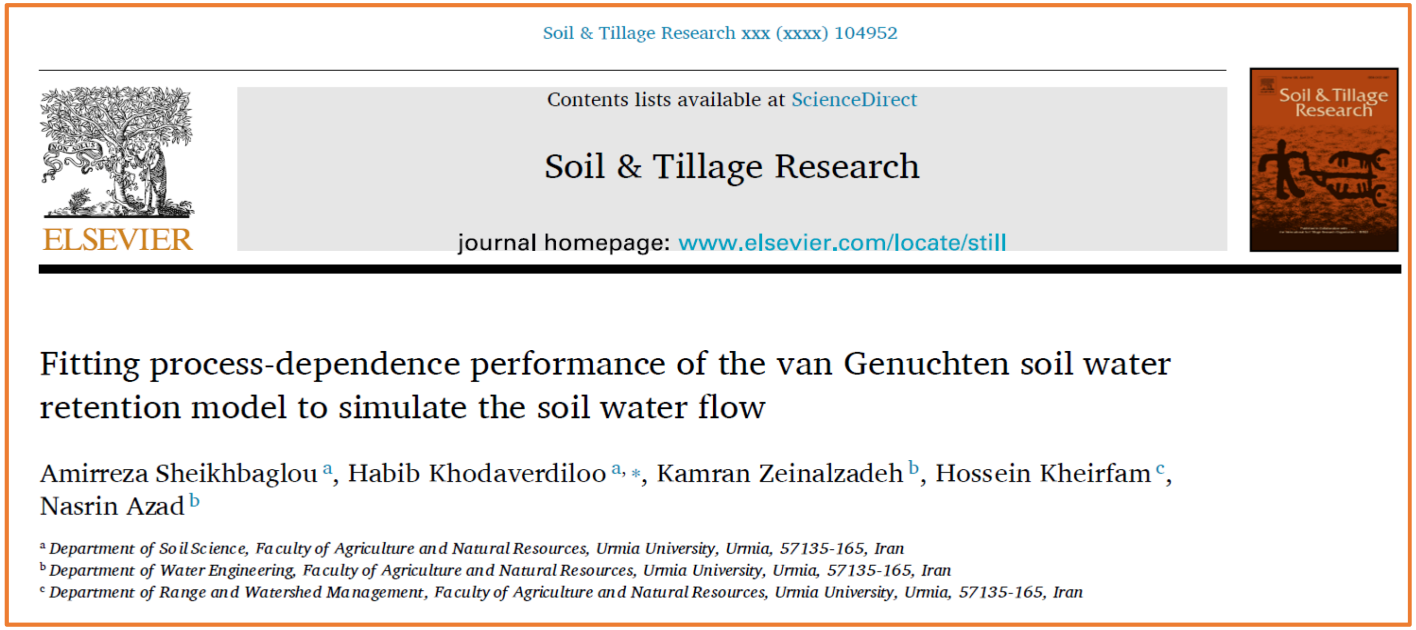
Example#2

Fitting process-dependence performance of the vG SWRC model to simulate the soil water flow and calculate the soil physical quality

#Inaccurate model fit

#Incorrect evaluation of model accuracy

#Misinterpretation of results



Soil & Tillage Research xxx (xxxx) 104952

Contents lists available at ScienceDirect

Soil & Tillage Research

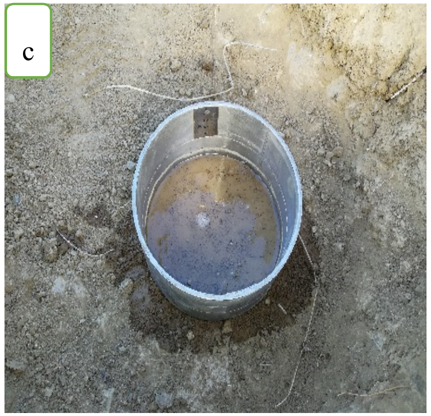
journal homepage: www.elsevier.com/locate/still

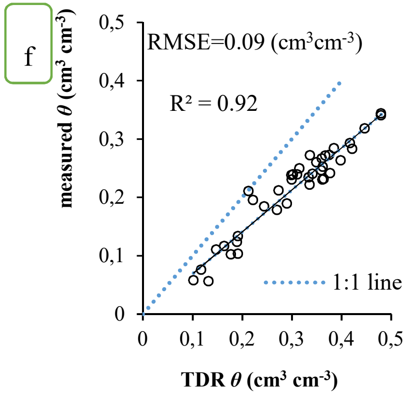
Fitting process-dependence performance of the van Genuchten soil water retention model to simulate the soil water flow

Amirreza Sheikhbaglou[a], Habib Khodaverdiloo[a,*], Kamran Zeinalzadeh[b], Hossein Kheirfam[c], Nasrin Azad[b]

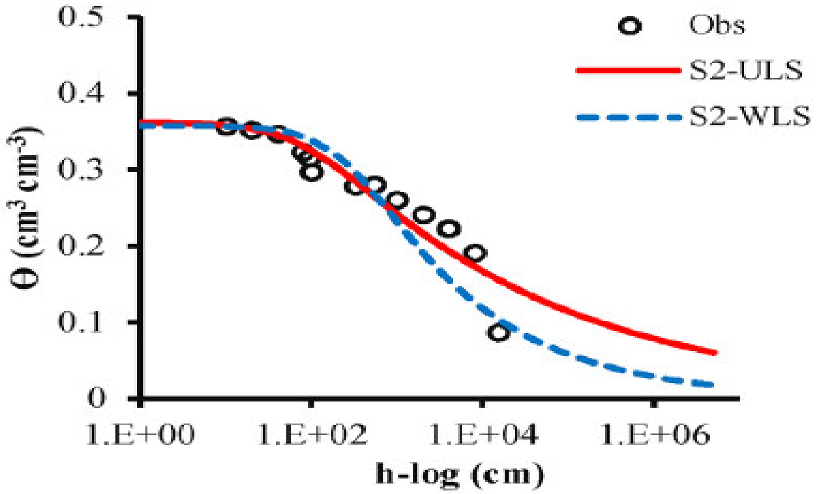
[a] Department of Soil Science, Faculty of Agriculture and Natural Resources, Urmia University, Urmia, 57135-165, Iran
[b] Department of Water Engineering, Faculty of Agriculture and Natural Resources, Urmia University, Urmia, 57135-165, Iran
[c] Department of Range and Watershed Management, Faculty of Agriculture and Natural Resources, Urmia University, Urmia, 57135-165, Iran

- **How varies the expected accuracy of water content measurements at different pressure heads in typical SWRC measurement experiments?**
- **How does the consideration of a different uncertainty of measured water contents affect the fit of hydraulic functions to the measured data (ULS vs. WLS)?**
- **How does the variability of the resulting SWRC then affect parameters that are commonly derived from it?**
- **What is the impact of using different hydraulic functions on the simulation of water movement (under different simulation scenarios)?**

