Title

The time validity of Philip's two-term infiltration equation: An elusive theoretical quantity?

Authors

Vrugt, Jasper A

Hopmans, Jan W

Gao, Yifu

et al.

Publication Date

2024

DOI

10.1002/vzj2.20309

Received: 20 September 2023 | Accepted: 29 December 2023

DOI: 10.1002/vzj2.20309

Vadose Zone Journal

ORIGINAL ARTICLE

Special Section: Tribute to Rien van Genuchten, Recipient of the 2023 Wolf Prize for Agriculture

The time validity of Philip's two-term infiltration equation: An elusive theoretical quantity?

Jasper A. Vrugt[1] | **Jan W. Hopmans**[2,5] | **Yifu Gao**[1] | **Mehdi Rahmati**[3,4] | **Jan Vanderborght**[4] | **Harry Vereecken**[4]

[1]Department of Civil and Environmental Engineering, University of California, Irvine, California, USA

[2]Department of Land Air and Water Resources, University of California, Davis, California, USA

[3]Department of Soil Science and Engineering, Faculty of Agriculture, University of Maragheh, Maragheh, Iran

[4]Forschungszentrum Jülich GmbH, Institute of Bio- and Geosciences: Agrosphere (IBG-3), Jülich, Germany

[5]Department of Agricultural & Life Sciences, Oregon State University, Corvallis, Oregon, USA

Correspondence
Jasper A. Vrugt, Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA.
Email: jasper@uci.edu

Assigned to Associate Editor Dr. Binayak Mohanty.

Abstract

The two-term infiltration equation $I(t) = S\sqrt{t} + At$ is commonly used to determine the sorptivity, S ($LT^{-1/2}$), and product, $A = cK_s$ (LT^{-1}), of the dimensionless multiple c and saturated soil hydraulic conductivity K_s (LT^{-1}) from cumulative vertical infiltration measurements $\tilde{I}_1, \ldots, \tilde{I}_n$ (L) at times $t_1, \ldots, t_n$ (T). This reduced form of the quasi-analytical power series solution of Richardson's equation of Philip enjoys a solid physical underpinning but at the expense of a limited time validity. Using simulated infiltration data, Jaiswal et al. have shown this time validity to equal about 2.5 cm of cumulative infiltration. The goals of this work are twofold. First, we investigate the extent to which cumulative infiltration measurements larger than 2.5 cm bias the estimates of S and K_s. Second, we investigate the impact of epistemic errors on the inferred time validities and parameters. Partial infiltration curves up to 2.5 cm of cumulative vertical infiltration improve substantially the agreement between actual and least squares estimates of S and K_s. But this only holds if the data generating infiltration process follows Richardson's equation and experimental conditions satisfy assumptions of soil homogeneity and a uniform initial water content. Otherwise, autocorrelated cumulative infiltration residuals will bias the least squares estimates of S and K_s. Our findings reiterate and reinvigorate earlier conclusions of Haverkamp et al. and show that epistemic errors deteriorate the physical significance of the coefficients of infiltration functions. As a result, the parameters of infiltration functions cannot simply be used in storm water and vadose zone flow models to forecast runoff and recharge at field and landscape scales unless these predictions are accompanied by realistic uncertainty bounds. We conclude that the time validity of Philip's two-term equation is an elusive theoretical quantity with arbitrary physical meaning.

Abbreviations: LM, Levenberg–Marquardt; MVG, Mualem–van Genuchten; RMSD, root mean square deviation; SWIG, Soil Water Infiltration Global.

1 | INTRODUCTION AND SCOPE

Hydrologists use the term *infiltration* to describe the process by which water enters the soil and moves downward under the influence of gravity and/or capillary action (Philip, 1954). In his classic series of papers on the mathematical-physical description of infiltration, Philip (1957b, 1957c, 1957d, 1957e, 1957f) presented a quasi-analytic solution of the general flow equation for cumulative infiltration, I (L), into homogeneous soils at uniform initial moisture content, θ_i (L^3L^{-3}). For vertical infiltration, this approximate solution of Richards (1931) equation equals a finite and convergent power series of $p > 3$ expansion terms (Philip, 1955, 1957a)

$$I(t) = \beta_1 t^{1/2} + \beta_2 t + \beta_3 t^{3/2} + \cdots + \beta_p t^{p/2} = \sum_{j=1}^{p} \beta_j t^{j/2}, \tag{1}$$

where t (T) denotes time and β_1 ($LT^{-1/2}$), β_2 (LT^{-1}), β_3 ($LT^{-3/2}$), and β_p ($LT^{-p/2}$) are soil-dependent constants that need to be determined by numerical methods. Philip (1969) showed that β_1 is synonymous to the sorptivity, S ($LT^{-1/2}$), a measure of the soil's capacity to take up and release liquids by capillarity, and β_2 is equal to a unitless multiple, c, of the saturated hydraulic conductivity, K_s (LT^{-1}), which measures a soil's ability to transmit water under the influence of gravity. The constant c is equal to 1/2, 2/3, and 0.38 depending upon the respective diffusivity model (i.e., linear, δ-function, and/or nonlinear) (Kunze & Nielsen, 1982; Philip, 1969; Rahmati et al., 2022). The sorptivity is not an invariant soil property but has meaning only in relation to the soil's uniform initial, θ_i, and final, θ_0 (L^3L^{-3}), moisture contents. Thus, we should write $S(\theta_i, \theta_0)$ but conveniently omit the two arguments of S in the remainder of this paper. Note that $K_s = K(\theta_s)$, where $K(\theta)$ is the soil's unsaturated soil hydraulic conductivity function and θ_s (L^3L^{-3}) signifies its saturated volumetric moisture content. The general soil water flow equation is commonly credited to Lorenzo Richards (1904–1993) but was originally introduced by Lewis Richardson (1881–1953) within the context of atmospheric heat and mass transfer (Richardson, 1922) as pointed out by Knight and Raats (2016). Thus, we refer to the general flow equation as Richardson's equation.

The coefficients $\beta_1, \ldots, \beta_p$ cannot be freely chosen but must satisfy constraints to preserve the physical underpinning of Equation (1). Specifically, coefficients cannot be negative and, thus, $\beta_j > 0$ for all $j = (1, \ldots, p)$. Moreover, Philip (1957b) cites empirical evidence that $\beta_3, \ldots, \beta_p$ should honor the following more stringent condition

$$\frac{\beta_j}{S} > \left(\frac{\beta_2}{S}\right)^{j-1} \quad \forall j \in (3, \ldots, p) \tag{2}$$

Core Ideas

- Structural model errors of infiltration functions bias the values of S and K_s.
- Least squares estimates of K_s compare poorly to measured values.
- Epistemic uncertainty turns coefficients of infiltration functions into fitting parameters.
- The time validity of Philip's two-term equation is an elusive theoretical quantity.
- Treatment of epistemic errors enhances usefulness of plot-scale infiltration experiments for large-scale prediction.

so that the series approximation of Equation (1) converges for $t < (S/\beta_2)^2$. These constraints may guarantee the physical integrity of Philip's infiltration function but have far reaching implications. First of all, strict positivity of the coefficients $\beta_1, \ldots, \beta_p$ implies a monotonically increasing infiltration rate with time t. This conflicts with measured cumulative infiltration curves which usually portray a constant infiltration rate at late times. As a result, the series expansion (1) will have a limited time validity. In this context, Philip (1969) introduced the so-called characteristic time

$$t_{char} = \left[\frac{S}{K(\theta_0) - K(\theta_i)}\right]^2, \tag{3}$$

at which "*…the effect of gravity on the process can be expected to be as great as that of capillarity*" (p. 250). Philip (1969) postulated that $t \leq t_{char}$ is commensurate with the time range of useful convergence for Equation (1). Rahmati et al. (2022) provide an alternative formulation of the characteristic time which multiplies the right-hand side of Equation (3) with a soil type dependent function $F(\phi)$ where ϕ is the dimensionless shape parameter of the infiltration equation of Parlange et al. (1982). This new formulation increases the value of t_{char} by a factor of three, on average, with largest increments observed for fine-textured soils. A second problem is that the coefficient constraints of $\beta_3, \ldots, \beta_p$ demand knowledge of the soil sorptivity, S, the very variable that is subject to inference using the measured infiltration data. Simple numerical methods such as presented by Warrick (2003) and Sayah et al. (2016) provide a means to estimate the sorptivity S but require knowledge of soil hydraulic properties. This increases experimental demands and limits the practical applicability of Equation (1). Last, the constraints of Equation (2) turn curve fitting of the series approximation of Equation (1) into a box-constrained optimization problem (Bristow & Savage, 1987)

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\mathbf{b}^- \le \boldsymbol{\beta} \le \mathbf{b}^+} \frac{1}{2} F(\boldsymbol{\beta}), \tag{4}$$

with $p \times 1$ vectors $\mathbf{b}^-$ and $\mathbf{b}^+$ of lower and upper bounds of $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ and sum of squares objective function $F(\boldsymbol{\beta})$ of the n-vector of cumulative infiltration residuals, $\mathbf{e}(\boldsymbol{\beta}) = \left(e_1(\boldsymbol{\beta}), \ldots, e_n(\boldsymbol{\beta})\right)^\top$

$$F(\boldsymbol{\beta}) = \left\| \widetilde{\mathbf{I}}_n - \mathbf{D}\boldsymbol{\beta} \right\|_2^2 = \mathbf{e}(\boldsymbol{\beta})^\top \mathbf{e}(\boldsymbol{\beta}), \tag{5}$$

where $\mathbf{D}$ is the $n \times p$ design matrix whose column-wise entries are the basis functions of Equation (1), $\widetilde{\mathbf{I}}_n = (\tilde{I}_1, \ldots, \tilde{I}_n)^\top$, is the $n \times 1$ vector of cumulative infiltration measurements at times $\tilde{t}_1, \ldots, \tilde{t}_n$ and the $\| \cdot \|_2$ operator returns the Euclidean or 2-norm of the n-vector of cumulative infiltration residuals, $\mathbf{e}(\boldsymbol{\beta})$. Unfortunately, methods such as bounded-variable least squares (Lawson & Hanson, 1995; Stark & Parker, 1995) are not equipped to handle the coefficient-dependent lower bounds, $b_3^-, \ldots, b_p^-$, of $\beta_3, \ldots, \beta_p$ in pursuit of the least squares coefficients $\widehat{\boldsymbol{\beta}}$ and, thus, minimum of $F(\boldsymbol{\beta})$. Nonlinear least squares with the Levenberg–Marquardt algorithm (Levenberg, 1944; Marquardt, 1963) is to no avail as the coefficient-dependent lower bounds introduce discontinuities in $F(\boldsymbol{\beta})$ which frustrate an elegant iterative solution for $\widehat{\boldsymbol{\beta}}$. As a side note, Equation (5) makes the critical but convenient assumption that the cumulative infiltration residuals $\mathbf{e}(\boldsymbol{\beta})$ are independent with a zero mean and constant variance. Epistemic errors due to a structurally deficient infiltration function are assumed to be inconsequential and absorbed into the residuals.

The use of a large number of expansion terms p encourages Philip's series approximation of the general flow equation to converge to the exact infiltration solution assuming adequate knowledge of the values of the β coefficients. Thus, intuitively, one would expect a larger number of expansion terms p to enlarge the time interval over which Philip's series expression is valid (Rahmati et al., 2019), but possibly at the potential risk of overfitting. This is not uncommon for polynomial functions and is also known as Runge's phenomenon. The most popular and perhaps robust variant of Equation (1) retains only the first two terms of the power series (Philip, 1957e)

$$I(t) = St^{1/2} + cK_s t, \tag{6}$$

with strictly positive coefficients $\beta_1 = S$ and $\beta_2 = cK_s$. This two-parameter variant satisfies the plea of Philip (1957e) who argues (p. 257) that many situations in applied hydrology "*…require that the dynamics of infiltration be characterized by a small number of parameters*" and reduces box-constrained least squares estimation to a simple iterative solution with non-negativity constraints, $\beta_1, \beta_2 > 0$. The two-term expression is very simple and based on physical theory (Philip, 1969; Hunt et al., 2017) but the elimination of the higher-order terms of the series expansion is expected to reduce the time validity of Equation (6) "*…to all but very large t*" (Philip, 1957e).

This limitation has been well recognized in the infiltration literature of the 1970s and 1980s, for example, in the publications of Talsma and Parlange (1972), Kunze and Nielsen (1982), and Haverkamp et al. (1988) but these contributions are not always used and/or cited in cavalier applications of Philip's two-term expression. The hardly precise language "*…to all but very large t*" of Philip (1957e) may have contributed to the misuse of Equation (6) to all infiltration data, possibly excluding cumulative infiltration measurements taken at $t > t_{\text{char}}$ and/or relatively late times. This approach will inevitably corrupt the estimates of S and the product, cK_s, in curve fitting. The extent to which cumulative infiltration measurements, $\tilde{I}(t)$, beyond the valid time interval, t_{valid}, of Philip's two-term expression bias the estimates of the soil sorptivity and saturated hydraulic conductivity has not been investigated in the vadose zone literature in the absence of detailed guidelines and/or knowledge on the time validity of Equation (6). The closest to our work is the study of Haverkamp et al. (1988) who analyzed the time dependence of the coefficients of Philip's two-term expression and algebraic infiltration functions of Green and Ampt (1911), Kostiakov (1932), Horton (1941), Mezencev (1948), and Parlange et al. (1982) by gradually increasing the length of the infiltration data record starting at very short times. The authors concluded that the coefficients of all infiltration functions but that of Parlange et al. (1982) violate the constancy assumption and, thus, should be considered fitting parameters without physical significance.

In Jaiswal et al. (2022), we have introduced a new method coined parasite inversion for determining S, c, K_s, and t_{valid} of Philip's two-term infiltration Equation (6). The power and usefulness of this approach was illustrated by application to HYDRUS-1D (Šimůnek et al., 2008) simulated infiltration data using the soil hydraulic functions of Mualem (1976) and van Genuchten (1980) with parameter values of Carsel and Parrish (1988) and uniform initial moisture contents documented in Table 1 for all 12 soils of the textural triangle. The last four columns report the results of Jaiswal et al. (2022) and list the characteristic time t_{char} of Philip (1969) using Equation (3) with $K(\theta_i) = 0$, the least squares value of the dimensionless coefficient, c, and corresponding estimates of the time validity, t_{valid}, and infiltration validity, $I_{\text{valid}} = I(t_{\text{valid}})$, of Philip's two-term equation.

Note that the characteristic time of each soil type exceeds (by far) its respective time validity. Thus, past studies that relied on t_{char} as guideline of the valid time interval of Philip's two-term infiltration equation may have yielded biased

TABLE 1 HYDRUS-1D simulated infiltration experiments: Initial moisture content, θ_i, hydraulic parameter values, θ_s, θ_r, α_{VG}, n_{VG}, and K_s, of the constitutive equations of Mualem (1976); van Genuchten (1980) and corresponding sorptivity estimate, $\widetilde{S}$, for each soil type of the textural triangle. The last four columns pertain to Philip's two-term infiltration Equation (6) and list the characteristic time, t_{char}, of Equation (3), the multiplicative coefficient, c, and associated values of the time validity, t_{valid}, and infiltration validity, I_{valid}, at $t = t_{valid}$. The sorptivity was computed numerically using HYDRUS-1D simulated soil moisture profiles for a horizontal infiltration experiment with initial moisture content, θ_i. We treat S and K_s as measured quantities, thus, supply them with a tilde symbol.