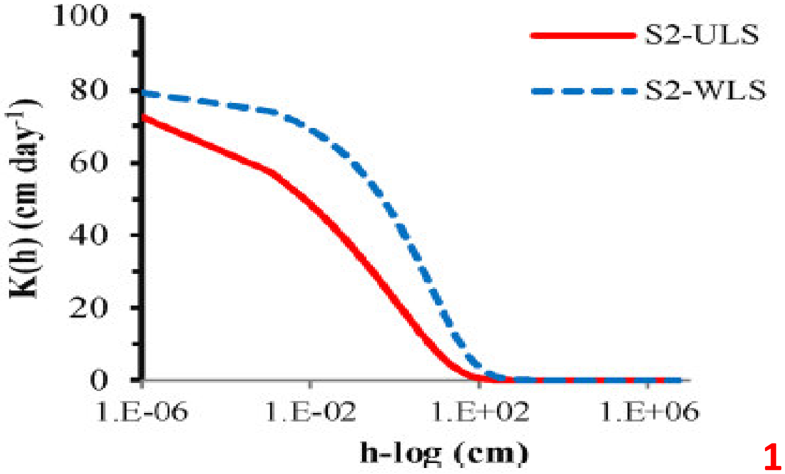


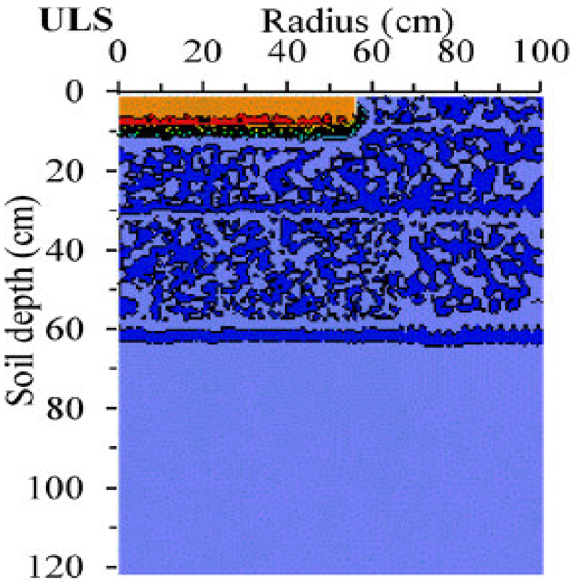


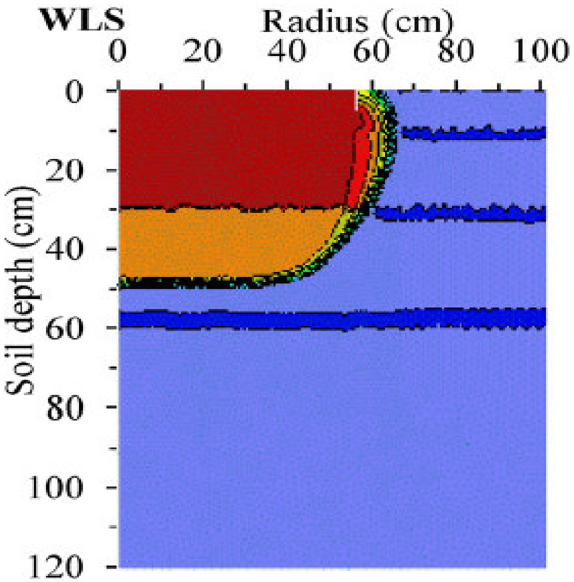
(a) SWRC-S2

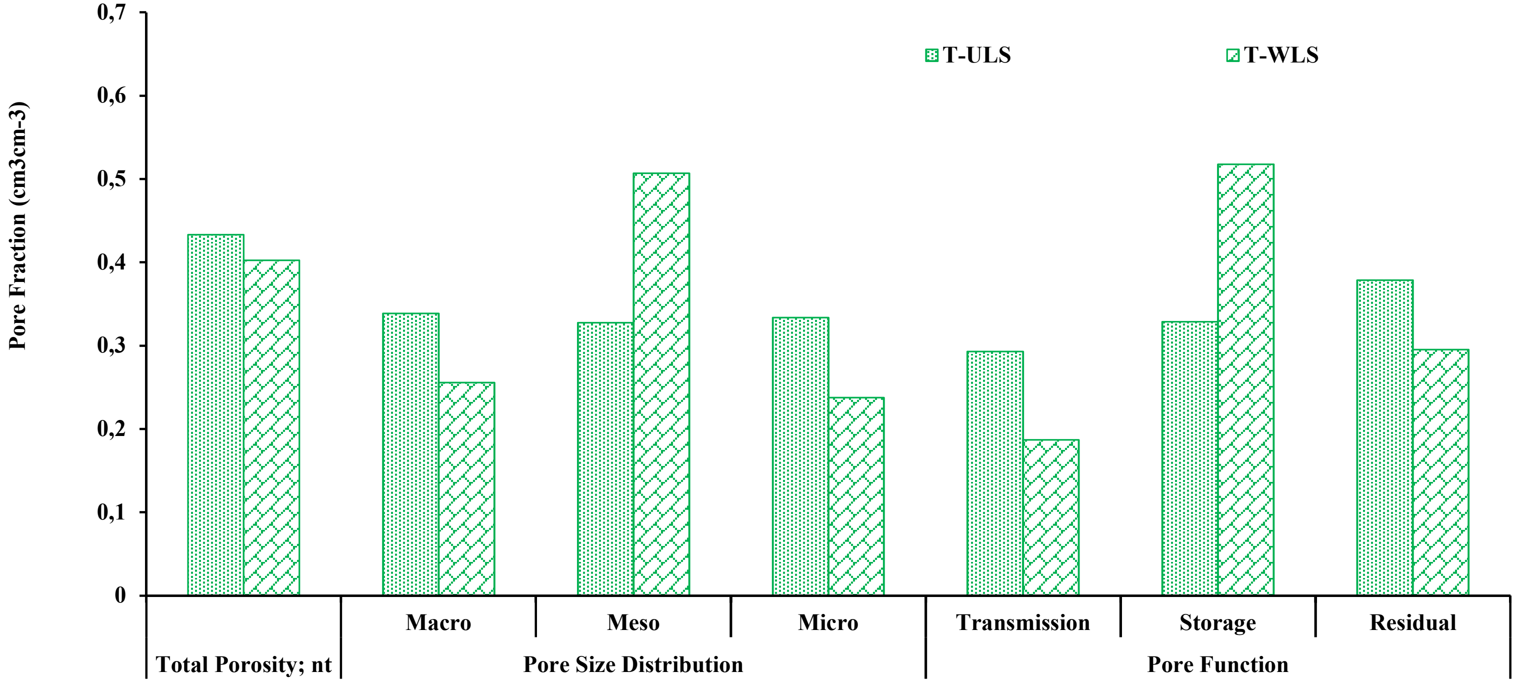


(c) K(h)-S2









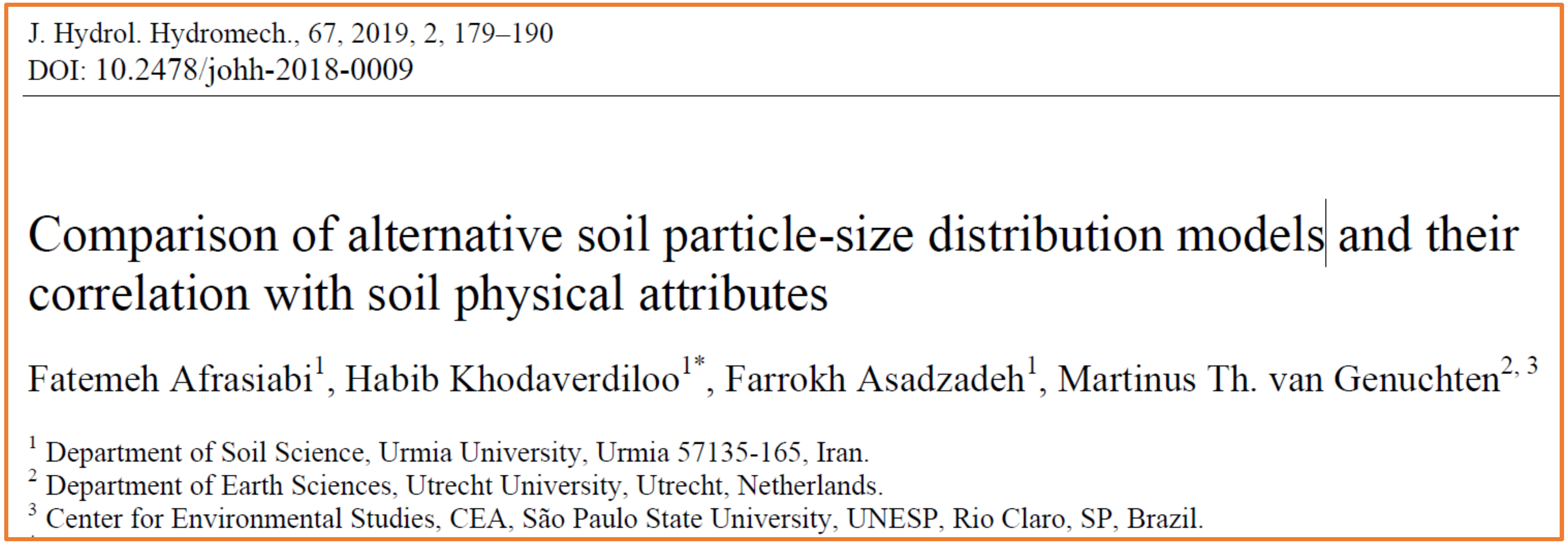
Method	θ_S (cm^3 cm^{-3})	θ_{FC} (cm^3 cm^{-3})	θ_{PWP} (cm^3 cm^{-3})	S-Index (-)	AC (cm^3 cm^{-3})	RFC (-)	PAW (330) (cm^3 cm^{-3})
ULS	0.432	0.246	0.142	0.030	0.186	0.577	0.104
WLS	0.407	0.237	0.096	0.041	0.170	0.589	0.141