Soil type	MVG soil hydraulic parameters							Philip 2-term equation			
	θ_i	θ_s	θ_r	α_{VG}	n_{VG}	$\widetilde{K}_s$	$\widetilde{S}$	t_{char}	$\hat{c}$	t_{valid}	I_{valid}
	cm^3cm^{-3}			cm^{-1}		$cm\ h^{-1}$	$cm\ h^{-1/2}$	h		h	cm
Clay	0.271	0.380	0.068	0.008	1.09	0.200	1.02	23.75	0.210	4.909	2.550
Clay loam	0.150	0.410	0.095	0.019	1.31	0.260	1.45	23.62	0.235	3.799	3.200
Loam	0.088	0.430	0.078	0.036	1.56	1.040	2.19	5.471	0.245	1.169	2.750
Loamy sand	0.057	0.410	0.057	0.124	2.28	14.59	6.20	0.171	0.458	0.099	2.700
Sand	0.045	0.430	0.045	0.145	2.68	29.70	9.21	0.086	0.498	0.055	3.100
Sandy clay	0.170	0.380	0.100	0.027	1.23	0.120	0.78	34.46	0.282	7.800	2.550
Sandy clay loam	0.111	0.390	0.100	0.059	1.48	1.310	1.60	1.818	0.207	0.795	1.750
Sandy loam	0.066	0.410	0.065	0.075	1.89	4.421	3.83	0.773	0.395	0.369	3.050
Silt	0.090	0.460	0.034	0.016	1.37	0.250	1.34	45.45	0.157	5.541	3.450
Silt loam	0.104	0.450	0.067	0.020	1.41	0.450	1.65	20.20	0.181	2.949	3.150
Silt clay	0.266	0.360	0.070	0.005	1.09	0.020	0.35	288.2	0.192	54.45	2.850
Silty clay loam	0.197	0.430	0.089	0.010	1.23	0.070	0.52	79.80	0.178	31.43	3.250

estimates of S and K_s. Large differences are observed in the values of t_{valid} among the different soils. Instead, it is more convenient to talk about the infiltration validity of Equation (6). The cumulative infiltration values at $t = t_{valid}$ are in close agreement and range between 2.5 and 3.2 cm, with the exception of 1.8 cm for sandy clay loam which was identified as dissonant soil by Jaiswal et al. (2022). Based on the tabulated findings, Jaiswal et al. (2022) concluded that 2.5 cm of cumulative infiltration serves as a simple proxy for the time validity of Equation (6). At this time, most of the soils have reached an approximately constant infiltration rate. This guideline of about 2.5 cm (or 1 in.) of cumulative infiltration guarantees an unbiased estimation of the soil sorptivity and saturated hydraulic conductivity for all soils but sandy clay loam. Indeed, soils with a value of I_{valid} larger than 2.5 cm (see Figure 1) are protected by a valid model formulation and, consequently, values of S and K_s derived from curve fitting should be bias free, if, of course, the $I(t)$ relationship for $0 < I \leq 2.5$ cm does not suffer systematic measurement errors and the infiltration process is exactly described by Richardson's equation.

In this paper, we revisit HYDRUS-1D simulated infiltration data and the measured infiltration experiments of the Soil Water Infiltration Global (SWIG) database of Rahmati et al. (2018) and compare and contrast least squares estimates of S and K_s of Philip's two-term infiltration equation derived from current practice using all cumulative infiltration data and their values obtained separately using data only up to the infiltration validity of 2.5 cm. The simulated infiltration data allows us to benchmark the inferred values of S and K_s against their known values. Section 2 describes briefly the data and our methodology. This is followed by Section 3 with a presentation and discussion of our results. Here, we are especially concerned with the role of epistemic errors on the least squares estimates of S and K_s, simulated infiltration curves, and their confidence intervals. This is followed by Section 4 which discusses the implications of our results. Finally, Section 5 concludes this paper with a summary of our main findings.

Before moving on to the materials and methods we state first that some of our findings on the physical significance of the infiltration parameters are not new and have been reported in earlier publications, most notably the previously cited work of Haverkamp et al. (1988). These authors used the time dependence of the coefficients of the algebraic infiltration equations of Kostiakov (1932), Horton (1941), Mezencev (1948), Green and Ampt (1911), and Philip (1957e) to warn readers about their lack of physical significance. Yet, Haverkamp et al. (1988) assumed each of the infiltration equations to be valid for the entire duration of the experiment and did not recognize the role of epistemic errors in controlling the time validity and physical underpinning of the coefficients. Furthermore, Haverkamp et al. (1988) analyzed only two infiltration experiments involving numerical and experimental data from a clay and sandy soil, respectively (see their Table 2, p. 321). By definition, the clay soil will always favor the infiltration equation of Parlange et al. (1982) as this is a quasi-exact implicit solution of the very flow equation used

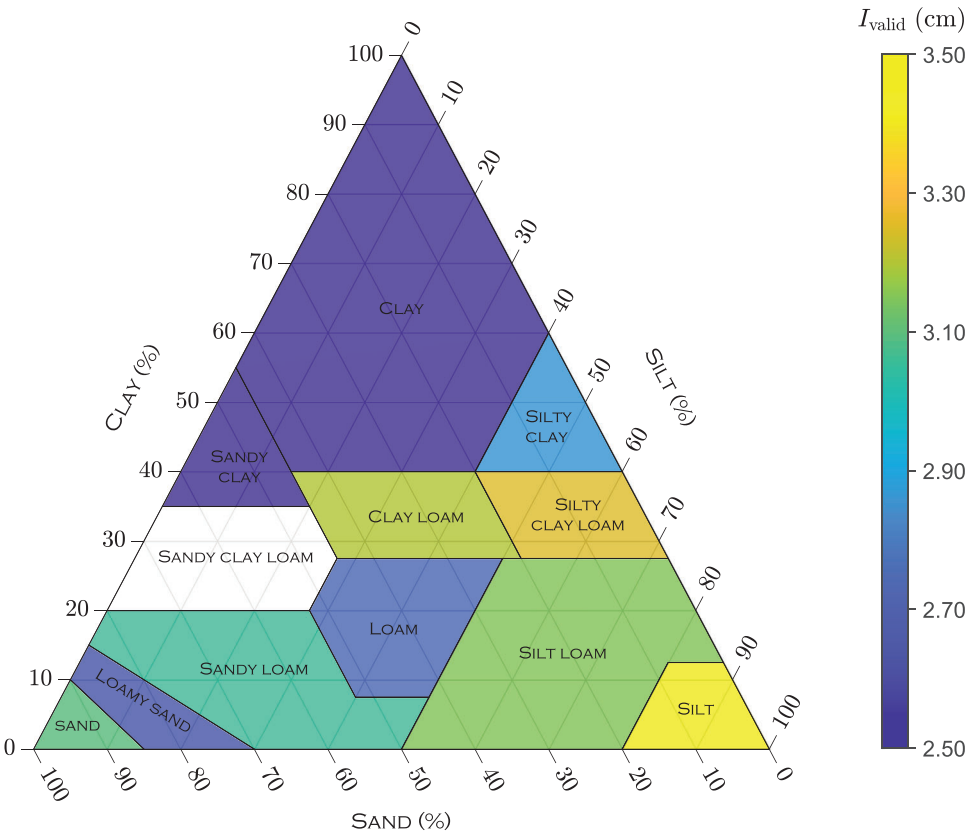

FIGURE 1 Projection of the cumulative infiltration validity of Philip's two-term equation onto the 12 soil types of the textural triangle. The dissonant soil, sandy clay loam, is intentionally left blank so as to not suppress subtle variations in I_{valid}. The colorbar assigns values (in cm) to the infiltration validity, I_{valid}, at $t = t_{valid}$.

to simulate the data. But as our results suggests, the parameters of this infiltration equation are not exempt from a loss of physical significance if the infiltration process does not satisfy Richardson's equation.

2 | MATERIALS AND METHODS

In this section, we briefly discuss the experimental data and methods used to determine the least squares values of the sorptivity and saturated soil hydraulic conductivity.

2.1 | Materials

2.1.1 | Synthetic infiltration data

Vertical infiltration into a 200 cm deep homogeneous soil was simulated with HYDRUS-1D (Šimůnek et al., 2008, 2016) for the 12 different soil types of the textural triangle using the Mualem–van Genuchten (MVG) hydraulic functions with parameter values and initial moisture content, θ_i, listed in Table 1. For each soil, the initial hydraulic head of the column was set equal to -15,000 cm (except for sand we used -1000 cm). Then cumulative infiltration was simulated for a period of 240 h using a constant pressure head at the surface and free drainage condition at the bottom of the soil profile (see also Rahmati et al., 2020). The raw output of HYDRUS-1D was post-processed to yield constant infiltration increments of 0.05 cm between successive measurement times up to a maximum of 5 cm of cumulative infiltration. The $n = 100$ measurement times of the $I(t)$ relationship were derived from linear interpolation of the raw HYDRUS-1D output. Thus, the measurement times, $\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_{100}$ correspond to $\tilde{I}_1 = 0.05, \tilde{I}_2 = 0.10, \ldots, \tilde{I}_{100} = 5.00$ cm. The resulting data set is discussed at length in Vrugt and Gao (2022), and interested readers are referred to this publication for further details. What suffices to say is that the simulated infiltration curves differ substantially among the different soil types. Sand, for example, needs only a handful of minutes to infiltrate 5 cm of water, whereas silty clay requires about 150 h to do so.

2.1.2 | Measured infiltration data

We also evaluate our methodology using measured infiltration experiments from the SWIG database of Rahmati et al. (2018). This database documents more than 5000 infiltration data sets

involving experiments on all 12 soil types (listed in Table 1) of the texture triangle. These experiments were conducted with a variety of different measurement devices and experimental techniques including infiltrometers, permeameters, and rainfall simulators. We only consider experiments with a double ring infiltrometer (coded with instrument 1 in the SWIG database) as the water flow underneath the inner ring most closely approximates the assumptions of vertical water flow of Philip's infiltration equation. The final collection is comprised of 646 data sets and includes a few experiments with suspicious data and/or units as detailed in Vrugt and Gao (2022). Of this collection, a total of 161 infiltration experiments provide measured values of the saturated hydraulic conductivity, $\widetilde{K}_s$. We cannot appraise the listed values, hence, simply use them as is in the present paper.

2.1.3 | Reference infiltration data

The series expansion of Philip (1955) and Philip (1957a) in Equation (1) assumes that vertical infiltration is accurately described with Richardson's equation. But what if this critical assumption does not hold in practice and the inference of S and K_s is subject to structural error? Epistemic errors shake the physical–mathematical foundation of our analysis and will violate the statistical underpinning of the infiltration validities of Philip's two-term Equation (6) presented in Jaiswal et al. (2022) and listed in the last column of Table 1.

To evaluate the impact of epistemic errors on the least squares values of S and K_s derived from the full and partial experiments, respectively, we create a third data set using the infiltration equation of Horton (1941)

$$I(t) = i_f t + (i_0 - i_f)\left[1 - \exp(-\gamma t)\right], \tag{7}$$

where $i_0 > 0$ and $i_f > 0$ (cm h^{-1}) signify the initial and final (or equilibrium) infiltration capacity, respectively, and $\gamma > 0$ (h^{-1}) is a soil-dependent coefficient which governs the rate of decline of the infiltration capacity with time. This equation performs well for field and laboratory experiments (Mishra et al., 2003). We use nonlinear least squares with the Levenberg–Marquardt (LM) algorithm to determine the optimal values of the coefficients, i_f, i_0, and γ for our collection of 646 experiments of the SWIG database. Then, for each SWIG experiment, we simulate the least squares $I(t)$ relationship of Horton's equation with constant increments of $\tilde{I}_{max}/n$ until the maximum measured cumulative infiltration, $\tilde{I}_{max}$, is reached after $n = 100$ successive print times. In our third data set, the hundred (t, I) data pairs of Horton's equation replace the measured data of the SWIG experiments. In the remainder of this paper, this collection of least squares simulated infiltration curves is also referred to as Horton or reference infiltration data set. Certainly, as Philip's two-term expression differs in function, coefficient definition, and dimensionality from the data generating infiltration process of Equation (7), the inference of S and K_s will be subject to model error. This situation is commonplace for measured infiltration data. Hence, this third data set will help confirm and generalize our findings for the SWIG database about the impact of structural errors on the estimates of S and K_s from Philip's two-term infiltration equation.

Note that we did not add measurement noise to the simulated cumulative infiltration data sets of HYDRUS-1D and Horton's equation. Such random errors do not affect our main conclusions, except may increase somewhat the uncertainty of the least squares coefficients, $\widehat{\boldsymbol{\beta}}$, and corresponding infiltration curve, $\widehat{\mathbf{I}}_n = (\hat{I}_1, \ldots, \hat{I}_n)^\top$.

2.2 | Methods

The two-term infiltration function of Philip in Equation (6) can be written as an inner product, $I(t) = \mathbf{d}(t)^\top \boldsymbol{\beta}$, of a 1×2 design vector, $\mathbf{d}(t)^\top = (t^{1/2}, ct)$, and the 2×1 vector, $\boldsymbol{\beta} = (\beta_1, \beta_2)^\top$, of coefficients. Table 1 documents the value of the multiplicative coefficient, c, for each soil type. While Talsma and Parlange (1972) has shown that $\frac{1}{3} < c < \frac{2}{3}$, its value must approach unity at large times to yield the correct limiting behavior. Calculations for a Yolo light clay soil by Kunze and Nielsen (1982) demonstrate that c changes with time and depends on θ_i and θ_0, nevertheless, its value is assumed constant over the valid time interval of Philip's two-term expression.

If we stack the transposed design vectors, $\mathbf{d}(\tilde{t})^\top$, of the n measurement times, $\tilde{t}_1, \ldots, \tilde{t}_n$, in the $n \times 2$ design matrix, $\mathbf{D}$, then least squares estimation of the Philip coefficients amounts to

$$\widehat{\boldsymbol{\beta}} = \begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix} = \arg\min_{\boldsymbol{\beta} \geq \mathbf{b}^-} \frac{1}{2} \left\| \widetilde{\mathbf{I}}_n - \mathbf{D}\boldsymbol{\beta} \right\|_2^2, \tag{8}$$

where $\widehat{\beta}_1 = S$ and $\widehat{\beta}_2 = K_s$ and non-negativity constraint, $\mathbf{b}^- = (0, 0)^\top$. The sum of squares residuals objective function stipulated above makes the critical assumption that the cumulative infiltration residuals are independent with a zero mean and constant variance, $\sigma_{\tilde{I}}^2$. Thus, the residuals are expected to behave exactly similar to the measurement errors of the cumulative infiltration measurements.

The above optimization problem is solved in MATLAB (The Mathworks, 2021) using the built-in `lsqnonneg` function. For each infiltration experiment, we execute this function two times. In the first trial we use all data, $0 < \tilde{I}(\tilde{t}) \leq \tilde{I}_f$, where $\tilde{I}_f$ signifies the final cumulative infiltration at the end of the experiment. The second trial only uses the $\tilde{I}(\tilde{t})$ data up to

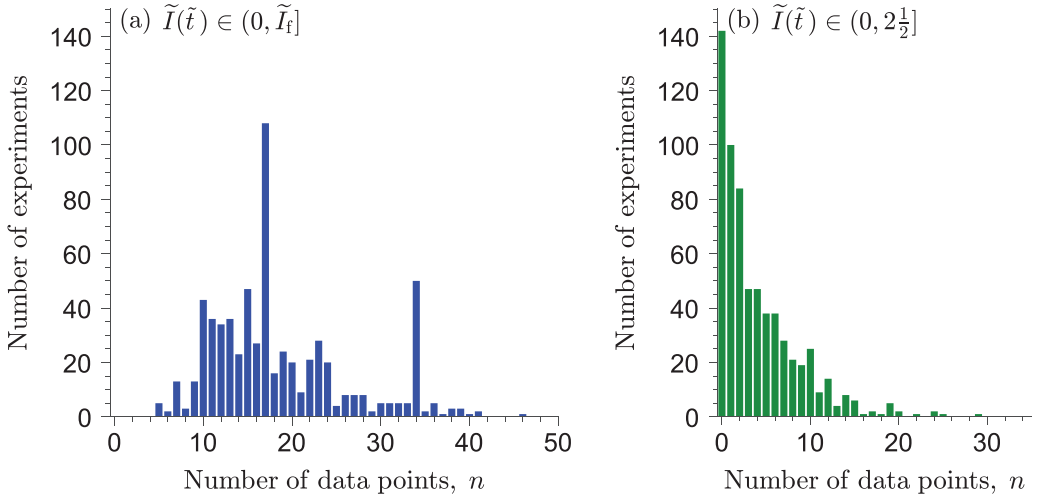

FIGURE 2 Soil Water Infiltration Global (SWIG) database: Frequency distribution of the number of cumulative infiltration measurements of the 646 experiments, (a) original data sets, and (b) data sets limited to the infiltration validity of 2.5 cm of Philip's two-term expression.

the infiltration validity, $I_{\text{valid}} = 2.5$ cm, of Philip's two-term equation. We use the wording full and partial experiments to differentiate between the data sets of the two trials. For the 12 HYDRUS-1D data sets, the cumulative infiltration increases with constant increments of 0.05 cm between successive measurement times until $\tilde{I}_f$ reaches 5 cm at the end of the simulation. Thus, the full and partial infiltration experiments of HYDRUS-1D will consist of $n = 100$ and $n = 50$ different $(\tilde{t}, \tilde{I})$ data pairs, respectively. The infiltration experiments of the SWIG database, on the contrary, do not have a common measurement protocol and therefore, we must verify that the number of $\tilde{I}(\tilde{t})$ data pairs is large enough to warrant joint estimation of the sorptivity and saturated soil hydraulic conductivity. Figure 2 presents histograms of the number, n, of $(\tilde{t}, \tilde{I})$ data pairs, of the (a) full and (b) partial experiments of the SWIG database. The color coding in blue and green for both data sets is used throughout the remainder of this paper. The number of $\tilde{I}(\tilde{t})$ measurements of the infiltration experiments of the SWIG database ranges between $n = 6$ and $n = 47$, and decreases to $n = 0$ and $n = 29$ when we limit the cumulative infiltration of each SWIG experiment to $I_{\text{valid}} = 2.5$ cm. There are $M = 226$ experiments with at least five $(\tilde{t}, \tilde{I})$-data pairs, the minimum data length (without origin) we deem necessary to provide adequate estimates of the sorptivity, S, and saturated soil hydraulic conductivity, K_s. The partial experiments cover all soil types except for sand, sandy clay, and silt (see Table 2 and Figure 3). For a fair comparison of the results, our reference infiltration data set of Horton's equation matches the final collection of 226 SWIG experiments.