Example#3

Comparison of alternative models: effects of used efficiency criteria

#Incorrect evaluation of model accuracy

#Misinterpretation of results



J. Hydrol. Hydromech., 67, 2019, 2, 179–190
DOI: 10.2478/johh-2018-0009

Comparison of alternative soil particle-size distribution models and their correlation with soil physical attributes

Fatemeh Afrasiabi[1], Habib Khodaverdiloo[1*], Farrokh Asadzadeh[1], Martinus Th. van Genuchten[2, 3]

[1] Department of Soil Science, Urmia University, Urmia 57135-165, Iran.
[2] Department of Earth Sciences, Utrecht University, Utrecht, Netherlands.
[3] Center for Environmental Studies, CEA, São Paulo State University, UNESP, Rio Claro, SP, Brazil.

Objectives

- To evaluate the performance of 19 models for describing PSD data of selected soils,
- To provide a functional evaluation of the models to predict selected practically-important PSD points or parameters using different efficiency criteria,
- To compare results obtained with the general and functional evaluations

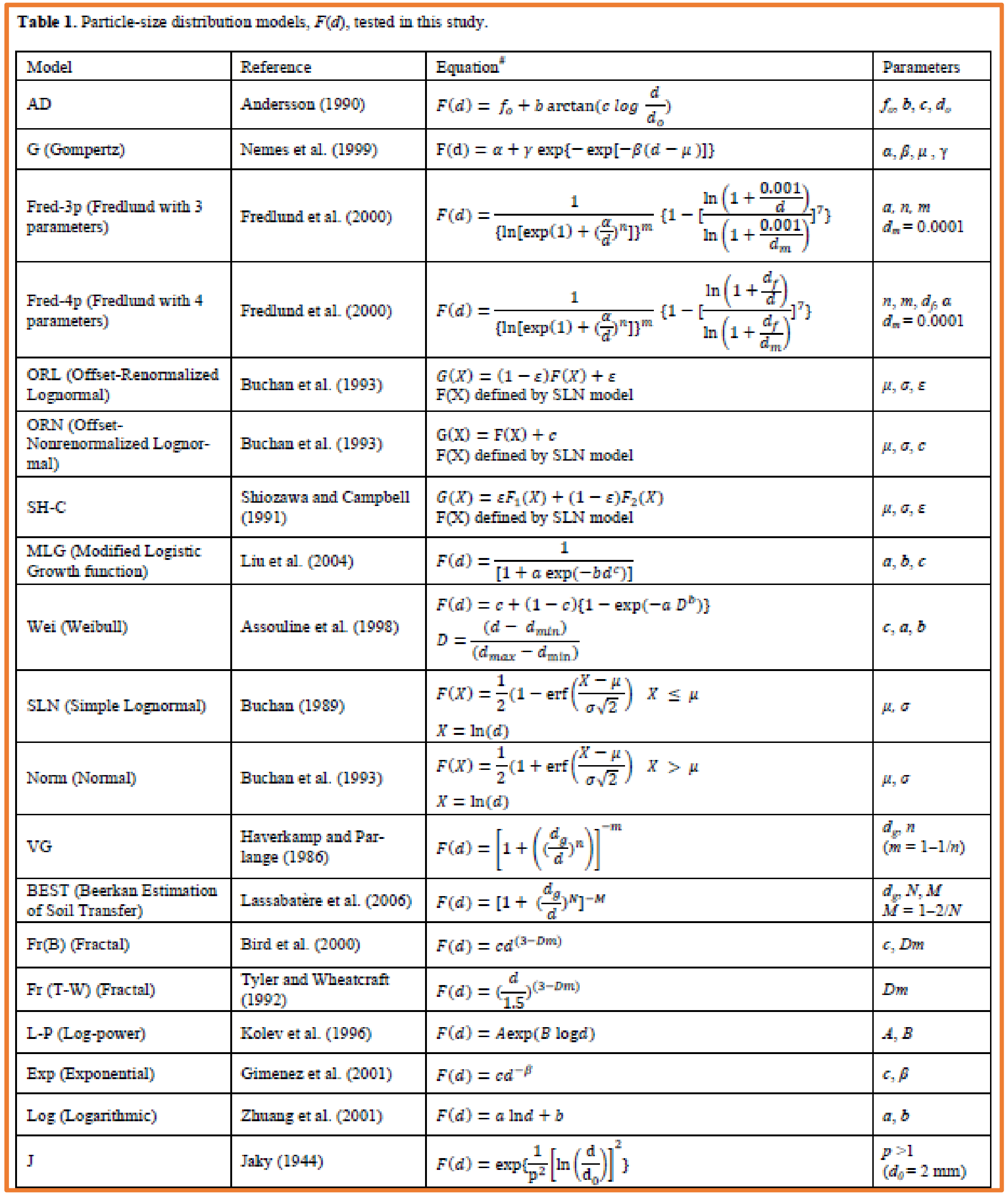


Table 1. Particle-size distribution models, *F*(*d*), tested in this study.