Once the least squares coefficients, $\hat{\boldsymbol{\beta}}$, of Philip's two-term expression are known, the corresponding $n \times 1$ vector of simulated cumulative infiltration may be computed for the full and partial experiments through matrix-vector multiplication using, $\hat{\mathbf{I}}_n = \mathbf{D}\hat{\boldsymbol{\beta}}$. The so-obtained least squares infiltration curves are entirely deterministic and do not characterize simulation uncertainty. In doing so, we must first quantify the uncertainty of the least squares estimates, $\hat{\boldsymbol{\beta}}$, of the soil sorptivity and saturated hydraulic conductivity. This requires us to make an assumption about the distribution of the cumulative infiltration residuals. If we make the common and convenient assumption that the measurement errors a normally distributed, the 2×2 covariance matrix, $\mathbf{C}(\hat{\boldsymbol{\beta}})$, of the least squares parameter values of the two-term expression simplifies to

$$\mathbf{C}(\hat{\boldsymbol{\beta}}) = \mathbb{E}\left[(\hat{\boldsymbol{\beta}} - \mathbb{E}[\hat{\boldsymbol{\beta}}])(\hat{\boldsymbol{\beta}} - \mathbb{E}[\hat{\boldsymbol{\beta}}])^{\top}\right] = \hat{\sigma}_{\tilde{I}}^2 (\mathbf{D}^{\top}\mathbf{D})^{-1}, \quad (9)$$

TABLE 2 Soil Water Infiltration Global (SWIG) database: The number of infiltration experiments with at least five $\tilde{I}(\tilde{t})$ measurements in the range of $\tilde{I} \in (0, 2.5]$ cm for each soil type.

Soil	*n*
Clay	79
Clay loam	38
Loam	17
Loamy sand	4
Sand	0
Sandy clay	0
Sandy clay loam	8
Sandy loam	38
Silt	0
Silt loam	23
Silt clay	10
Silty clay loam	9
Total	226

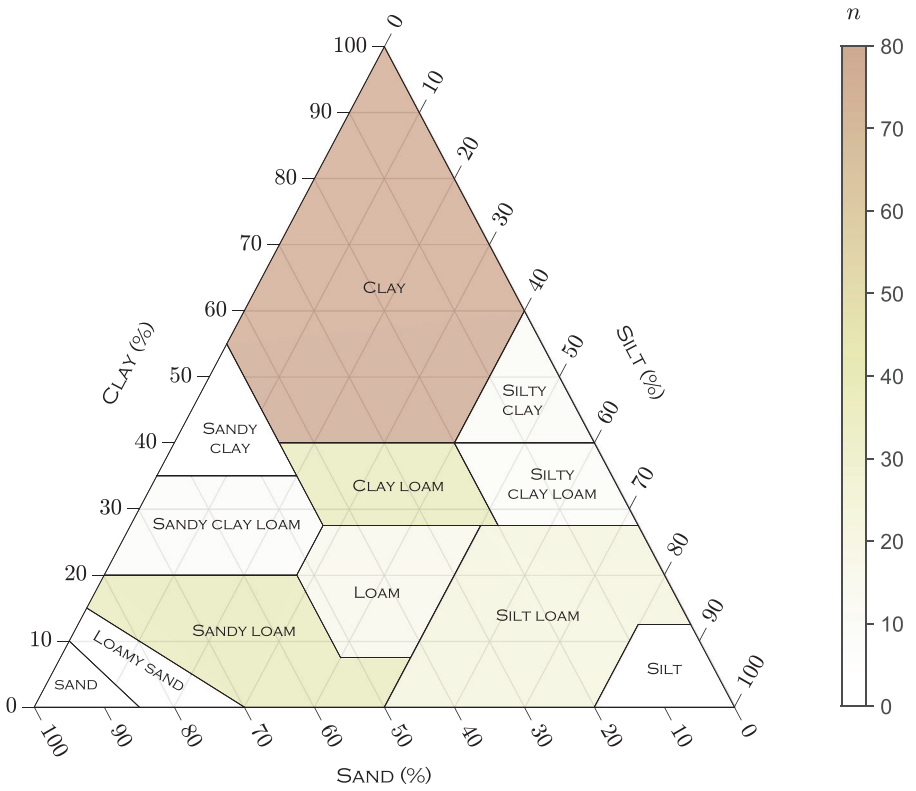

FIGURE 3 Projection of the tabulated results onto the textural triangle.

where

$$\hat{\sigma}_{\tilde{I}}^2 = \frac{\mathbf{e}(\hat{\boldsymbol{\beta}})^\top \mathbf{e}(\hat{\boldsymbol{\beta}})}{n-2}, \tag{10}$$

is our sample estimate of $\sigma_{\tilde{I}}^2$. The standard deviation, $\hat{\sigma}_{\tilde{I}}$, is the familiar root mean square deviation (RMSD) of the least squares fit. The $100(1-\alpha)\%$ confidence interval, $\beta_j \in [\hat{\beta}_j^-, \hat{\beta}_j^+]$, of the jth coefficient, $j = (1, 2)$, is now equal to

$$\hat{\beta}_j^{\pm} = \hat{\beta}_j \pm F_{\mathcal{T}}^{-1}(p_\alpha|0,1,\nu)\sqrt{c_{j,j}} \tag{11}$$

where $c_{j,j}$ signifies the (j,j)th element of the parameter covariance matrix, $\mathbf{C}(\hat{\boldsymbol{\beta}})$, $F_{\mathcal{T}}^{-1}(p_\alpha|0,1,\nu)$ is the inverse of the standardized Student's t-cumulative distribution function at percentile, $p_\alpha = \frac{1}{2} \pm (1-\alpha)/2$, and degrees of freedom, $\nu = n-2$, and $\alpha \in (0,1)$ denotes the significance level. For $\alpha = 0.05$ we yield a 95% confidence interval and the critical t-value, $F_{\mathcal{T}}^{-1}(p_\alpha|0,1,\nu)$, equals 12.71, 2.57, and 1.96 for $n = 1$, $n = 5$, and $n \to \infty$, respectively. Next, we can turn the parameter uncertainty into confidence intervals of the simulated infiltration curve. Specifically, the $100(1-\alpha)\%$ confidence limits, $\hat{I}_c^{\pm}(t) = [\hat{I}_c^-(t), \hat{I}_c^+(t)]$, of the least squares infiltration, $\hat{I}(t)$, simulated by Philip's two-term expression can be computed as follows

$$\hat{I}_c^{\pm}(t) = \hat{I}(t) \pm F_{\mathcal{T}}^{-1}(p_\alpha|0,1,\nu)\sqrt{\mathbf{d}(t)^\top \mathbf{C}(\hat{\boldsymbol{\beta}})\mathbf{d}(t)}, \tag{12}$$

for all t. These confidence limits summarize the effect of coefficient uncertainty on the model output and make up the $100(1-\alpha)\%$ confidence interval of the simulated infiltration curve. For any time, t, the confidence interval follows a Gaussian distribution with mean equal to the simulated cumulative infiltration, $\hat{I}(t) = \mathbf{d}(t)^\top \hat{\boldsymbol{\beta}}$, of $\hat{\boldsymbol{\beta}}$ and variance determined by $\mathbf{d}(t)$ and the parameter covariance matrix, $\mathbf{C}(\hat{\boldsymbol{\beta}})$, of the least squares coefficients. Note that the element of the square root in Equation (12) is equal to the variance of the simulated cumulative infiltration at time t.

To inspire confidence in the optimal coefficients, $\hat{\boldsymbol{\beta}}$, we must evaluate whether the cumulative infiltration residuals satisfy the assumptions of homogeneity and independence of the least squares estimator. We are particularly concerned about residual autocorrelation, a common byproduct of systematic errors. We use the unitless Durbin and Watson (1950), Durbin and Watson (1951) test statistic

$$d_w = \frac{\sum_{i=2}^{n}\left[e_i(\hat{\boldsymbol{\beta}}) - e_{i-1}(\hat{\boldsymbol{\beta}})\right]^2}{\sum_{i=1}^{n} e_i(\hat{\boldsymbol{\beta}})^2} \approx 2(1-\hat{\rho}), \tag{13}$$

to verify whether the residuals, $\mathbf{e}(\widehat{\boldsymbol{\beta}})$, of the least squares coefficients, $\widehat{\boldsymbol{\beta}}$, satisfy the null hypothesis, $\rho = 0$, against the alternative hypothesis, $\rho \neq 0$, that there is autocorrelation among them. The d_w statistic varies between 0 and 4. A value of $d_w \in (0, 2)$ suggests that the successive infiltration residuals are positively correlated, whereas a value of $d_w \in (2, 4)$ is indicative of negative autocorrelation. The null hypothesis is rejected if $d_w < d_w^-(\alpha|n, p)$ or $(4 - d_w) < d_w^-(\alpha|n, p)$, where $d_w^-(\alpha|n, p)$ signifies the critical value at α significance level and n and $p = 2$ degrees of freedom. Farebrother (1980) present tables of the critical values for sample sizes ranging from $n = 2$ to $n = 200$ data points and for $p = 0$ to $p = 21$ coefficients in the class of regression models without an intercept. For $\alpha = 0.05$, $n = 10$, and $p = 2$, the value of $d_w^-(\alpha|n, p) = 0.697$. As a simple rule of thumb, a value of $d_w < 1$ or $d_w > 3$ is cause for alarm, that is, there is sufficient statistical evidence for positive or negative residual autocorrelation, respectively, and the null hypothesis is rejected at $100(1 - \alpha)\%$ confidence interval.

In Appendix A we present a MATLAB subroutine called `fit_Philip2` which implements the above methods for Philip's two-term infiltration equation and returns the desired output. For the HYDRUS-1D data sets, the values of the soil sorptivity and saturated hydraulic conductivity are known beforehand. We use the RMSD to quantify the overall distance between the least squares estimates, $\widehat{S}$ and $\widehat{K}_s$, and the "*true*" values of the sorptivity, $\widetilde{S}$, and saturated soil hydraulic conductivity, $\widetilde{K}_s$, respectively. For the sorptivity we can write

$$\text{RMSD} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(\widetilde{S}_i - \widehat{S}_i\right)^2}, \tag{14}$$

where $m = 12$ is the number of soil types.

3 | RESULTS AND DISCUSSION

This section presents the results of our analysis. We first focus our attention on the HYDRUS-1D data set, then followed by the SWIG database, and present last our findings for the reference data set. In the text and figures we assign the subscripts 1 and 2 to the least squares values of the the full and partial experiments, respectively.

Figure 4 presents scatter plots of the measured sorptivity, $\widetilde{S}$, and saturated hydraulic conductivity, $\widetilde{K}_s$, against the least squares values of S and K_s derived from (a1, a2) current practice using all infiltration measurements and (b1, b2) those obtained from using cumulative infiltration up to $I_{\text{valid}} = 2.5$ cm only.

The least squares values of the sorptivity and saturated hydraulic conductivity derived from the HYDRUS-1D simulated infiltration data correlate well with their measured counterparts. The use of all measured $(\tilde{t}, \tilde{I})$ data pairs for least squares estimation of S and K_s tends to underestimate somewhat the "*measured*" sorptivity, $\widetilde{S}$, and systematically overestimate the saturated hydraulic conductivity, $\widetilde{K}_s$. The underestimation of the sorptivity is most apparent for coarse-textured soils, sand and loamy sand, which have the largest "*measured*" sorptivities of 6.20 and 9.21 cm $h^{-1/2}$, respectively. The RMSD of the least squares values of the soil sorptivity and saturated hydraulic conductivity of the full experiments, $\widehat{S}_1$ and $\widehat{K}_{s,1}$, respectively, are on the order of 0.2 cm $h^{-1/2}$ and 3.2 cm h^{-1}, respectively. The use of the partial experiments with infiltration data limited to the time validity of 2.5 cm of Philip's two-term expression improves the compliance between the fitted sorptivities, $\widehat{S}_2$, and saturated soil hydraulic conductivities, $\widehat{K}_{s,2}$, and their true values from Table 1. The $(\widetilde{S}, \widehat{S}_2)$ and $(\widetilde{K}_s, \widehat{K}_{s,2})$ data points cluster more closely around the 1:1 line and exhibit substantially lower values of the RMSD than their counterparts of the full experiments. These results provide support for the claim that we should only use infiltration data up to the time (infiltration) validity in curve fitting of Philip's two-term expression. Cumulative infiltration measurements beyond 2.5 cm corrupt the estimates of S and K_s. The 95% confidence intervals of the least squares estimates of S and K_s are negligible, hence do not show in the scatter plots.

Next, we move on to the experiments of the SWIG database. Figure 5 compares the least squares values of (a) the sorptivity, $\widehat{S}_1$ and $\widehat{S}_2$, and (b) saturated soil hydraulic conductivity, $\widehat{K}_{s,1}$ and $\widehat{K}_{s,2}$, of the full and partial experiments, respectively. The associated 95% confidence intervals of S and K_s are separately indicated with the horizontal blue and vertical green lines.

The two scatter plots demonstrate considerable dispersion of the $(\widehat{S}_1, \widehat{S}_2)$ and $(\widehat{K}_{s,1}, \widehat{K}_{s,2})$ data points around the 1:1 line of the respective quantities. This confirms our earlier finding that the length of the infiltration experiment has a considerable impact on the inferred values of S and K_s. Notice that a relatively large number of $(\widehat{S}_1, \widehat{S}_2)$ data points appear under the 1:1 line. The use of $(\tilde{t}, \tilde{I})$ data pairs beyond the infiltration validity of Philip's two-term equation tends to overestimate the sorptivity, S. The opposite is true for the saturated soil hydraulic conductivity, K_s. A large proportion of the $(\widehat{K}_{s,1}, \widehat{K}_{s,2})$ data points fall above the 1:1: line on the right-hand side graph. Thus, measurements of the $I(t)$ relationship that extent beyond the infiltration validity of Philip's two-term equation will underestimate the K_s value. The confidence intervals of the coefficients, S and K_s, are largest, on average, for the partial experiments with green lines that often extent beyond their blue counterparts. This is particularly true for the soil sorptivity, $\widehat{S}_2$, and a result of using a relatively small data set of $(\tilde{t}, \tilde{I})$ measurements for coefficient

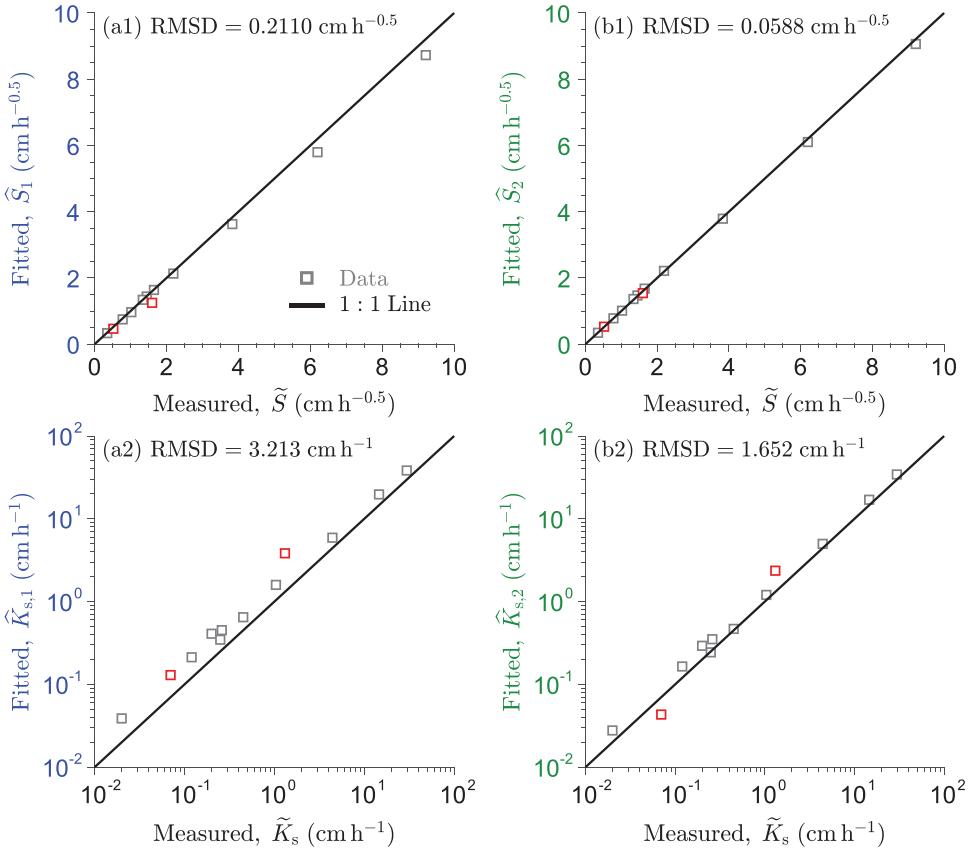

FIGURE 4 HYDRUS-1D infiltration data set: Scatter plots of the observed and estimated values of the (a1, b1) soil sorptivity, S, in cm h$^{-1/2}$ and (a2, b2) saturated soil hydraulic conductivity, K_s, in units of cm h^{-1} using (a1, a2) current practice with all infiltration data and (a2, b2) infiltration data limited to 2.5 cm. The solid black line characterizes the 1:1 relationship between the measured and estimated quantities. The two red squares correspond to sandy clay loam and silty clay loam, the two dissonant soils identified by Jaiswal et al. (2022).

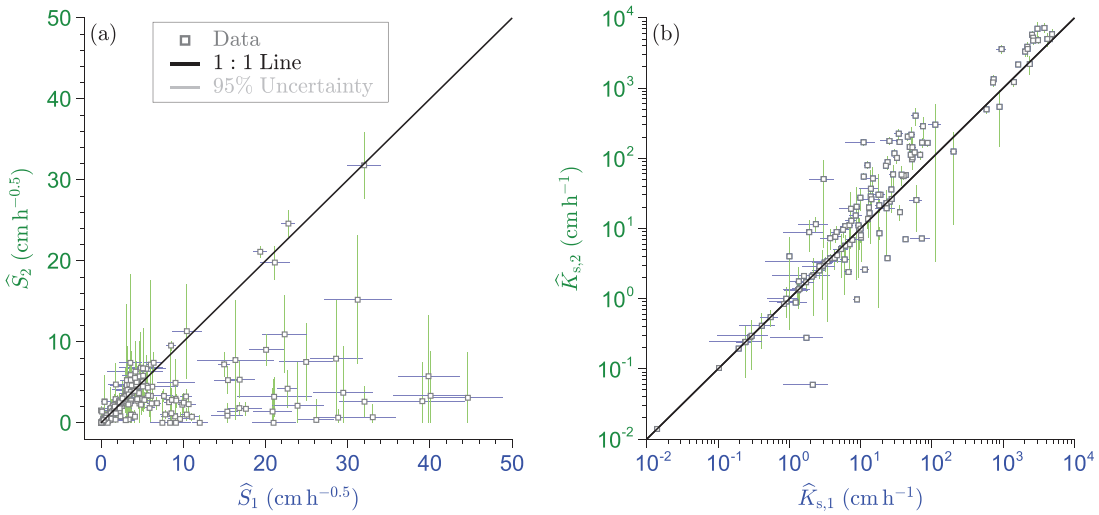

FIGURE 5 Soil Water Infiltration Global (database: Scatter plots of the (a) soil sorptivity, S, and (b) saturated soil hydraulic conductivity, K_s, derived from current practice, $\widehat{S}_1$ and $\widehat{K}_{s,1}$ (on x-axes), using all infiltration data against their counterparts, $\widehat{S}_2$ and $\widehat{K}_{s,2}$ (on y-axes) for the partial experiments up to the infiltration validity of 2.5 cm of Philip's two-term equation. The solid black line characterizes the 1:1 relationship between the plotted quantities. The horizontal blue and green vertical lines quantify the width of the 95% confidence intervals of the least squares coefficients.

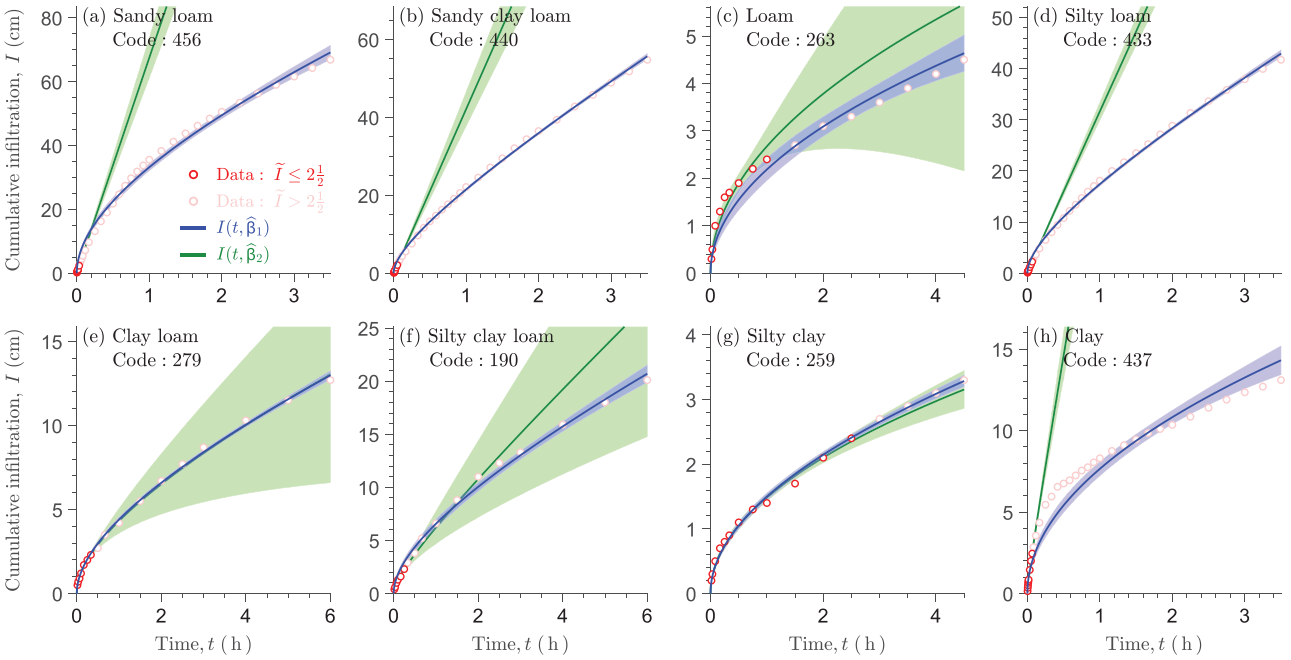

FIGURE 6 Soil Water Infiltration Global (SWIG) database: Comparison of observed (red dots) and simulated cumulative infiltration curves for a selection of eight representative soils using the least squares values of S, K_s, and c (Table 1) of Philip's two-term infiltration equation. Each graph corresponds to a different soil type, specifically (a) sandy loam, (b) sandy clay loam, (c) loam, (d) silty loam, (e) clay loam, (f) silty clay loam, (g) silty clay and (h) clay. Measured cumulative infiltration data beyond the time validity of Philip's two-term equation are displayed with a tint of red. The light blue and green colored regions are the 95% confidence intervals of the simulated infiltration curves. The bottom right corner of each graph documents the SWIG code of each experiment.

estimation. Note that a consequence of the logarithmic scaling of the axes of the $(\widehat{K}_{s,1}, \widehat{K}_{s,2})$ scatter plot is that the 95% uncertainty is not symmetric around each data point.

Figure 6 compares measured (red dots) and simulated cumulative infiltration curves derived from current practice using all infiltration data (solid blue line) of the experiments of the SWIG database and data up to the infiltration validity of Philip's two-term equation (solid green line). Measured data beyond the infiltration validity, 2.5 cm, of Philip's two-term equation, are separately displayed in light red. The light colored regions portray the 95% confidence intervals of the simulated cumulative infiltration curves. The least squares values of the sorptivity and saturated soil hydraulic conductivity, $\widehat{\beta}_1 = (\widehat{S}_1, \widehat{K}_{s,1})^\top$, derived from current practice (solid blue lines) provide an excellent match of Philip's two-term equation to the measured infiltration curves of the different soil types. As all $(\tilde{t}, \tilde{I})$ measurements are used for coefficient estimation, Philip's two-term equation is in close agreement with the measured data and the 95% confidence intervals (blue region) appear tight and center on the least squares infiltration curves. This demonstrates that the coefficients of Philip's two-term infiltration equation are well defined by calibration against the observed $\tilde{I}(\tilde{t})$ relationship.

The least squares values of the coefficients, $\widehat{\beta}_2 = (\widehat{S}_2, \widehat{K}_{s,2})^\top$, obtained from the partial experiments describe well the measured $I(t)$ relationship at early infiltration times (red dots) but do not match the infiltration data (light red dots) at larger measurement times beyond the infiltration validity of 2.5 cm of Philip's two-term equation. Notable exceptions are (e) clay loam and (g) silty clay for which the solid blue and green lines almost coalesce into a single curve. This close correspondence between $\widehat{\beta}_1$ and $\widehat{\beta}_2$ is expected for silty clay as the full and partial experiments only differ in a few $(\tilde{t}, \tilde{I})$ data points. The confidence intervals of the partial experiment (medium green) are narrow at early measurement times but grow rapidly for (c) loam, (e) clay loam, and (f) silty clay loam after the infiltration validity of Philip's two-term equation has been exceeded. These large confidence limits simply articulate a large uncertainty in the saturated hydraulic conductivity, $\widehat{K}_{s,2}$, derived from the early infiltration data, $0 < \tilde{I}(\tilde{t}) \leq 2.5$ cm displayed with the red dots. The small sample size, n, for a large majority of the partial experiments will certainly contribute to the uncertainty of $\widehat{K}_{s,2}$, and for that matter, $\widehat{S}_2$.

The apparent mismatch between the solid blue and green lines at later infiltration times does not come as a surprise. As Philip's two-term equation is deficient in describing the cumulative infiltration at late times, the full and partial experiments may return different estimates of the soil sorptivity, S, and saturated hydraulic conductivity, K_s. As we are looking at the cumulative infiltration, the divergence of the blue and green lines is expected to increase with time, t. For (a) sandy loam, (b) sandy clay loam, (d) silty loam, and (h)

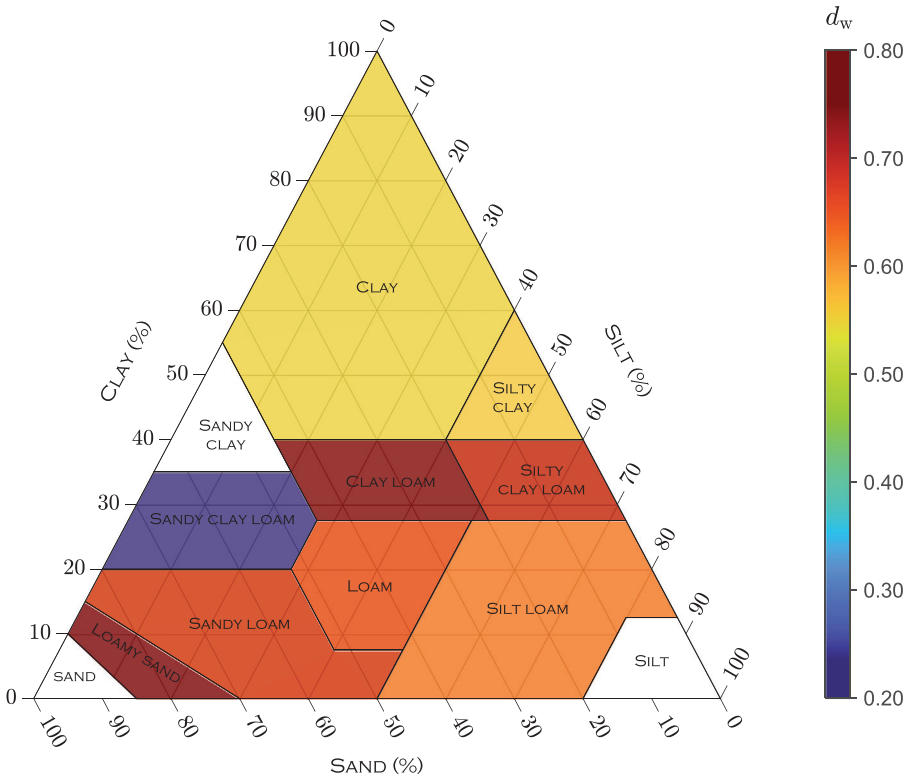

FIGURE 7 Soil Water Infiltration Global (SWIG) database: Graphical projection of the Durbin and Watson (1950), Durbin and Watson (1951) statistic for Philip's two-term infiltration equation onto the 12 soil types of the textural triangle. Values correspond to the mean d_w statistic for all the full SWIG experiments of each soil type. The colorbar assigns values to the test statistic. Sand, sandy clay and silt are intentionally left blank in the absence of sufficiently long experimental records (see Table 2).

clay, however, the differences in the cumulative infiltration curves of the full and partial experiments are much larger than one would expect from an equivalent analysis of the synthetic infiltration data simulated with HYDRUS-1D (see Figure B.1).

Figure 7 projects the mean value of the Durbin and Watson (1950), Durbin and Watson (1951) test statistic, $d_w \in (0, 4)$, for each soil type of the SWIG database onto the textural triangle. The colorbar assigns values to d_w. The mean values of d_w for experiments with the same soil type are consistently smaller than unity. In almost all experiments of the SWIG database we find strong statistical evidence that the cumulative infiltration residuals are positively correlated. This violates our basic assumption of residual independence, and highlights the consequences of model structural errors. Epistemic uncertainty is persistent across all soil types. This introduces bias in the least squares estimates of S and K_s and is responsible for the rather poor agreement in the infiltration curves of the full and partial experiments.

To help understand the large differences in the simulated cumulative infiltration curves of the full and partial experiments for some of the soils, we turn our attention to the reference data set of Horton's equation. When we compare the soil sorptivity and saturated hydraulic conductivity of the full and partial experiments of this reference set (see Figure 8) a pattern of data scattering emerges familiar to the $(\hat{S}_1, \hat{S}_2)$ and $(\hat{K}_{s,1}, \hat{K}_{s,2})$ bubble graphs in Figure 5a,b for the SWIG database. The soil sorptivities of the full experiments, $\hat{S}_1$, almost always exceed their counterparts, $\hat{S}_2$, of the partial experiments. The opposite is true for the saturated hydraulic conductivity. Indeed, the values of $\hat{K}_{s,1}$ tend to underestimate their values, $\hat{K}_{s,2}$, obtained from the partial experiments. Table 7 of Haverkamp et al. (1988) demonstrates an exactly opposite behavior of S and K_s (but listed as transmissivity A). We cannot comment so as to why this is true in the absence of information about the fitting procedure and objective function used in Haverkamp et al. (1988). The 95% confidence intervals of S and K_s are considerably smaller than those depicted previously in Figure 5 for the measured infiltration data of the SWIG database. This is a result of the much larger number of $(\tilde{t}, \tilde{I})$ data points in the reference data set, and, thus, lower critical t-values in Equations (11) and (12). The data length and duration of the infiltration experiment will soon become evident in our comparison of measured and simulated infiltration curves.