Model	Reference	Equation[#]	Parameters
AD	Andersson (1990)	$F(d) = f_o + b\arctan(c\log\frac{d}{d_o})$	f_∞, b, c, d_o
G (Gompertz)	Nemes et al. (1999)	$\mathrm{F(d)} = \alpha + \gamma\exp\{-\exp[-\beta(d-\mu)]\}$	$\alpha, \beta, \mu, \gamma$
Fred-3p (Fredlund with 3 parameters)	Fredlund et al. (2000)	$F(d) = \frac{1}{\{\ln[\exp(1) + (\frac{\alpha}{d})^n]\}^m}\{1 - [\frac{\ln\left(1+\frac{0.001}{d}\right)}{\ln\left(1+\frac{0.001}{d_m}\right)}]^7\}$	α, n, m $d_m = 0.0001$
Fred-4p (Fredlund with 4 parameters)	Fredlund et al. (2000)	$F(d) = \frac{1}{\{\ln[\exp(1) + (\frac{\alpha}{d})^n]\}^m}\{1 - [\frac{\ln\left(1+\frac{d_f}{d}\right)}{\ln\left(1+\frac{d_f}{d_m}\right)}]^7\}$	n, m, d_f, α $d_m = 0.0001$
ORL (Offset-Renormalized Lognormal)	Buchan et al. (1993)	$G(X) = (1-\varepsilon)F(X) + \varepsilon$ F(X) defined by SLN model	μ, σ, ε
ORN (Offset-Nonrenormalized Lognormal)	Buchan et al. (1993)	$\mathrm{G(X)} = \mathrm{F(X)} + c$ F(X) defined by SLN model	μ, σ, c
SH-C	Shiozawa and Campbell (1991)	$G(X) = \varepsilon F_1(X) + (1-\varepsilon)F_2(X)$ F(X) defined by SLN model	μ, σ, ε
MLG (Modified Logistic Growth function)	Liu et al. (2004)	$F(d) = \frac{1}{[1 + a\exp(-bd^c)]}$	a, b, c
Wei (Weibull)	Assouline et al. (1998)	$F(d) = c + (1-c)\{1 - \exp(-a\,D^b)\}$ $D = \frac{(d - d_{min})}{(d_{max} - d_{\min})}$	c, a, b
SLN (Simple Lognormal)	Buchan (1989)	$F(X) = \frac{1}{2}(1 - \mathrm{erf}\left(\frac{X-\mu}{\sigma\sqrt{2}}\right) \quad X \le \mu$ $X = \ln(d)$	μ, σ
Norm (Normal)	Buchan et al. (1993)	$F(X) = \frac{1}{2}(1 + \mathrm{erf}\left(\frac{X-\mu}{\sigma\sqrt{2}}\right) \quad X > \mu$ $X = \ln(d)$	μ, σ
VG	Haverkamp and Parlange (1986)	$F(d) = \left[1 + \left((\frac{d_g}{d})^n\right)\right]^{-m}$	d_g, n $(m = 1\text{–}1/n)$
BEST (Beerkan Estimation of Soil Transfer)	Lassabatère et al. (2006)	$F(d) = [1 + (\frac{d_g}{d})^N]^{-M}$	d_g, N, M $M = 1\text{–}2/N$
Fr(B) (Fractal)	Bird et al. (2000)	$F(d) = cd^{(3-Dm)}$	c, Dm
Fr (T-W) (Fractal)	Tyler and Wheatcraft (1992)	$F(d) = (\frac{d}{1.5})^{(3-Dm)}$	Dm
L-P (Log-power)	Kolev et al. (1996)	$F(d) = A\exp(B\log d)$	A, B
Exp (Exponential)	Gimenez et al. (2001)	$F(d) = cd^{-\beta}$	c, β
Log (Logarithmic)	Zhuang et al. (2001)	$F(d) = a\ln d + b$	a, b
J	Jaky (1944)	$F(d) = \exp\{\frac{1}{\mathrm{p}^2}\left[\ln\left(\frac{\mathrm{d}}{\mathrm{d}_0}\right)\right]^2\}$	$p > 1$ $(d_0 = 2\ \mathrm{mm})$

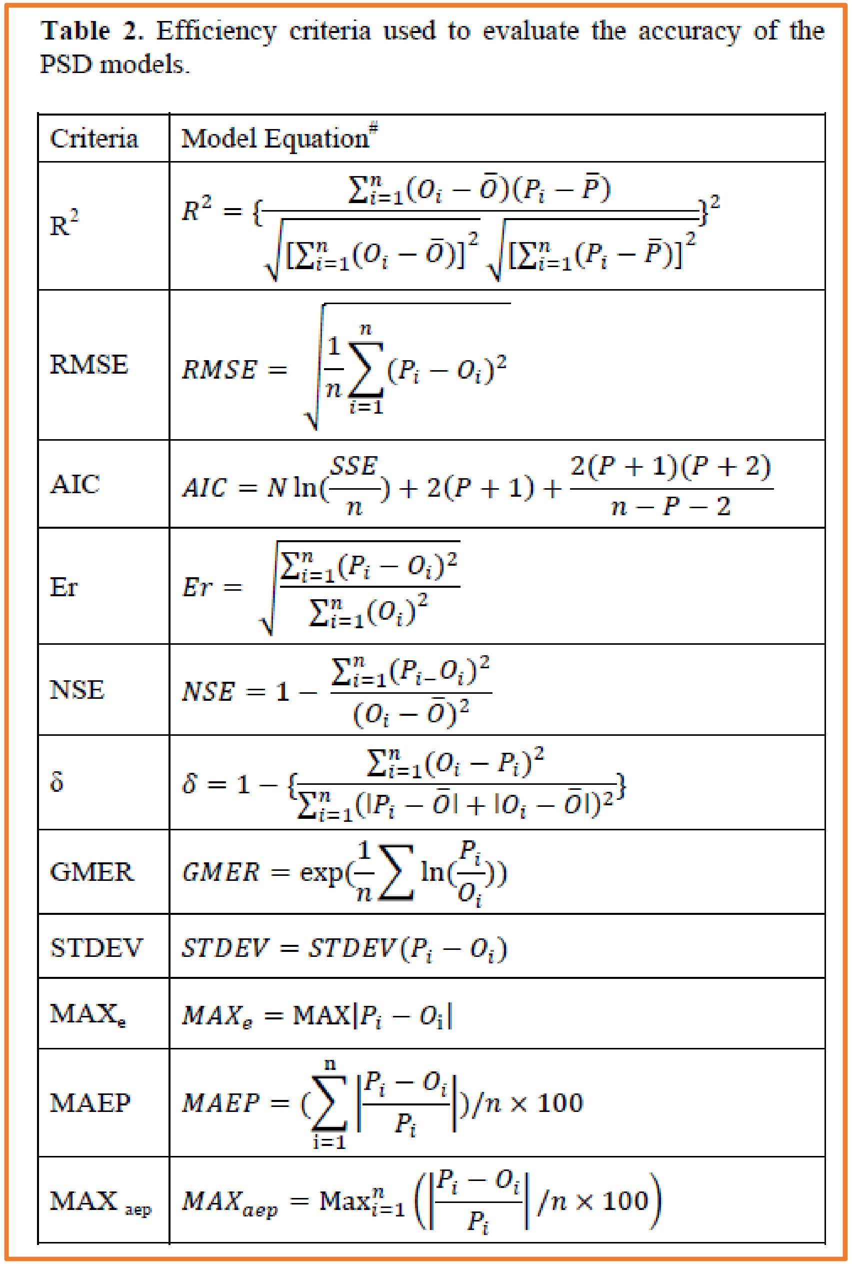


Table 2. Efficiency criteria used to evaluate the accuracy of the PSD models.

Criteria	Model Equation[#]
R^2	$R^2 = \{\frac{\sum_{i=1}^{n}(O_i - \bar{O})(P_i - \bar{P})}{\sqrt{[\sum_{i=1}^{n}(O_i - \bar{O})]^2}\sqrt{[\sum_{i=1}^{n}(P_i - \bar{P})]^2}}\}^2$
RMSE	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2}$
AIC	$AIC = N\ln(\frac{SSE}{n}) + 2(P+1) + \frac{2(P+1)(P+2)}{n-P-2}$
Er	$Er = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{\sum_{i=1}^{n}(O_i)^2}}$
NSE	$NSE = 1 - \frac{\sum_{i=1}^{n}(P_{i-}O_i)^2}{(O_i - \bar{O})^2}$
δ	$\delta = 1 - \{\frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(\lvert P_i - \bar{O}\rvert + \lvert O_i - \bar{O}\rvert)^2}\}$
GMER	$GMER = \exp(\frac{1}{n}\sum\ln(\frac{P_i}{O_i}))$
STDEV	$STDEV = STDEV(P_i - O_i)$
MAX$_e$	$MAX_e = \mathrm{MAX}\lvert P_i - O_i\rvert$
MAEP	$MAEP = (\sum_{i=1}^{n}\left\lvert\frac{P_i - O_i}{P_i}\right\rvert)/n \times 100$
MAX $_{aep}$	$MAX_{aep} = \mathrm{Max}_{i=1}^{n}\left(\left\lvert\frac{P_i - O_i}{P_i}\right\rvert/n \times 100\right)$

Table 4. Rankings of the six most accurate PSD models in terms of various statistical criteria.

Ranking	R^2		AIC	RMSE	Er	d	MAX_{aep}	NSE	STDEV	GMER	MAX_e	MAEP
1	AD		Wei	AD	Fred-4p	Fred-4p	Fred-4p	VG	Fred-4p	ONL	Fred-4p	AD
2	Wei		Fred-3p	Fred-4p	MLG	Wei	Wei	Fred-4p	AD	ORL	Fred-3p	Wei
3	Fred-4p		AD	Wei	Fred-3p	Fred-3p	MLG	Wei	Wei	AD	MLG	Fred-4p
4	Fred-3p		MLG	Fred-3p	AD	MLG	Fred-3p	AD	Fred-3p	Wei	AD	ONL
5	MLG		Fred-4p	MLG	ONL	ONL	AD	Fred-3p	MLG	VG	Wei	Fred-3P
6	ONL		ORL	ONL	ORL	ORL	ONL	MLG	ONL	Log	ORL	MLG

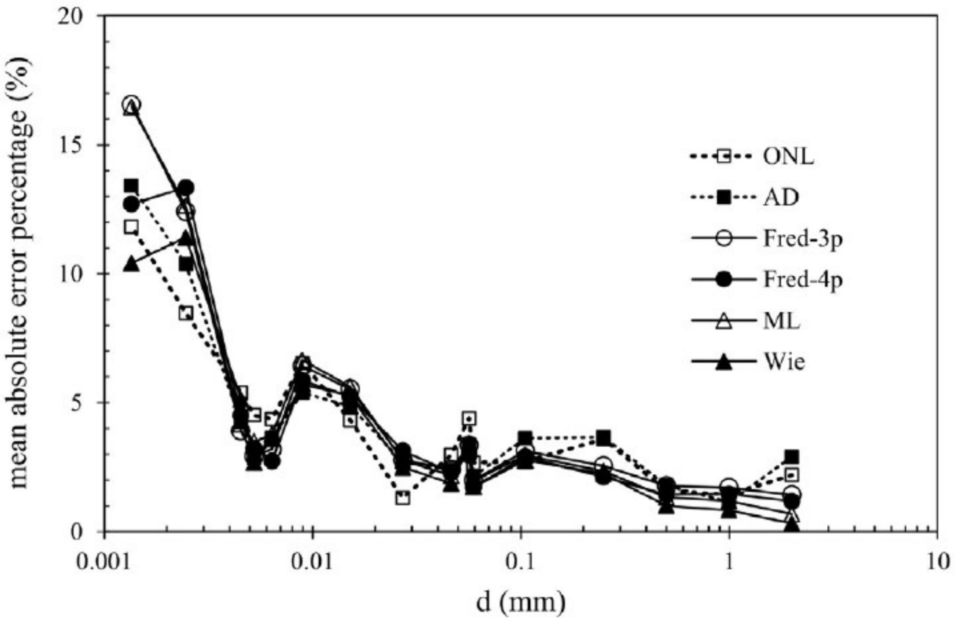


Fig. 3. Mean absolute error percentages of the six most accurate models as a function of soil particle diameter.

Example#4

Sorption Studies and Mass Transport Simulations

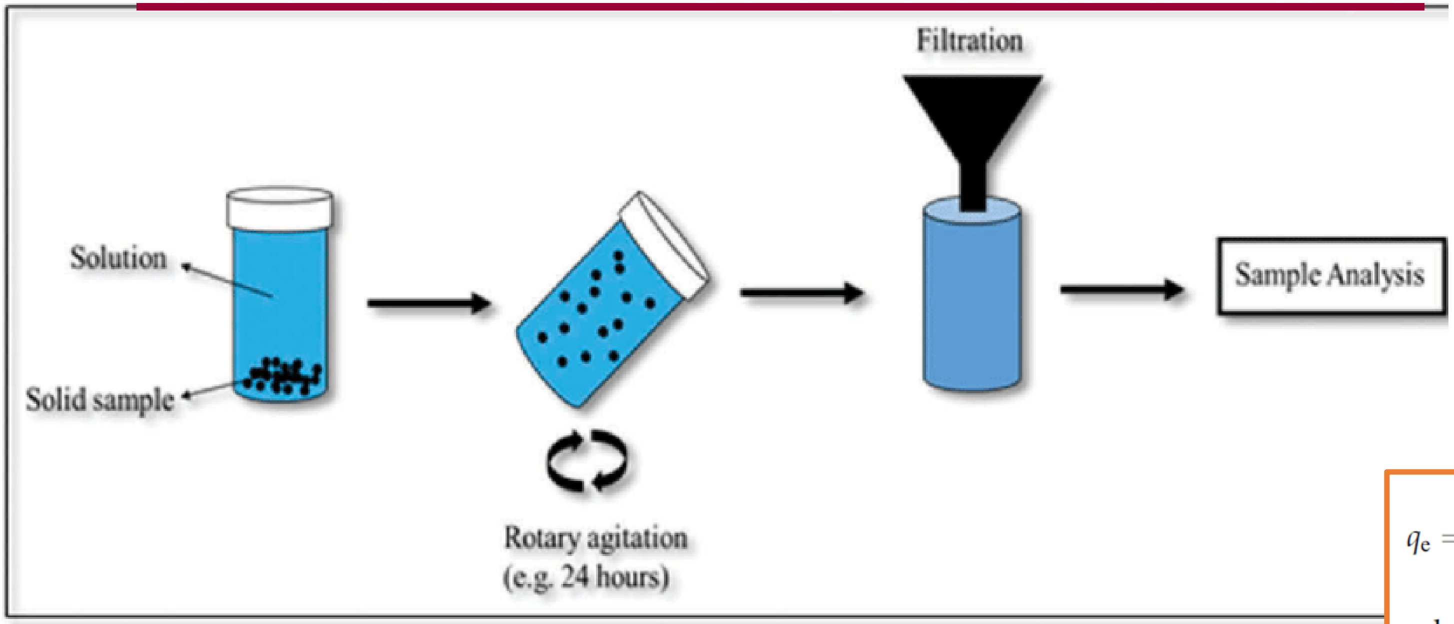
#Inaccurate Measurements

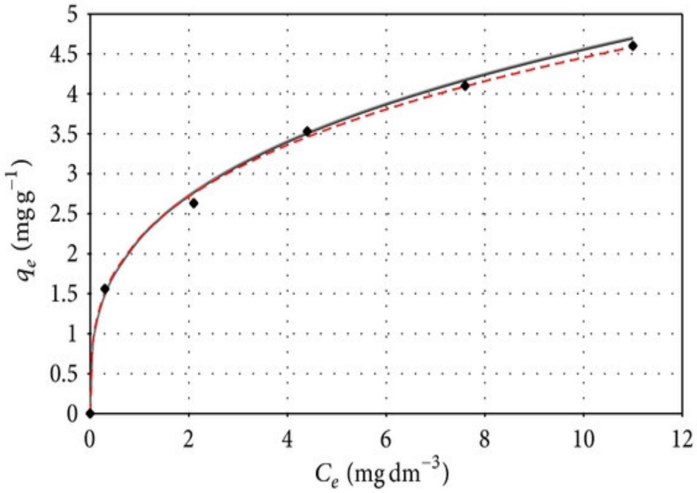
#Oversimplified Models

#Misevaluation

#Misinterpretation

PRINCIPLE OF THE METHOD





$$q_e = \frac{(C_i - C_e) \times V}{1000 \times m} \quad (3)$$

where q_e is the equilibrium sorption capacity (mg/g), V is the volume of phosphate solution (ml), m is the mass of peat used (g), and C_i and C_e are the initial and final (equilibrium) concentrations (mg/l).