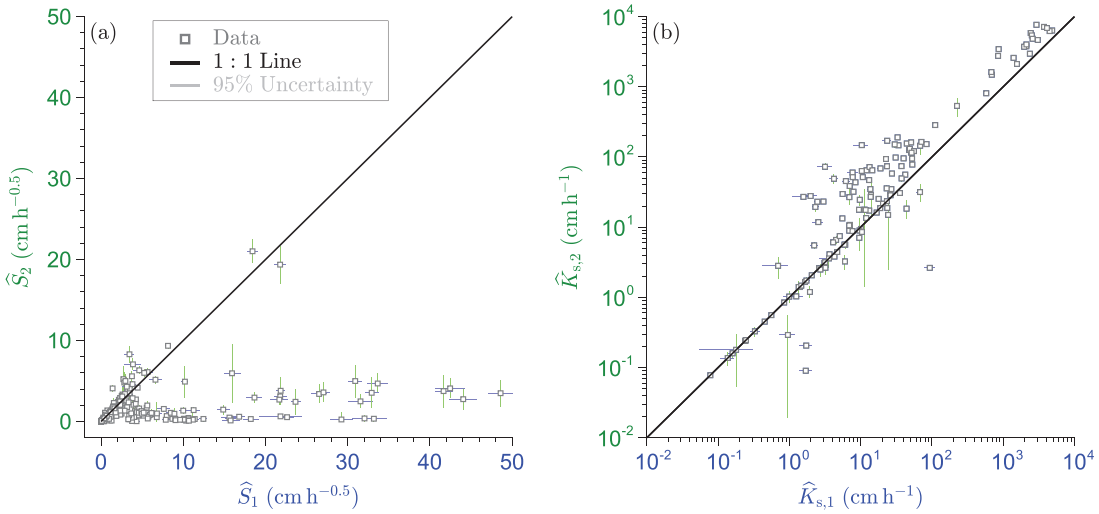

FIGURE 8 Horton infiltration data set: Scatter plots of the (a) soil sorptivity, S, and (b) saturated soil hydraulic conductivity, K_s, derived from the full experiment, $\widehat{S}_1$ and $\widehat{K}_{s,1}$ (on x-axes), against their counterparts, $\widehat{S}_2$ and $\widehat{K}_{s,2}$ (on y-axes) of the partial experiments up to the infiltration validity of 2.5 cm of Philip's two-term equation. The solid black line characterizes the 1:1 relationship between the plotted quantities. The horizontal blue and vertical green lines display the 95% confidence intervals of the two coefficients of Philip's two-term equation.

The striking resemblance of the $(\widehat{S}_1, \widehat{S}_2)$ and $(\widehat{K}_{s,1}, \widehat{K}_{s,2})$ scatter plots in Figures 5 and 8 highlights a fundamental weakness in our methodology. The assumption that model errors are inconsequential and absorbed into the cumulative infiltration residuals is convenient in applying statistical theory, but not borne out of the actual properties of the residuals which exhibit considerable serial correlation (shown next) and may deviate from normality. The culprit is structural error, that is, Philip's two-term equation does not describe perfectly the measured infiltration process. There may be many reasons for this. For example, our mathematical–physical description of the data generating may be deficient. But even if this description is accurate, our assumptions about the homogeneity or initial conditions of the soil may violate modeling assumptions. Indeed, (i) the soil may not be homogeneous, (ii) the initial moisture content, θ_i, may not be uniform, and/or (iii) the soil is hydrophobic and/or has a crust that resists infiltration.

The sobering reality is that the infiltration process will violate Richardson's equation and/or assumptions about soil homogeneity and uniformity of the initial moisture content. Theory dictates that Philip's two-term equation is deficient at late times, but this deficiency may be picayune in practice when confronted with structural model errors. Graphical eyeballing of fitted infiltration curves may lead us to conclude that model errors are small and inconsequential. Yet, almost surely will epistemic errors have impacted the estimates of S and K_s. Furthermore, the infiltration validities of Jaiswal et al. (2022) and documented in Table 1 are not exempt from structural errors. This is a disheartening prospect.

Unfortunately, epistemic errors do not admit a convenient mathematical–statistical description that may be exploited in the construction of a suitable objective function, $F(\Theta)$ (Gupta et al., 1998; Vrugt et al., 2005; Vrugt & Beven, 2018). Hence, structural errors complicate and frustrate tasks such as parameter estimation, uncertainty characterization, hypothesis testing, and scientific discovery (learning) with dynamical system models (Gupta et al., 2008; Vrugt & Sadegh, 2013). But even for simple polynomial functions such as Philip's infiltration Equation (1) it is well known that ignoring model structural errors, or wrongly specifying their spatiotemporal structure, will lead to bias in the parameter estimates (Beven, 2006). This discourages any attempts to seeking a physical underpinning of the regression coefficients and diminishes any hope that the least squares estimates of S and K_s represent innate soil properties.

Further evidence for the important role epistemic errors play in our analysis is provided in Figure 9 which compares measured and simulated infiltration curves of eight different soil types of Horton's reference data set. The soil codes (experiments) are equivalent to those used previously in Figure 6 for the measured data of the SWIG database. The results are in strong qualitative agreement with our findings for the SWIG database. For most of the soils, we observe a similar divergence of the simulated infiltration curves of the full and partial experiments. This divergence is much larger than anticipated from the HYDRUS-1D infiltration data set (see Figure B.1). This is a testament to epistemic error. The 95% confidence intervals of the simulated infiltration curves are substantially smaller than their counterparts of the measured infiltration experiments of the SWIG database displayed in Figure 6. This is particularly evident for the (c) loam, (e) clay loam, and (f) silty clay loam soils and a result of the much larger number of $(\tilde{t}, \tilde{I})$ data pairs of the reference infiltration set.

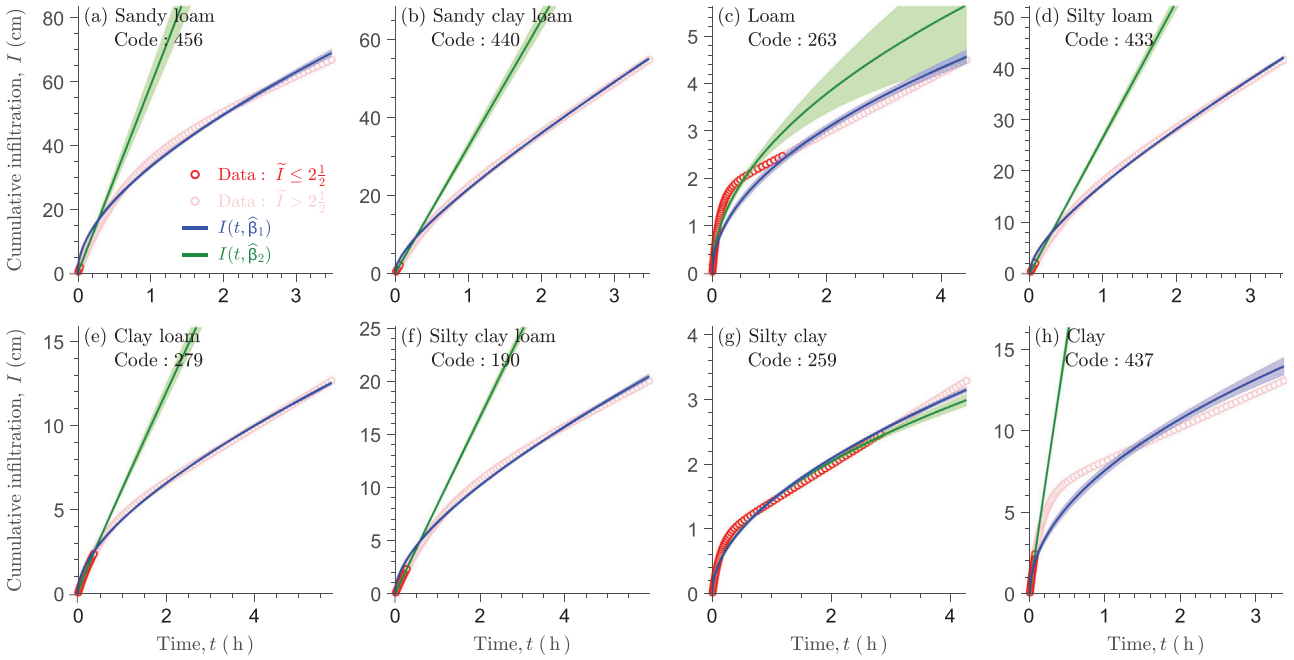

FIGURE 9 Horton infiltration data set: Observed (red dots) and simulated cumulative infiltration curves for the selection of eight Soil Water Infiltration Global (SWIG) soils using the least squares values of S, K_s, and c (Table 1) of Philip's two-term infiltration equation for the full (solid blue line) and partial (solid green line) experiments. Each graph corresponds to a different soil type, specifically (a) sandy loam, (b) sandy clay loam, (c) loam, (d) silty loam, (e) clay loam, (f) silty clay loam, (g) silty clay, and (h) clay. Measured cumulative infiltration data beyond the time validity of Philip's two-term equation are displayed with a tint of red. The light colored regions portray the 95% confidence intervals of the simulated infiltration curves.

The relatively long infiltration records of the reference infiltration data set help in illuminating the limitations of Philip's two-term equation. Structural errors are readily apparent in the simulated infiltration curves of the (c) loam, (g) silty clay, and (h) clay soils. The solid blue lines for these three soils do not pass well through the measured data (red circles). Specifically, Philip's two-term equation underestimates the cumulative infiltration at early times and overestimates the $(\tilde{t}, \tilde{I})$ data points at intermediate and/or late measurement times in an effort to minimize the 2-norm of the cumulative infiltration residuals. This alternating pattern of under- and overprediction is the result of epistemic errors and can be detected with residual diagnostics.

The quality of fit of Philip's two-term infiltration equation is much better for the other five soils, (a) sandy loam, (b) sandy clay loam, (d) silty loam, (e) clay loam, and (f) silty clay loam. Yet, upon closer inspection of the model–data mismatch, a similar, albeit visually less apparent, temporal pattern emerges of consecutive segments with over- and underestimation of the cumulative infiltration data. Although the simulated infiltration curves meander tightly around the measured data (red dots), there is little doubt that the residuals will violate the assumption of independence. Indeed, the d_w statistic is close to zero for all blue lines and near zero for most of the infiltration curves of the partial experiments (green lines). The structural errors will corrupt the estimates of S and K_s.

For all soils but (c) loam and (g) silty clay, the infiltration validity of Philip's two-term equation is reached before the reference curves attain a constant infiltration rate. This is the reason so as to why the solid blue and green lines diverge so much for the (a) sandy loam, (b) sandy clay loam, (d) silty loam, (e) clay loam, (f) silty clay loam, and (h) clay soils and reinforces that the infiltration validities of Table 1 have meaning only in the context of Richardson's equation. Indeed, if we repeat the analysis of Jaiswal et al. (2022) but use instead the infiltration data of the reference set displayed in Figure 9 then we expect an enhanced infiltration validity of Philip's two-term equation for all soils but (c) loam and (g) silty clay. This will substantially decrease the differences between the least squares coefficients, $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$, of the full and partial experiments, respectively, and improve considerably the compliance of the blue and green lines. Thus, the infiltration validity of 2.5 cm of Philip's two-term equation lacks a solid physical foundation if the data generating infiltration process does not follow Richardson's equation.

To verify our conclusions, we repeat our analysis with Philip's two-term equation for a second reference infiltration set created with the infiltration model of Green and Ampt (1911). In Appendix C we present theory and a MATLAB implementation of an efficient, exact, and robust numerical solution of the Green and Ampt model. Appendix D summarizes how this second reference set was created and Figure E.1

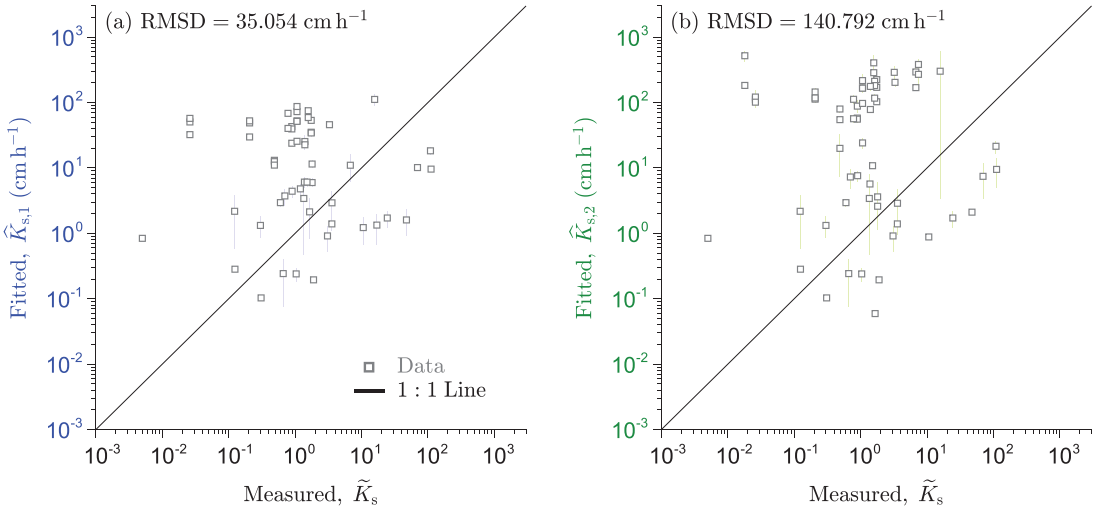

FIGURE 10 Soil Water Infiltration Global (SWIG) database: Scatter plots of the measured saturated hydraulic conductivity, $\widetilde{K}_s$, and their counterparts (a) $\widehat{K}_{s,1}$ and (b) $\widehat{K}_{s,2}$ derived from the full and partial experiments, respectively. The solid black line characterizes the 1:1 relationship between the plotted quantities. The horizontal blue and vertical green lines display the 95% confidence intervals of the coefficients.

displays the results in a format similar to Figures B.1, 6, and 9 for the HYDRUS-1D, SWIG, and Horton infiltration data sets, respectively. We observe a similar divergence between the infiltration curves of the full and partial experiments, but with one fundamental difference. The infiltration curves of the full experiment (solid blue lines) are in perfect agreement with the "*measured*" data (solid red dots). Philip's two-term equation can describe exactly the simulated infiltration data of the Green and Ampt model. The divergence between the solid blue and green lines is a result of an insufficiently large infiltration validity for all soils but (c) loam, (e) clay loam, and (g) silty clay. This confirms our earlier conclusions that the infiltration validity of Philip's two-term equation depends on assumptions made about the data generating infiltration process. As this process is not known exactly, the infiltration validity should be considered a random variable with probability distribution defined by the epistemic uncertainty.

Finally, we turn our attention to the measured K_s values of the SWIG database (see Figure 10). Of the collection of 226 experiments with at least $n = 5$ data points in the range of $0 < \tilde{I}(\tilde{t}) \leq 2.5$ cm of the partial experiments, 76 experiments report a measured value of the saturated hydraulic conductivity. As expected, we find a rather poor agreement between the measured saturated hydraulic conductivities, $\widetilde{K}_s$, of the SWIG experiments and their least squares estimates, $\widehat{K}_{s,1}$ and $\widehat{K}_{s,2}$, of the full and partial experiments, respectively. This highlights the consequences of epistemic uncertainty. In our desire to minimize the sum of squares of the cumulative infiltration residuals, the least squares values of S and K_s will compensate for structural errors of Philip's two-term equation. The resulting estimates of S and K_s are subject to bias and relate poorly to their measured counterparts. In fact, the listed RMSD values suggest that the values of $\widehat{K}_{s,1}$ of the full experiments are in much better agreement with the measured saturated hydraulic conductivities than their counterparts, $\widehat{K}_{s,2}$, of the partial experiments. Thus, there appears to be no benefits in using a partial infiltration curve up to $\tilde{I} = 2.5$ cm for inverse estimation of S and K_s, Seemingly, epistemic uncertainty supports the familiar credo and/or paradigm, "*the more data the better*." With an infiltration function that is known to be deficient at late times, this conclusion is counter-intuitive and at odds with our earlier findings for the HYDRUS-1D infiltration data set. As epistemic uncertainty is more the rule than exception, we should not expect the coefficients S and K_s to represent innate soil properties and the infiltration validity to be a uniquely defined theoretical quantity. Certainly, S and K_s should be treated similar to fitting coefficients in a statistical regression function.

To minimize the impact of structural errors on the estimates of S and K_s there is one data-mining methodology that we have not explored in this paper. Strictly speaking, we do not need to rely on an infiltration function to obtain estimates of S and K_s. The sorptivity, S, is synonymous to the slope of $(\sqrt{\tilde{t}}, \tilde{I})$ data pairs at early infiltration times. The product of the multiple, c, and saturated soil hydraulic conductivity, K_s, should match the slope of the $(\tilde{t}, \tilde{I})$ data pairs at late times of infiltration. This data-driven method is not uncommon but requires a formal definition of early and late infiltration times and their dependence on soil type. We have not investigated this data-mining approach in the present paper.

A final remark is appropriate. The past two decades have witnessed the development of Bayesian methods for parameter estimation and inverse modeling. The Bayesian approach offers an arsenal of methodological advances over box-constrained least squares estimation used herein. Specifically, the distribution-adaptive likelihood (objective) functions of

Schoups and Vrugt (2010) and Vrugt et al. (2022) will much better characterize the non-normal distributions of the cumulative infiltration residuals in the face of epistemic errors. This should reduce the impact of model misspecification on the parameter estimates. Unfortunately, we cannot recommend this approach in the present context as most SWIG experiments have an insufficient number of $(\tilde{t}, \tilde{I})$ data pairs to warrant an accurate characterization of the cumulative infiltration residuals and corresponding bivariate posterior probability distribution of the sorptivity and saturated soil hydraulic conductivity.

4 | IMPLICATIONS

While the definition of infiltration is seemingly plain and simple, it is a complicated process with different flow types (steady state, transient, saturated, and unstable flow) acting together to move fluid through a heterogeneous soil medium with different degrees of anisotropy. Infiltration determines not only the amount of water that will enter a soil, but also governs the incoming flux of dissolved passenger chemicals such as nutrients and pollutants. The rate and pattern of infiltration will depend on rainfall intensity and distribution, the depth of the water table, the physical properties of the underlying soil column, and its antecedent moisture content before wetting (Ferré & Warrick, 2005). The physical properties are often understood to be the water retention and hydraulic conductivity functions of the soil matrix, but infiltration is also controlled by surface sealing and crusting, hydrophobicity, the ionic composition of the infiltrating water (Hopmans et al., 2006), soil layering, and variability (Steenhuis et al., 1991). Epistemic errors are unavoidable in this pursuit of a convenient and tractable physical– mathematical description of the infiltration process. We wish to briefly discuss the implications of our findings.

We cannot control and/or resolve epistemic errors of infiltration functions but can do a better job in constraining our physical–mathematical description of the infiltration process. The soil's initial moisture content determines its sorptivity but is irrelevant to Philip's two-term infiltration function and the infiltration functions of Green and Ampt (1911), Kostiakov (1932), Horton (1941), and Mezencev (1948). This neglect of the soil's initial state will impact the estimates of the sorptivity and saturated soil hydraulic conductivity and favor the three-parameter infiltration equation of Parlange et al. (1982). This quasi-exact implicit solution of Richardson's equation accounts for the soil's initial state and is valid for the entire infiltration event (Haverkamp et al., 1994). Its constants S, K_s, and ϕ exhibit a solid mathematical–physical underpinning and should be thought of as super parameters of the hydraulic functions (Vrugt & Gao, 2022). These parameters may characterize well the hydraulic behavior of the soil wetted by the infiltration experiment but should not be expected to yield accurate predictions of infiltration, runoff, and recharge for nearby soil pedons let alone extrapolate well to the hillslope and watershed scale. The culprit is spatial heterogeneity, specifically, the large variability in soil physical properties such as surface sealing and crusting, hydrophobicity, and the hydraulic functions. Juxtaposed with spatial variability in rainfall intensity and amount and differences in initial moisture content give rise to complex patterns of infiltration, runoff, and recharge at pedon to watershed scales.

In summary, though infiltration measurements and modeling have significantly improved our scientific understanding of key processes of soil hydrology, the super parameters of the hydraulic functions embedded in the infiltration functions cannot simply be used in storm water and vadose zone flow models to forecast recharge and runoff at the field and landscape scales unless we can accompany these predictions with realistic uncertainty bounds. The joint use of multiple different algebraic infiltration equations coupled with Bayesian inference is one way to account for measurement and epistemic uncertainty.

5 | CONCLUSIONS

Philip's two-term infiltration equation is widely used by researchers and practitioners to describe the cumulative vertical infiltration into unsaturated soils. This quasi-analytical approximation of Richardson's equation is known to be deficient at late times of infiltration, nevertheless, it is quite common to use all $(\tilde{t}, \tilde{I})$ data pairs of the measured $\tilde{I}(\tilde{t})$ relationship to estimate its unknown coefficients, the soil sorptivity, S, and product, cK_s, of the soil-dependent multiple, c, and saturated hydraulic conductivity, K_s. Inevitably, this cavalier practice will bias the least squares estimates of S and K_s. This is true even if the infiltration data is restricted to the characteristic time obtained from Philip (1969) and Rahmati et al. (2022) as t_{char} exceeds by far the valid time interval of Philip's two-term infiltration equation. In this paper, we focused our attention on the impacts of time validity violation on the estimates S and K_s. We used synthetic and measured infiltration experiments to compare, contrast, and evaluate estimates of S and K_s derived from cavalier practice of using all measured (I, t) data pairs against those obtained from a partial infiltration curve using only the $\tilde{I}(\tilde{t})$ relationship up to the infiltration validity of 2.5 cm of Philip's two-term equation. Specific attention in this paper was given to the role of epistemic error on the inferred parameters and time validities.

The results of our synthetic case study with vertical infiltration data simulated from a numerical solution of Richardson's equation, demonstrate the advantages of a partial infiltration curve for least squares estimation of S and K_s. The parameter estimates derived from partial experiments are in closer

agreement with their "*true*" counterparts. Full experiments tend to overestimate the sorptivity and underestimate the saturated hydraulic conductivity. Albeit important, the differences in the parameter estimates of the full and partial experiments do not lead to large differences in the simulated infiltration curves. Thus, a violation of the infiltration validity of Philip's two-term equation has a relatively minor impact on the optimal values of the coefficients, S and K_s.

The results of our second case study with measured data of the SWIG database demonstrate what happens if the data generating infiltration process does not follow Richardson's equation. Structural errors can bias the parameter estimates of Philip's two-term equation and even bring into question the accuracy and robustness of the inferred infiltration validities. For most soils, we observed rather large differences in the least squares estimates of S and K_s and corresponding infiltration curves of the partial and full experiments. The inferred saturated hydraulic conductivities exhibited a poor correspondence to their measured values.

The results of our third case study using infiltration data simulated with Horton's equation confirm our earlier findings. Structural errors corrupt the estimates of S and K_s and turn the time validity into an elusive theoretical quantity with poor physical underpinning. This expands on earlier conclusions of Haverkamp et al. (1988) who based the physical insignificance of coefficients S and K_s in Philip's two-term expression on their time dependence. In fact, this conclusion is not limited to Philip's power series approximation but will hold for all commonly used infiltration functions if the infiltration process does not follow Richardson's equation. Therefore, it would be more productive to treat the infiltration validity of Philip's two-term equation as a random variable with unknown probability density function (PDF). We can estimate this PDF by repeated application of the approach detailed in Jaiswal et al. (2022) to synthetic data created from a suite of different infiltration functions. If the data generating process is contained within the ensemble of infiltration functions, then the so-obtained 95% confidence intervals of the infiltration validity of Philip's two-term equation should be robust for each soil type. We can draw samples from the PDF of the infiltration validity, and determine the corresponding values of S and K_s. This results in a bivariate probability distribution of the soil sorptivity and saturated hydraulic conductivity for the partial and full infiltration experiments. This is the best we can do with Philip's two-term infiltration equation in the presence of epistemic uncertainty.

Finally, to enhance the practical use of plot-scale infiltration experiments for prediction of runoff and recharge at the field and watershed scales we must account explicitly for epistemic uncertainty.

DATA AND SOFTWARE AVAILABILITY

The data and software are available upon request from the corresponding author and can be downloaded using the following link: https://github.com/jaspervrugt/Projects/Infiltration

AUTHOR CONTRIBUTIONS

Jasper A. Vrugt: Conceptualization; data curation; formal analysis; investigation; methodology; software; supervision; validation; visualization; writing–original draft; writing–review and editing. **Jan W. Hopmans**: Conceptualization; writing–original draft. **Yifu Gao**: Software; visualization. **Mehdi Rahmati**: Conceptualization; data curation. **Jan Vanderborght**: Resources. **Harry Vereecken**: Resources.

ACKNOWLEDGMENTS

The reviews from Ty Ferré, Binayak Mohanty (AE) and an anonymous referee are greatly appreciated and have helped to further improve this paper.

ORCID

Jasper A. Vrugt https://orcid.org/0000-0003-2599-1165
Jan W. Hopmans https://orcid.org/0000-0002-4807-2172
Yifu Gao https://orcid.org/0000-0002-7055-4004
Mehdi Rahmati https://orcid.org/0000-0001-5547-6442
Jan Vanderborght https://orcid.org/0000-0001-7381-3211
Harry Vereecken https://orcid.org/0000-0002-8051-8517

REFERENCES

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1), 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007

Bristow, K. L., & Savage, M. J. (1987). Estimation of parameters for the philip two-term infiltration equation applied to field soil experiments. *Australian Journal of Soil Research*, *25*, 369–375. https://doi.org/10.1071/SR9870369

Carsel, R. F., & Parrish, R. S. (1988). Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*, *24*(5), 755–769. https://doi.org/10.1029/WR024i005p00755

Darcy, H. (1856). *Fontaines publiques de la ville de Dijon: Exposition et application des principes a suivre et des formulesa employer dans les questions de distribution d'eau*. https://gallica.bnf.fr/ark:/12148/bpt6k624312/f1n657.pdf

Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression. I. *Biometrika*, *37*(3-4), 409–428. https://doi.org/10.1093/biomet/37.3-4.409

Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika*, *38*(1-2), 159–179. https://doi.org/10.1093/biomet/38.1-2.159

Farebrother, R. W. (1980). The Durbin-Watson test for serial correlation when there is no intercept in the regression. *Econometrica*, *48*(6), 1553–1563. https://doi.org/10.2307/1912825

Ferré, T. P. A., & Warrick, A. W. (2005). Infiltration. In D. Hillel (Ed.), *Encyclopedia of soils in the environment* (pp. 254–260). Elsevier. https://doi.org/10.1016/B0-12-348530-4/00382-9

Green, W. H., & Ampt, G. A. (1911). Studies on soil physics. *The Journal of Agricultural Science*, *4*(1), 1–24. https://doi.org/10.1017/S0021859600001441

Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1998). Toward improved calibration of hydrologic models: Multiple and non-commensurable measures of information. *Water Resources Research*, *34*(4), 751–763. https://doi.org/10.1029/97WR03495

Gupta, H. V., Wagener, T., & Liu, Y. (2008). Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrological Processes*, *22*(18), 3802–3813. https://doi.org/10.1002/hyp.6989

Haverkamp, R., Kutilek, M., Parlange, J. Y., Rendon, L., & Krejca, M. (1988). Infiltration under ponded conditions: 2. Infiltration equations tested for parameter time-dependence and predictive use. *Soil Science*, *145*, 317–329. https://journals.lww.com/soilsci/fulltext/1988/05000/infiltration_under_pondedconditions_2.1.aspx

Haverkamp, R., Ross, P. J., Smettem, K. R. J., & Parlange, J.-Y. (1994). Three-dimensional analysis of infiltration from the disc infiltrometer: 2. Physically based infiltration equation. *Water Resources Research*, *30*(11), 2931–2935. https://doi.org/10.1029/94WR01788

Hopmans, J. W., Parlange, J.-Y., & Assouline, S. (2006). Infiltration. In J. W. Delleur (Ed.), *The handbook of groundwater engineering* (2 ed., pp. 18). CRC Press.

Horton, R. E. (1941). An approach toward a physical interpretation of infiltration capacity. *Soil Science Society of America Journal*, *5*, 339–417. https://doi.org/10.2136/sssaj1941.03615995000500C0075x

Hunt, A. G., Holtzman, R., & Ghanbarian, B. (2017). A percolation-based approach to scaling infiltration and evapotranspiration. *Water*, *9*(2), 104.

Jaiswal, P., Gao, Y., Rahmati, M., Vanderborght, J., Šimůnek, J., Vereecken, H., & Vrugt, J. A. (2022). Parasite inversion for determining the coefficients and time-validity of Philip's two-term infiltration equation. *Vadose Zone Journal*, *21*, e20166. https://doi.org/10.1002/vzj2.20166

Knight, J., & Raats, P. (2016). The contributions of Lewis Fry Richardson to drainage theory, soil physics, and the soil-plant-atmosphere continuum. *Geophysical Research Abstracts*, 18EGU2016-10980-1. https://meetingorganizer.copernicus.org/EGU2016/EGU2016-10980-1.pdf

Kostiakov, A. N. (1932). On the dynamics of the coefficients of water percolation in soils and on the necessity of studying it from a dynamic point of view for purpose of amelioration. In *Transactions of 6th Committee International Society of Soil Science, Russia, Part A* (pp. 17–21).

Kunze, R. J., & Nielsen, D. R. (1982). Finite-difference solutions of the infiltration equation. *Soil Science*, *34*, 81–88. https://journals.lww.com/soilsci/fulltext/1982/08000/finite_difference_solutions_of_the_infiltration.1.aspx

Lawson, C. L., & Hanson, R. J. (1995). *Solving least squares problems*. Society of Industrial and Applied Mathematics. https://doi.org/10.1137/1.9781611971217

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, *2*(2), 164–168.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, *11*(2), 431–441. https://doi.org/10.1137/0111030

Mezencev, V. J. (1948). Theory of formation of the surface runoff. *Meteorologiae Hidrologia*, *3*, 33–40.

Mishra, S. K., Tyagi, J. V., & Singh, V. P. (2003). Comparison of infiltration models. *Hydrological Processes*, *17*, 2629–2652. https://doi.org/10.1002/hyp.1257

Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, *12*(3), 513–522. https://doi.org/10.1029/WR012i003p00513

Parlange, J.-Y., Lisle, I., Braddock, R. D., & Smith, R. E. (1982). The three-parameter infiltration equation. *Soil Science*, *133*(6), 337–341. https://doi.org/10.1097/00010694-198206000-00001

Philip, J. R. (1954). An infiltration equation with physical significance. *Soil Science*, *77*, 153–158.

Philip, J. R. (1955). Numerical solution of equations of the diffusion type with diffusivity concentration-dependent. *Transactions of the Faraday Society*, *51*, 885–892. https://doi.org/10.1039/TF9555100885

Philip, J. R. (1957a). Numerical solution of equations of the diffusion type with diffusivity concentration-dependent. II. *Australian Journal of Physics*, *10*, 29–42. https://api.semanticscholar.org/CorpusID:121375936

Philip, J. R. (1957b). The theory of infiltration: 1. The infiltration equation and its solution. *Soil Science*, *83*(5), 345–357. https://doi.org/10.1097/00010694-195705000-00002

Philip, J. R. (1957c). The theory of infiltration: 2. The profile at infinity. *Soil Science*, *83*(6), 435–448. https://doi.org/10.1097/00010694-195706000-00003

Philip, J. R. (1957d). The theory of infiltration: 3. Moisture profiles and relation to experiment. *Soil Science*, *84*, 163–178. https://doi.org/10.1097/00010694-195708000-00008

Philip, J. R. (1957e). The theory of infiltration: 4. Sorptivity and algebraic infiltration equations. *Soil Science*, *84*, 257–264. https://api.semanticscholar.org/CorpusID:93230567

Philip, J. R. (1957f). The theory of infiltration: 5. The influence of the initial moisture content. *Soil Science*, *84*, 329–340.

Philip, J. R. (1969). Theory of infiltration. *Advances in Hydroscience*, *5*, 215–296. https://doi.org/10.1016/B978-1-4831-9936-8.50010-6

Rahmati, M., Latorre, B., Lassabatere, L., Angulo-Jaramillo, R., & Moret-Fernández, D. (2019). The relevance of philip theory to Haverkamp quasi-exact implicit analytical formulation and its uses to predict soil hydraulic properties. *Journal of Hydrology*, *570*, 816–826. https://doi.org/10.1016/j.jhydrol.2019.01.038

Rahmati, M., Latorre, B., Moret-Fernández, D., Lassabatere, L., Talebian, N., Miller, D., Morbidelli, R., Iovino, M., Bagarello, V., Neyshabouri, M. R., Zhao, Y., Vanderborght, J., Weihermüller, L., Jaramillo, R. A., Or, D., Th.van Genuchten, M., & Vereecken, H. (2022). On infiltration and infiltration characteristic times. *Water Resources Research*, *58*(5), e2021WR031600. https://doi.org/10.1029/2021WR031600

Rahmati, M., Vanderborght, J., Šimůnek, J., Vrugt, J. A., Moret-Fernández, D., Latorre, B., Lassabatere, L., & Vereecken, H. (2020). Soil hydraulic properties estimation from one-dimensional infiltration experiments using characteristic time concept. *Vadose Zone Journal*, *19*, 1–22. https://doi.org/10.1002/vzj2.20068

Rahmati, M., Weihermüller, L., Vanderborght, J., Pachepsky, Y. A., Mao, L., Sadeghi, S. H., Moosavi, N., Kheirfam, H., Montzka, C., Van Looy, K., Toth, B., Hazbavi, Z., Al Yamani, W., Albalasmeh, A. A., Alghzawi, M. Z., Angulo-Jaramillo, R., Antonino, A. C. D.,

Arampatzis, G., Armindo, R. A., … Vereecken, H. (2018). Development and analysis of the soil water infiltration global database. *Earth System Science Data*, *10*(3), 1237–1263. https://doi.org/10.5194/essd-10-1237-2018

Richards, L. A. (1931). Capillary conduction of liquids through porous mediums. *Journal of Applied Physics*, *1*(5), 318–333. https://doi.org/10.1063/1.1745010

Richardson, L. F. (1922). *Weather prediction by numerical process*. Cambridge University Press. https://doi.org/10.1017/CBO9780511618291

Sayah, B., Gil-Rodríguez, M., & Juana, L. (2016). Development of one-dimensional solutions for water infiltration. Analysis and parameters estimation. *Journal of Hydrology*, *535*, 226–234. https://doi.org/10.1016/j.jhydrol.2016.01.026

Schoups, G., & Vrugt, J. A. (2010). A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research*, *46*(10), W10531. https://doi.org/10.1029/2009WR008933

Stark, P. B., & Parker, R. L. (1995). Bounded-variable least-squares: An algorithm and applications. *Computational Statistics*, *10*, 129–141.