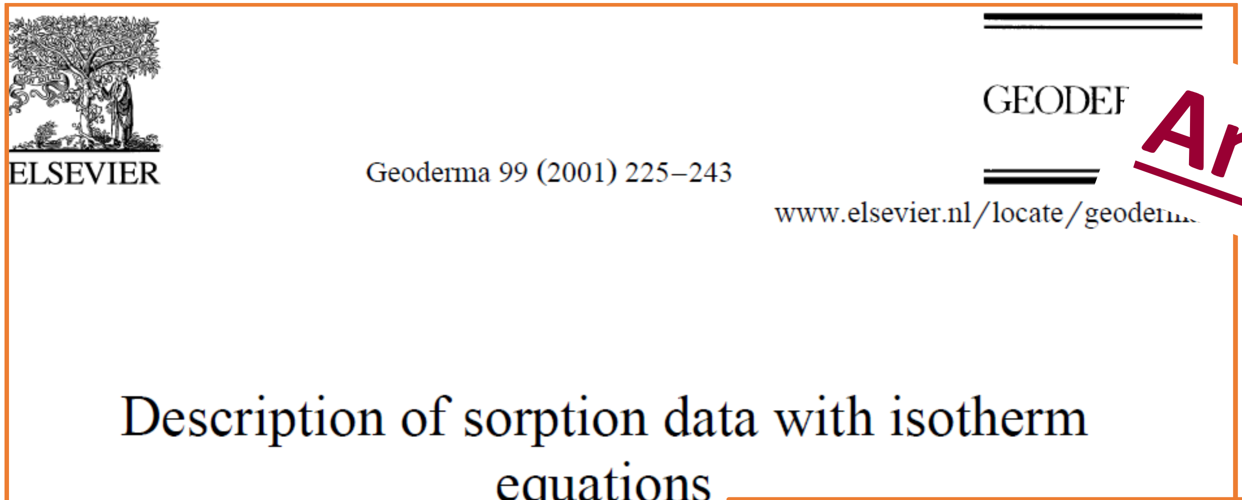
1. Known volumes of solutions of the test substance at known concentrations in a BG solution are added to soil samples of known dry weight.

2. The mixture is agitated for an appropriate time.

3. The soil suspensions are then separated by centrifugation and the aqueous phase is analysed.

4. The amount of test substance adsorbed on the soil sample is calculated as the difference between the amount of test substance initially present in solution and the amount remaining at the end of the experiment (indirect method).

Analyzing sorption data



ELSEVIER

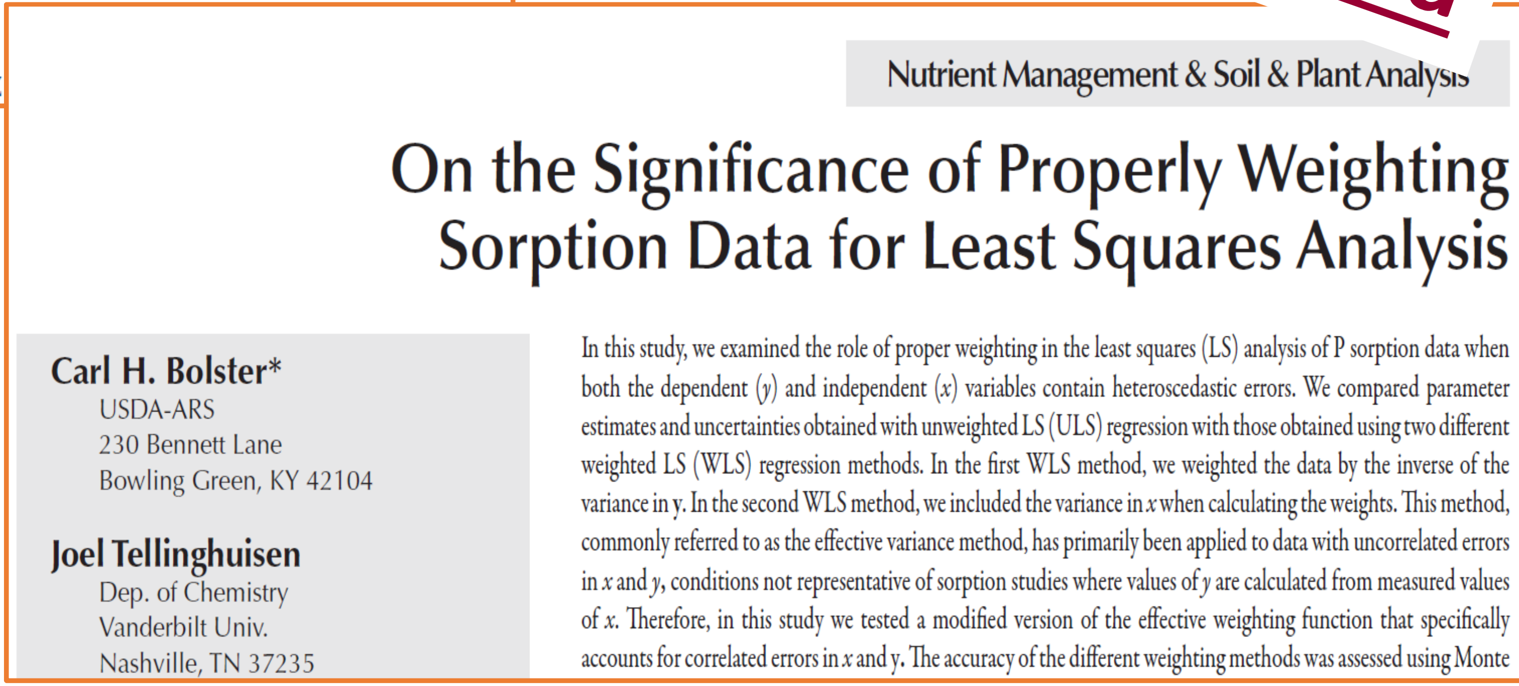
Geoderma 99 (2001) 225–243

GEODE

www.elsevier.nl/locate/geoderm

Description of sorption data with isotherm equations

Christoph Hinz



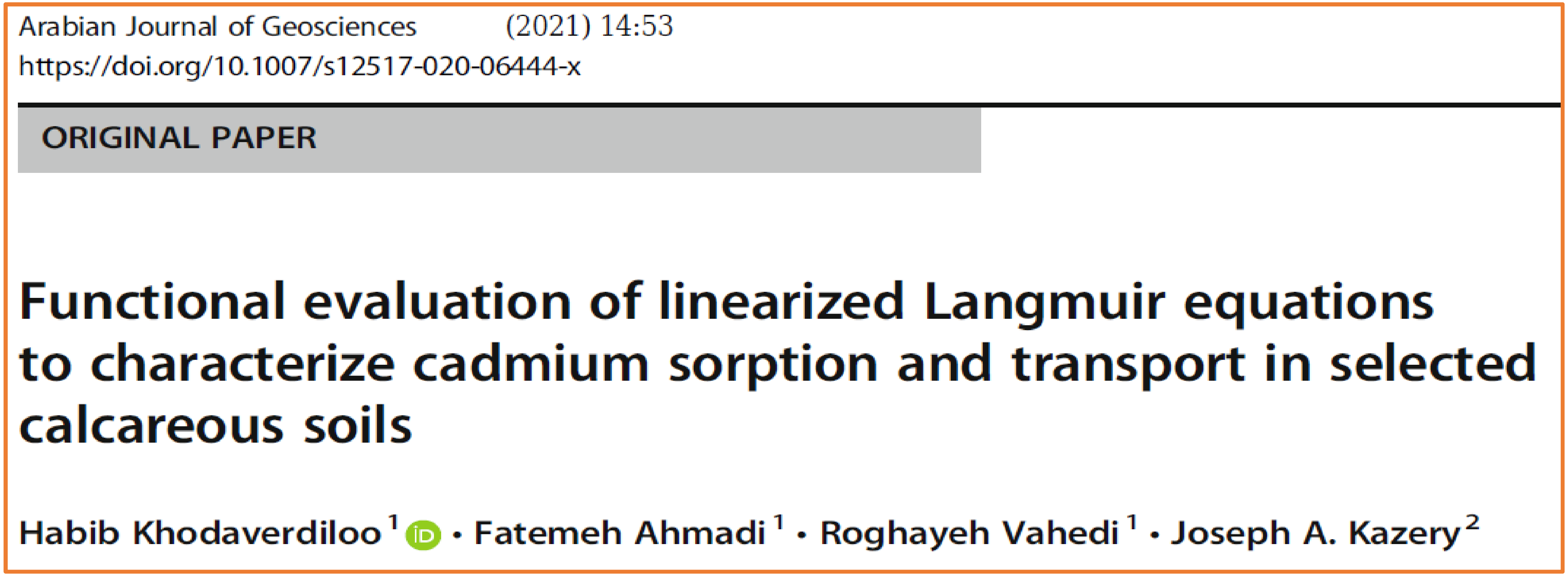
Nutrient Management & Soil & Plant Analysis