Steenhuis, T. S., Selker, J. S., Bell, J. L., Kung, K. J. S., & Parlange, J. -Y. (1991). Effects of soil layering on infiltration. In *Irrigation and drainage* (pp. 74–80). ASCE.

Talsma, T., & Parlange, J.-Y. (1972). One dimensional vertical infiltration. *Soil Research*, *10*, 143–150.

The Mathworks. (2021). MATLAB [Computer software]. https://www.mathworks.com/products/matlab.html

van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, *44*(5), 892–898. https://doi.org/10.2136/sssaj1980.03615995004400050002x

Vrugt, J. A., & Beven, K. (2018). Embracing equifinality with efficiency: Limits of acceptability sampling using the DREAM$_{(LOA)}$ algorithm. *Journal of Hydrology*, *559*, 954–971. https://doi.org/10.1016/j.jhydrol.2018.02.026

Vrugt, J. A., de Oliveira, D. Y., Schoups, G., & Diks, C. G. (2022). On the use of distribution-adaptive likelihood functions: Generalized and universal likelihood functions, scoring rules and multi-criteria ranking. *Journal of Hydrology*, *615*, 128542. https://doi.org/10.1016/j.jhydrol.2022.128542

Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W., & Verstraten, J. M. (2005). Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. *Water Resources Research*, *41*(1), W01017. https://doi.org/10.1029/2004WR003059

Vrugt, J. A., & Gao, Y. (2022). On the three-parameter infiltration equation of parlange et al. (1982): Numerical solution, experimental design, and parameter estimation. *Vadose Zone Journal*, *21*, e20167. https://doi.org/10.1002/vzj2.20167

Vrugt, J. A., & Sadegh, M. (2013). Toward diagnostic model calibration and evaluation: Approximate Bayesian computation. *Water Resources Research*, *49*(7), 4335–4345. https://doi.org/10.1002/wrcr.20354

Šimůnek, J., van Genuchten, M. T., & Sejna, M. (2008). Development and applications of the HYDRUS and STANMOD software packages and related codes. *Vadose Zone Journal*, *7*(2), 587–600. https://doi.org/10.2136/vzj2007.0077

Šimůnek, J., van Genuchten, M. T., & Sejna, M. (2016). Recent developments and applications of the HYDRUS computer software packages. *Vadose Zone Journal*, *15*(7), 1–25. https://doi.org/10.2136/vzj2016.04.0033

Warrick, A. W. (2003). *Soil water dynamics*. Oxford University Press.

How to cite this article: Vrugt, J. A., Hopmans, J. W., Gao, Y., Rahmati, M., Vanderborght, J., & Vereecken, H. (2024). The time validity of Philip's two-term infiltration equation: An elusive theoretical quantity? *Vadose Zone Journal*, e20309. https://doi.org/10.1002/vzj2.20309

APPENDIX A: PHILIP'S TWO-TERM INFILTRATION EQUATION IN MATLAB

In this Appendix we present a MATLAB subroutine, called fit_Philip2 which implements least squares estimation for the parameters β_1 and β_2 of Philip's two-term infiltration equation. For didactic purposes, we do not take advantage of built-in functions such as `nlparci` and `nlpredci` to determine confidence limits of the coefficients and simulated cumulative infiltration curve.

```
function [b,bc,Ip,Ic,dw] = fit_Philip2(tm,Im,c,Imax,alfa)
%% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% %%
%% This function computes the coefficients of Philip's two-term infiltration equation %%
%%     I(t) = b_{1}*t^(1/2) + b_{2}*c*t,                                            %%
%% where the coefficients, b_{1} and b_{2}, are strictly positive                   %%
%%                                                                                  %%
%%  SYNOPSIS: [b,bc,Iopt,Ic,dw] = fit_Philip2(tm,Im)                                 %%
%%            [b,bc,Iopt,Ic,dw] = fit_Philip2(tm,Im,c)                               %%
%%            [b,bc,Iopt,Ic,dw] = fit_Philip2(tm,Im,c,Imax)                          %%
%%            [b,bc,Iopt,Ic,dw] = fit_Philip2(tm,Im,c,Imax,alfa)                     %%
%% where                                                                            %%
%%  tm   [input]       nx1 vector with measured times, t, in hours                  %%
%%  Im   [input]       nx1 vector with measured cum. infiltration, I, in cm         %%
%%  c    [opt. input]  scalar with soil dependent value of c         (default: 1)   %%
%%  Imax [opt. input]  scalar with infiltration validity in cm      (default: inf)  %%
%%  alfa [opt. input]  scalar with significance level               (default: 0.05) %%
%%  b    [output]      1x2 vector with least squares coefficient values             %%
%%  bc   [output]      2x2 matrix with 100(1-alfa)% confidence limits of coefficients %%
%%  Ip   [output]      nx1 vector with least squares simulated inf. curve           %%
%%  Ic   [output]      nx2 matrix with 100(1-alfa)% confidence limits of inf. curve  %%
%%  dw   [output]      scalar with Durbin-Watson statistic                          %%
%% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% %%
if nargin < 5, alfa = 0.05; end
if nargin < 4, Imax = inf; end
if nargin < 3, c = 1; end
idx = (tm > 0) + (Im <= Imax);                     % Index of t > 0 and Im <= Imax
tm = tm(idx==2); Im = Im(idx==2);                  % Time and infiltration vector
n = size(tm,1); df = n - 2;                        % # data points, degrees freedom
switch df
    case {-1,0}
        error('fit_Philip2-error: insufficient data');
    case {1,2,3}
        fprintf('fit_Philip2-warning: only %d points, %d degrees freedom\n',...
            n,df);
end
options = optimoptions('lsqlin','display','none'); % no output writing to screen
D = [ sqrt(tm) , c*tm ];                           % nx2 design matrix
b = lsqlin(D,Im,[],[],[],[],eps*ones(2,1),...
    [],[],options);                                % box-constrained least squares
Ip = D*b;                                          % nx1 vector optimum inf. curve
r = Im - Ip;                                       % nx1 vector of residuals
s2_I = sum(r.^2)/df;                               % Sample variance of residuals
C = s2_I * inv(D'*D);                              % 2x2 coef. covariance matrix
t_crit = tinv(1/2 + (1 - alfa)/2,df);              % Critical t value
br = t_crit * sqrt(diag(C));                       % 1-alfa conf. ranges of coef.
bc = [ b - br , b + br ];                          % 1-alfa conf. limits of coef.
dIc = t_crit * sqrt(diag(D * C * D'));             % 1-alfa conf. ranges of inf. curve
Ic = [ Ip - dIc , Ip + dIc ];                      % 1-alfa conf. limits of inf. curve
dw = sum(diff(r).^2)/sum(r.^2);                    % Durbin-Watson statistic
```

The `fit_Philip2` subroutine has five input arguments, three of which are optional. Required inputs are the $n \times 1$ vectors, `tm` and `Im`, of measurement times and cumulative infiltration in units of hour and centimeter, respectively. Optional input arguments are the dimensionless soil-dependent coefficient `c`, the infiltration validity, `Imax`, of Philip's two-term equation in units of centimeter and the significance level, `alfa`. This latter input argument is required in the computation of the confidence limits of the coefficients and simulated infiltration curve. If the user does not specify the optional input arguments, default values of $c = 1$, $I_{\text{valid}} \to \infty$, and $\alpha = 0.05$ are assumed.

The function returns five output arguments, which are discussed in order. The variable `b` is synonymous to $\hat{\beta}$ and equals a 2×1 vector with the least squares values of the coefficients. The second output argument, `bc`, is equivalent to a 2×2 matrix, and stores the lower, $\hat{\beta}^-$, and upper, $\hat{\beta}^+$, limits of the $100(1-\alpha)\%$ confidence intervals of `b`. The third output argument, `Ip`, is a $n \times 1$ vector, $\hat{\mathbf{I}}_n$, with simulated cumulative infiltration curve of `b`. The fourth output

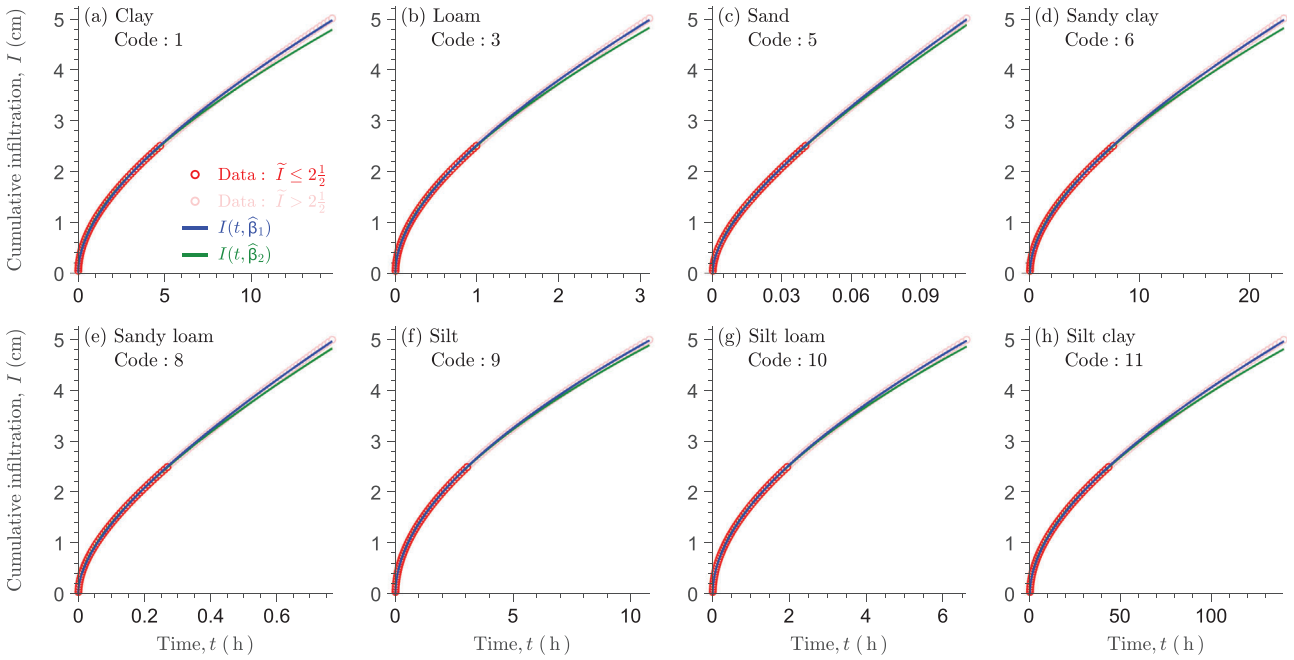

FIGURE B.1 HYDRUS-1D infiltration data set: Observed (red dots) and simulated cumulative infiltration curves for the selection of eight SWIG soils using the least squares values of S, K_s, and c (Table 1) of Philip's two-term infiltration equation for the full (solid blue line) and partial (solid green line) experiments. Each graph corresponds to a different soil type, specifically (a) clay, (b) loam, (c) sand, (d) sandy clay, (e) sandy loam, (f) silt, (g) silt loam and (h) silt clay. Measured cumulative infiltration data beyond the time validity of Philip's two-term equation are displayed with a tint of red. The 95% confidence intervals are too small to be visible..

argument, `Ic`, is a $n \times 2$ matrix with lower, $\widehat{\mathbf{I}}_c^-$, and upper, $\widehat{\mathbf{I}}_c^+$, limits of the $100(1-\alpha)\%$ confidence intervals of `Ip`. The fifth and last output argument, `dw`, equals the value of the Durbin and Watson (1950, 1951) statistic, d_w. Built-in functions are highlighted with a low dash.

APPENDIX B: SIMULATED INFILTRATION CURVES FOR HYDRUS-1D DATA SET

This Appendix (Figure B.1) presents the simulated cumulative infiltration curves of Philip's two-term equation for the HYDRUS-1D data set.

APPENDIX C: GREEN AND AMPT INFILTRATION MODEL: NUMERICAL SOLUTION

In this Appendix we review briefly the infiltration function of Green and Ampt (1911) and present a CPU- efficient and accurate numerical solution and implementation of this model in MATLAB.

The Green and Ampt (1911) model describes vertical infiltration into an unsaturated soil. This semi-analytic infiltration function relies on Darcy's law (Darcy, 1856) and assumes that the soil profile is homogeneous with (i) uniform initial volumetric moisture content, θ_i ($L^3\ L^{-3}$), (ii) horizontal wetting front, (iii) constant soil water pressure head, h_f (L), at the wetting front independent of the wetting front's vertical position, and (iv) constant volumetric water content, θ_0 ($L^3\ L^{-3}$), soil water pressure head, h_0 (L), and hydraulic conductivity, K_0 ($L\ T^{-1}$), of the wetted (= transmission) zone. Then, the cumulative vertical infiltration, I (L), at time t (T) equals

$$I(t) = \Delta h \Delta\theta \log\left(1 + \frac{I(t)}{\Delta h \Delta\theta}\right) + K_0 t \qquad \text{(C.1)}$$

where $\Delta h = h_0 - h_F$ and $\Delta\theta = \theta_0 - \theta_i$. As the entities, Δh and $\Delta\theta$ appear as a product in the above equation, we cannot infer their values separately from measured $(\tilde{t}, \tilde{I})$ data pairs. Rather, we can only estimate their product, $\xi = \Delta h \Delta\theta$ (L).

The Green and Ampt (1911) model of Equation (C.1) demands an iterative numerical method to solve for the $I(t)$ relationship for given values of the coefficients, ξ and K_0. To explicate the numerical solution, lets write Equation (C.1) in residual form

$$r(I, t, \xi, K_0) = I - \xi \log\left(1 + \frac{I}{\xi}\right) - K_0 t. \qquad \text{(C.2)}$$

We resort to Newton's method to solve for the zero points of the residual function

$$I_{(k+1)} = I_{(k)} - \frac{r(I_{(k)}, t, \xi, K_0)}{r'(I_{(k)}, t, \xi, K_0)} \qquad \text{(C.3)}$$

where $r'(I, t, \xi, K_0)$ signifies the derivative of the residual function in Equation (C.2) with respect to I and k

denotes the iteration counter. The derivative, $r'(I, t, \xi, K_0)$, will help determine the direction of movement to the zero point. This partial derivative may be derived using symbolic differentiation

$$\frac{\partial r}{\partial I} = 1 - \frac{1}{1 + I/\xi} = \frac{I/\xi}{1 + I/\xi}. \tag{C.4}$$

If so desired, we can substitute the expressions of $r(I, t, \xi, K_0)$ and $r'(I, t, \xi, K_0)$ in Equations (C.3) and (C.4). This produces the following recurrence relation

$$\begin{aligned}
I_{(k+1)} &= I_{(k)} - \frac{I_{(k)} - \xi \log(1 + I_{(k)}/\xi) - K_0 t}{(I_{(k)}/\xi)/(1 + I_{(k)}/\xi)} \\
&= I_{(k)} - \frac{(1 + I_{(k)}/\xi)\big(I_{(k)} - \xi \log(1 + I_{(k)}/\xi) - K_0 t\big)}{(I_{(k)}/\xi)} \\
&= I_{(k)} - \frac{I_{(k)} + I_{(k)}^2/\xi - \xi \log(1 + I_{(k)}/\xi) - I_{(k)} \log(1 + I_{(k)}/\xi) - K_0 t - (I_{(k)}/\xi) K_0 t}{(I_{(k)}/\xi)} \\
&= I_{(k)} - \left(\xi + I_{(k)} - \frac{\xi^2}{I_{(k)}} \log\left(1 + \frac{I_{(k)}}{\xi}\right) - \xi \log\left(1 + \frac{I_{(k)}}{\xi}\right) - \frac{\xi}{I_{(k)}} K_0 t - K_0 t\right) \\
&= -\xi + \frac{\xi^2}{I_{(k)}} \log\left(1 + \frac{I_{(k)}}{\xi}\right) + \xi \log\left(1 + \frac{I_{(k)}}{\xi}\right) + \frac{\xi}{I_{(k)}} K_0 t + K_0 t.
\end{aligned} \tag{C.5}$$

The above expression may be simplified to read

$$I_{(k+1)} = -\xi + \left(1 + \frac{\xi}{I_{(k)}}\right)\left(\xi \log\left(1 + \frac{I_{(k)}}{\xi}\right) + K_0 t\right). \tag{C.6}$$

Now we have available a numerical solution for Equation (C.1) we present its implementation in MATLAB. The source code consists of the main function called `Green_Ampt`, which uses the iterative recipe of Equation (C.6) to find the zero points of the residual function, $r(I, t, \xi, K_0)$, of Equation (C.2) for each of the n entries of the time vector, $\mathbf{t} = (t_1 \ \dots \ t_n)^\top$.