On the Significance of Properly Weighting Sorption Data for Least Squares Analysis

Carl H. Bolster*
USDA-ARS
230 Bennett Lane
Bowling Green, KY 42104

Joel Tellinghuisen
Dep. of Chemistry
Vanderbilt Univ.
Nashville, TN 37235

In this study, we examined the role of proper weighting in the least squares (LS) analysis of P sorption data when both the dependent (*y*) and independent (*x*) variables contain heteroscedastic errors. We compared parameter estimates and uncertainties obtained with unweighted LS (ULS) regression with those obtained using two different weighted LS (WLS) regression methods. In the first WLS method, we weighted the data by the inverse of the variance in y. In the second WLS method, we included the variance in *x* when calculating the weights. This method, commonly referred to as the effective variance method, has primarily been applied to data with uncorrelated errors in *x* and *y*, conditions not representative of sorption studies where values of *y* are calculated from measured values of *x*. Therefore, in this study we tested a modified version of the effective weighting function that specifically accounts for correlated errors in *x* and y. The accuracy of the different weighting methods was assessed using Monte



Arabian Journal of Geosciences (2021) 14:53
https://doi.org/10.1007/s12517-020-06444-x

ORIGINAL PAPER

Functional evaluation of linearized Langmuir equations to characterize cadmium sorption and transport in selected calcareous soils

Habib Khodaverdiloo[1] · Fatemeh Ahmadi[1] · Roghayeh Vahedi[1] · Joseph A. Kazery[2]

- **We compared the performance of the nonlinear and linearized Langmuir models to fit the experimental data of Cd sorption to soils.**

- **Using HYDRUS program, we functionally evaluated if the sorption parameter values obtained, either by nonlinear fitting of the Langmuir model or by its linearized alternatives, will affect the simulation of Cd transport in soil.**

$$S = \frac{S_{max} \times K \times C}{1 + K \times C}$$

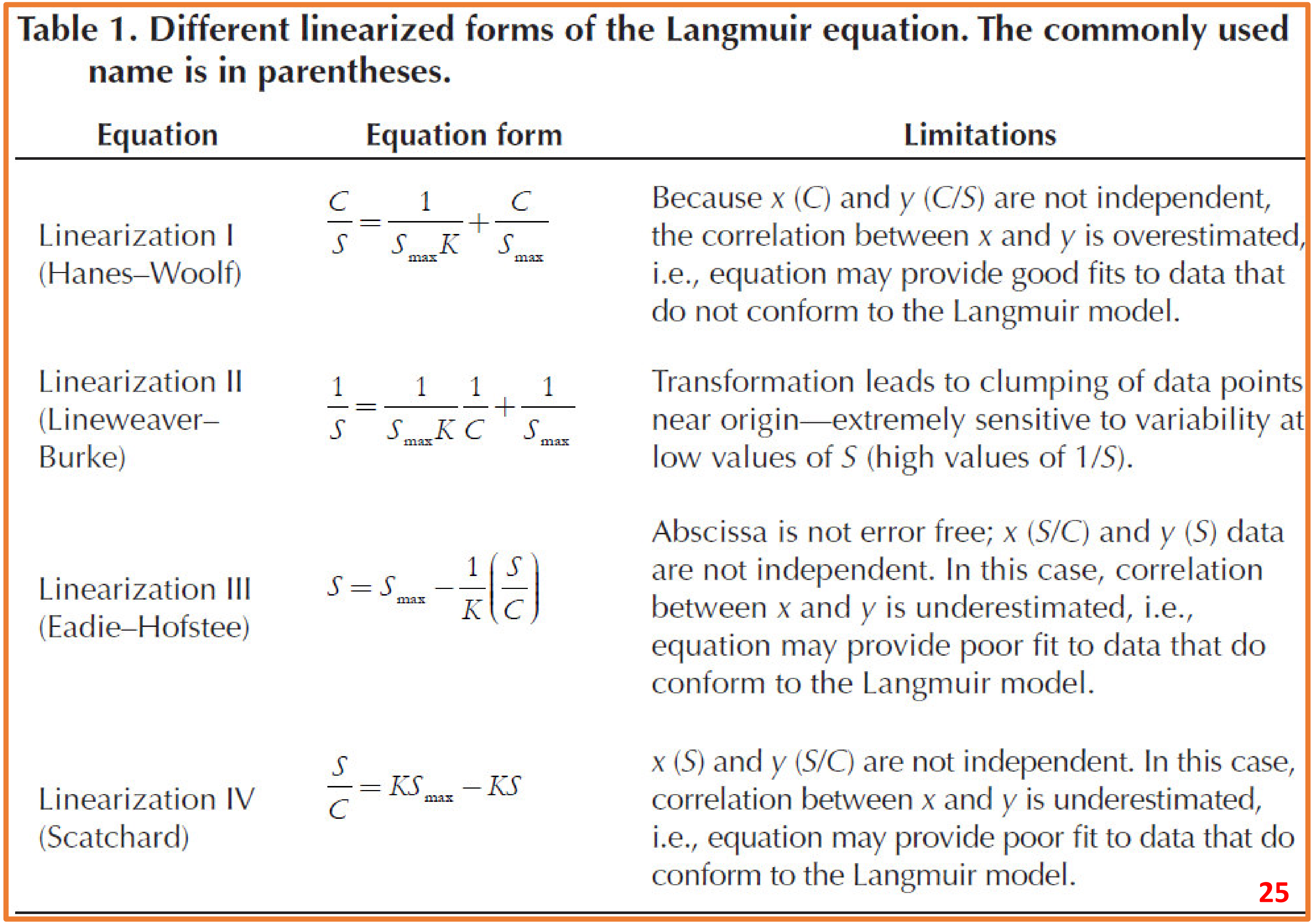


Table 1. Different linearized forms of the Langmuir equation. The commonly used name is in parentheses.

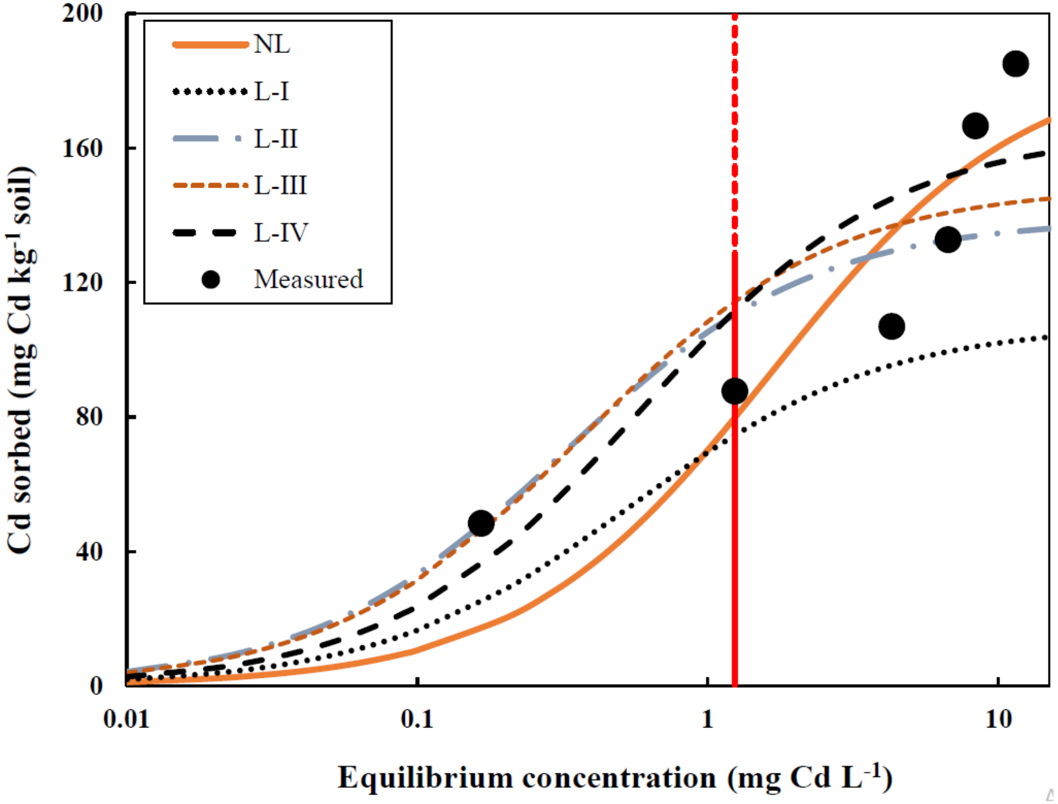
Equation	Equation form	Limitations
Linearization I (Hanes–Woolf)	$\frac{C}{S} = \frac{1}{S_{max}K} + \frac{C}{S_{max}}$	Because *x* (*C*) and *y* (*C*/*S*) are not independent, the correlation between *x* and *y* is overestimated, i.e., equation may provide good fits to data that do not conform to the Langmuir model.
Linearization II (Lineweaver–Burke)	$\frac{1}{S} = \frac{1}{S_{max}K}\frac{1}{C} + \frac{1}{S_{max}}$	Transformation leads to clumping of data points near origin—extremely sensitive to variability at low values of *S* (high values of 1/*S*).
Linearization III (Eadie–Hofstee)	$S = S_{max} - \frac{1}{K}\left(\frac{S}{C}\right)$	Abscissa is not error free; *x* (*S*/*C*) and *y* (*S*) data are not independent. In this case, correlation between *x* and *y* is underestimated, i.e., equation may provide poor fit to data that do conform to the Langmuir model.
Linearization IV (Scatchard)	$\frac{S}{C} = KS_{max} - KS$	*x* (*S*) and *y* (*S*/*C*) are not independent. In this case, correlation between *x* and *y* is underestimated, i.e., equation may provide poor fit to data that do conform to the Langmuir model.

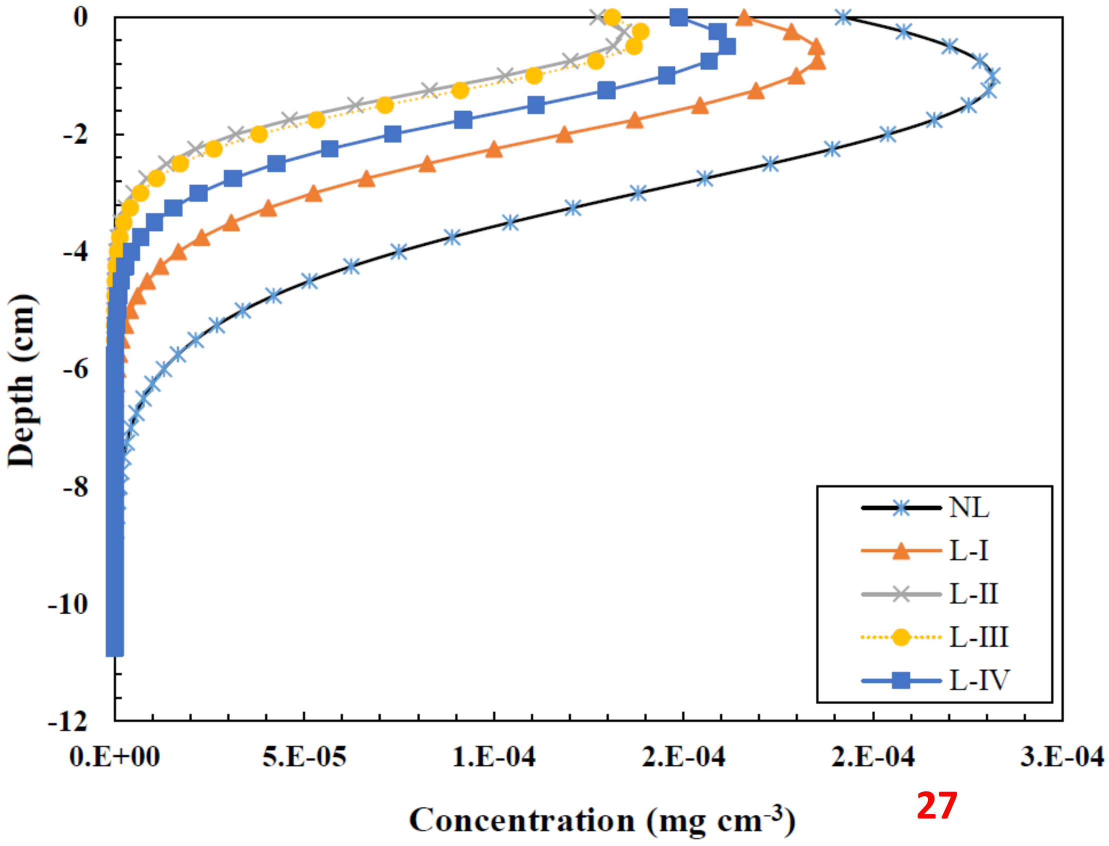
Table 3. Fitted sorption parameters, *RMSE* and R^2 for different Langmuir isotherm models (n = 19).

Langmuir models[#]	K (L mg^{-1})	S_m (mg kg^{-1})	R^2	*RMSE* (mg kg^{-1})
NL	0.6484^{b}	396.65^{a}	0.89263^{a}	14.58^{c}
L-I	1.6495^{a}	220.99^{c}	0.89211^{a}	37.42^{c}
L-II	1.3571^{a}	250.91^{c}	0.95474^{a}	27.53^{c}
L-III	1.3476^{a}	274.56^{bc}	0.70211^{b}	250.36^{b}
L-IV	0.8771^{b}	323.17^{b}	0.70211b	1518.46a

Means followed by the same letters in the each column are not significantly different according to Duncan's multiple range test at the level $p<0.05$.

[#] NL specifies nonlinear and L-I to L-IV specifies linearized Langmuir models (see Table 1).


Cd sorbed (mg Cd kg^{-1} soil)
200
160
120
80
40
0
0.01
0.1
1
10
Equilibrium concentration (mg Cd L^{-1})
NL
L-I
L-II
L-III
L-IV
Measured


Depth (cm)
0
-2
-4
-6
-8
-10
-12
0.E+00
5.E-05
1.E-04
2.E-04
2.E-04
3.E-04
Concentration (mg cm^{-3})
NL
L-I
L-II
L-III
L-IV

ENVOI

R. Hamming: “The purpose of computing is insight, not numbers”

- **We should use models but not have absolute trust in them.**
- **We, as teachers, need to shift from the passive “Sage-on-the-Stage” teaching approach to the more active “Guide-on-the-Side” method.**
- **All of us, both as teachers or students, must critically think about the existing models/methods/principals and try to rebuild a more realistic and scientificly-sound pillars for the soil science.**

Thank You

For Your Attention