The `Green_Ampt` subroutine uses as input arguments, the 2×1-vector, `eta`, of parameter values, ξ, cm, and K_0 (cm h^{-1}), and the $n \times 1$ vector of measurement times, $\mathbf{t} = (t_1 \ \dots \ t_n)^\top$, in units of hour. The function returns as output arguments the $n \times 1$ vectors of cumulative infiltration, $\mathbf{I} = (I_1 \ \dots \ I_n)^\top$ in units of cm, and the corresponding infiltration rate, $\mathbf{i} = (i_1 \ \dots \ i_n)^\top$ in cm h^{-1}. Built-in functions are highlighted with a low dash.

```
function [I,i] = Green_Ampt(eta,t,rtol,kmax)
%% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% This function evaluates the infiltration equation of Green and Ampt                %%
%%  SYNOPSIS: I = Green_Ampt(eta,t);                                                   %%
%%            I = Green_Ampt(eta,t,rtol);                                              %%
%%            I = Green_Ampt(eta,t,rtol,kmax);                                         %%
%% where                                                                               %%
%%  eta: [input]       2x1 vector with xi (cm) and K0 (cm/h)                          %%
%%  t:   [input]       nx1 vector with time, t, in hours (h)                          %%
%%  rtol [opt. input]  tolerance on function value at root        (default: 1e-12)    %%
%%  kmax [opt. input]  maximum number of Newton iterations        (default: 20)       %%
%%  I:   [output]      Cumulative infiltration in cm as function of time, t (h)       %%
%%  i:   [output]      Infiltration rate in cm/h as function of time, t (h)           %%
%% %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if nargin < 4, kmax = 20; end
if nargin < 3, rtol = 1e-12; end
%% Initialization
n = numel(t); I = nan(n,1);                          % Length data and init. cum. inf.
switch any(isnan(eta))
    case 0, xi = eta(1); K0 = eta(2);                    % Unpack coefficients
    otherwise, return                                    % Return
end
r = @(I,t) I-xi*log(1+I/xi)-K0*t;                    % Residual function
switch t(1)==0                                       % Check first time index
    case 0, j = 1;
    otherwise, I(1) = 0; j = 2;
end
I_low = t(2)*max(K0,0.1);                            % Initial root guess
%% Dynamic part
for z = j:n                                          % Loop over time
    y(1) = I_low; k = 1;                                 % Root guess and iter.
    while (abs(r(y(k),t(z)))>rtol) && (k<kmax)           % While loop
        y(k+1) = -xi + (1 + xi/y(k)) * ...                   % New iterate
            (xi*log(1 + y(k)/xi) + K0*t(z));
        k = k + 1;                                           % Increment iteration
    end                                                  % End while loop
    [I(z),I_low] = deal(y(k));                           % I(z) is equal to root
end                                                  % End for loop
i = (1 + xi./I)*K0;                                  % Infiltration rate
```

The for loop cycles over each element of the time vector. For each time, `t(z)`, the while loop executes repeatedly Newton's method in Equation (C.3) using as initial guess, `y(1)`. Newton's method keeps iterating until the solution, `y(k)`, satisfies `rtol`, the desired tolerance on the function value of the root. To prevent an infinite while loop and promote computational efficiency, Newton's method cannot exceed `kmax` iterations. The values of `rtol` and `kmax` may be specified by the user as optional input arguments of the `Green_Ampt` function. If omitted in the function call, default values of `rtol = 1e-12` and `kmax = 20` will be assigned.

The `Green_Ampt` function returns as second output argument a $n \times 1$ vector of infiltration rates, i (LT^{-1}), for each t. The infiltration rate must satisfy the following relationship

$$i = \frac{\partial I}{\partial t} = -\frac{\partial r/\partial t}{\partial r/\partial I}. \tag{C.7}$$

The denominator, $\partial r/\partial I$, is given by Equation (C.4) and the numerator of Equation (C.7) may be derived by differentiating the residual function, $r(I, t, \xi, K_0)$, with respect to t to yield

$$\frac{\partial r}{\partial t} = -K_0. \tag{C.8}$$

If we enter Equations (C.8) and (C.4) into Equation (C.7), then we yield the following expression for the infiltration rate

$$i = \frac{\partial I}{\partial t} = \frac{(1 + I/\xi)K_0}{I/\xi} = \left(1 + \frac{\xi}{I}\right)K_0, \tag{C.9}$$

in units of length per time. This concludes the numerical solution of Equation (C.1).

APPENDIX D: REFERENCE SET USING GREEN AND AMPT INFILTRATION MODEL

This Appendix describes how we created the reference infiltration data set for the Green and Ampt (1911) infiltration model. This model describes cumulative vertical infiltration,

I (L), using the following equation (from Appendix C)

$$I(t) = \xi \log\left(1 + \frac{I(t)}{\xi}\right) + K_0 t \tag{D.1}$$

where the scalar, $\xi = \Delta h \Delta\theta$ (L) and hydraulic conductivity of the transmission (wetted) zone, K_0 (LT^{-1}) are treated as unknown coefficients. As is evident from the above equation, the Green and Ampt (1911) model does not have valid basis functions. As a result, we must resort to nonlinear least squares to estimate the coefficients, $\xi > 0$ and $K_0 > 0$, of Equation (D.1) for each measured infiltration record of the SWIG database.

We use the Levenberg–Marquardt (LM) algorithm (Levenberg, 1944; Marquardt, 1963) to find the minimum sum of squares of the cumulative infiltration residuals. In Vrugt and Gao (2022) we use the LM algorithm for estimation of the unknown coefficients of the infiltration equation of Parlange et al. (1982). This implementation assumes an analytic Jacobian matrix and uses proactive boundary control to enforce the parameter boundaries. This exact same implementation is used to estimate the parameters, $\xi > 0$ and $K_0 > 0$, of the Green and Ampt (1911) model in Equation (D.1). We suffice to say here that the LM algorithm belongs to the class of gradient-based local optimization methods and uses information about the local shape of the sum of squares residuals objective function to determine the most productive downhill search direction in pursuit of the least squares estimates of ξ and K_0.

Once the two coefficients are known, we simulate the least squares $I(t)$ relationship of Equation (D.1) with a high temporal resolution until $I(t)$ reaches the maximum measured cumulative infiltration, $\tilde{I}_{\max}$. Next, we use linear interpolation to determine the successive times at which the simulated cumulative infiltration equals, $\tilde{I}_{\max}/j$, where $j = (1, \ldots, n)$. We replace the measured infiltration data of each SWIG experiment with $n = 100$ different (t, I) data pairs of the least squares $I(t)$ relationship of Equation (D.1). This data set then serves as our second reference infiltration set.

What is left to do is to determine the partial derivatives of Equation (D.1) with respect to its two parameters, ξ and K_0. The values of these two partial derivatives at the different measurement times, $\tilde{\mathbf{t}} = (\tilde{t}_1 \ \ldots \ \tilde{t}_n)^\top$, make up the columns of the Jacobian matrix. This matrix determines the search direction of the LM algorithm. In the derivation below, we use the function, $f(t, \xi, K_0)$, for the Green and Ampt (1911) infiltration model in Equation (C.1).

To determine the partial derivatives of the Green and Ampt (1911) infiltration model, we write Equation (C.1) in the familiar residual form (see also Appendix C)

$$r(I, t, \xi, K_0) = I - \xi \log\left(1 + \frac{I}{\xi}\right) - K_0 t. \tag{D.2}$$

We can take advantage of the chain rule to write the partial derivative of $r(I, t, \xi, K_0)$ with respect to ξ as follows

$$\frac{\partial r}{\partial \xi} = \frac{\partial r}{\partial I}\frac{\partial I}{\partial \xi} + \frac{\partial r}{\partial \xi}\frac{\partial \xi}{\partial \xi} + \frac{\partial r}{\partial K_0}\frac{\partial K_0}{\partial \xi} + \frac{\partial r}{\partial t}\frac{\partial t}{\partial \xi} = 0. \tag{D.3}$$

As the two coefficients, ξ and K_0, are independent, $\partial K_0/\partial \xi = 0$, and time invariant, $\partial t/\partial \xi = 0$, the above expression simplifies to

$$\begin{aligned} 0 &= \frac{\partial r}{\partial I}\frac{\partial f(t, \xi, K_0)}{\partial \xi} + \frac{\partial r}{\partial \xi} \cdot 1 + \frac{\partial r}{\partial K_0} \cdot 0 + \frac{\partial r}{\partial t} \cdot 0 \\ &= \frac{\partial r}{\partial I}\frac{\partial f(t, \xi, K_0)}{\partial \xi} + \frac{\partial r}{\partial \xi}. \end{aligned} \tag{D.4}$$

Thus, we yield the following analytic expression for the partial derivative of the Green and Ampt model, $\partial f(t, \xi, K_0)$, with respect to ξ

$$\frac{\partial f(t, \xi, K_0)}{\partial \xi} = -\frac{\partial r/\partial \xi}{\partial r/\partial I}. \tag{D.5}$$

We can differentiate the residual function, $r(I, t, \xi, K_0)$, with respect to ξ and I. Symbolic differentiation with respect to ξ results in the following expression

$$\frac{\partial r}{\partial \xi} = \frac{I}{\xi(1 + I/\xi)} - \log\left(1 + \frac{I}{\xi}\right). \tag{D.6}$$

If δ is the dimensionless input variable of the natural logarithm

$$\delta = 1 + \frac{I}{\xi}, \tag{D.7}$$

the expression for $\partial r/\partial \xi$ simplifies to

$$\frac{\partial r}{\partial \xi} = \frac{I}{\delta \xi} - \log(\delta) = \frac{I - \delta \xi \log(\delta)}{\delta \xi}. \tag{D.8}$$

The partial derivative of the residual function, $r(I, t, \xi, K_0)$, with respect to I equals (see Appendix C)

$$\frac{\partial r}{\partial I} = 1 - \frac{1}{\delta} = \frac{\delta - 1}{\delta}. \tag{D.9}$$

We can now substitute Equations (D.8) and (D.9) into (D.5) to yield the partial derivative of the Green and Ampt infiltration model, $f(t, \xi, K_0)$, with respect to ξ

$$\frac{\partial f(t, \xi, K_0)}{\partial \xi} = -\frac{\big(I - \delta \xi \log(\delta)\big)/\delta \xi}{(\delta - 1)/\delta} = \frac{I - \delta \xi \log(\delta)}{(1 - \delta)\xi}. \tag{D.10}$$

We can follow a similar derivation for the second coefficient, K_0, of the Green and Ampt function. Application of the

chain rule to Equation (D.2) produces the following identity

$$\frac{\partial r}{\partial K_0} = \frac{\partial r}{\partial I}\frac{\partial I}{\partial K_0} + \frac{\partial r}{\partial K_0}\frac{\partial K_0}{\partial K_0} + \frac{\partial r}{\partial \xi}\frac{\partial \xi}{\partial K_0} + \frac{\partial r}{\partial t}\frac{\partial t}{\partial K_0} = 0, \tag{D.11}$$

which simplifies to the following expression

$$\frac{\partial r}{\partial K_0} = \frac{\partial r}{\partial I}\frac{\partial f(t,\xi,K_0)}{\partial K_0} + \frac{\partial r}{\partial K_0} = 0 \Rightarrow \frac{\partial f(t,\xi,K_0)}{\partial K_0} = -\frac{\partial r/\partial K_0}{\partial r/\partial I}. \tag{D.12}$$

The numerator of Equation (D.12) may be derived from symbolic differentiation of the residual function, $r(I,t,\xi,K_0)$, with respect to K_0. We yield

$$\frac{\partial r}{\partial K_0} = -t. \tag{D.13}$$

We can now enter the partial derivatives in Equations (D.12) and (D.9) in the numerator and denominator of Equation (D.12). This results in the following expression for the partial derivative of the Green and Ampt infiltration model, $f(t,\xi,K_0)$, with respect to K_0

$$\frac{\partial f(t,\xi,K_0)}{\partial K_0} = \frac{\delta t}{\delta - 1}. \tag{D.14}$$

This concludes our derivation of the partial derivatives of the Green and Ampt (1911) model.

APPENDIX E: SIMULATED INFILTRATION CURVES FOR GREEN-AMPT DATA SET

This Appendix (Figure E.1) displays the simulated infiltration curves of Philip's two-term expression for the full and partial infiltration experiments of the Green-Ampt data set.

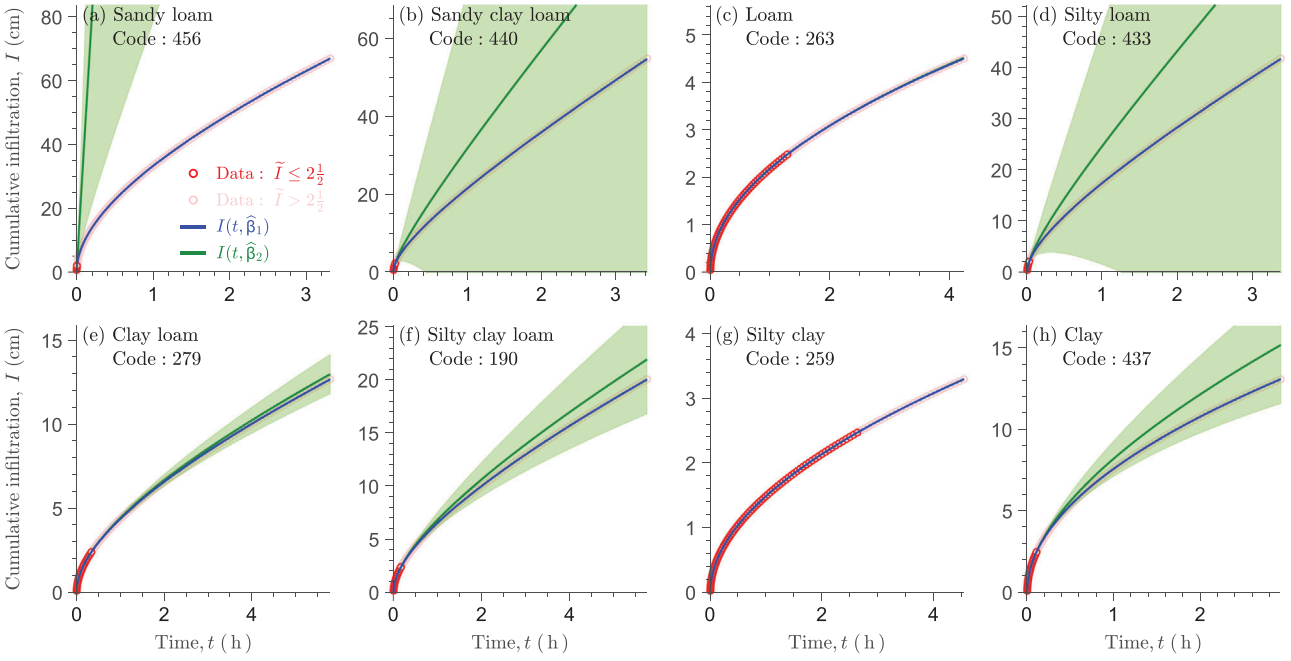

FIGURE E.1 Green and Ampt infiltration data set: Observed (red dots) and simulated cumulative infiltration curves for the selection of eight SWIG soils using the least squares values of S, K_s, and c (Table 1) of Philip's two-term infiltration equation for the full (solid blue line) and partial (solid green line) experiments. Each graph presents a different soil type, specifically (a) sandy loam, (b) sandy clay loam, (c) loam, (d) silty loam, (e) clay loam, (f) silty clay loam, (g) silty clay and (h) clay. Measured data beyond the time validity of Philip's two-term equation are displayed with a tint of red. The green region displays the 95% confidence intervals of the partial experiments..