Using a Neural Network – Physics-based Hybrid Model to Predict Soil Reaction Fronts

Tao Wen[†,1,*], Chacha Chen[†,2,6], Guanjie Zheng[2], Joel Bandstra[3], Susan L. Brantley[4,5]

[1]Department of Earth and Environmental Sciences, Syracuse University, Syracuse NY 13244
[2]College of Information, Science and Technology, Pennsylvania State University, University Park PA 16802
[3]Department of Mathematics, Engineering, and Computer Science, Saint Francis University, Loretto, PA 15940
[4]Earth and Environmental Systems Institute, Pennsylvania State University, University Park PA 16802
[5]Department of Geosciences, Pennsylvania State University, University Park PA 16802
[6]Now at Department of Computer Science, The University of Chicago, Chicago, IL 60637

[†]These authors contributed equally to this work

***Corresponding author:**
Tao Wen, Department of Earth and Environmental Sciences, Syracuse University, Syracuse, New York 13244, United States
Phone: 734-730-8814; E-mail: twen08@syr.edu

CRediT authorship contribution statement

Tao Wen: Supervision, Conceptualization, Investigation, Writing – original draft, Writing – review & editing. **Chacha Chen:** Conceptualization, Investigation, Writing – original draft, Writing – review & editing. **Guanjie Zheng:** Conceptualization, Investigation, Writing - review & editing. **Joel Bandstra:** Supervision, Conceptualization, Investigation, Writing - review & editing. **Susan L. Brantley:** Funding acquisition, Supervision, Conceptualization, Investigation, Writing – review & editing.

Abstract

Analytical and numerical solutions have been proposed to model reaction fronts to study soil formation. With growing access to large geo-datasets and powerful computational capacity, data-driven models are becoming increasingly useful. We therefore explored the use of a neural network (NN) guided by a physics-based model (PBM) to simulate the depth profile of feldspar dissolution in soils. Specifically, we explored this hybrid neural network (HNN) to see if it could predict reaction fronts as a function of important variables known from domain knowledge: site climate characteristics (temperature T; precipitation P), geomorphic parameters (soil residence time t; erosion rate E), and parent material mineralogy (quartz content Q; albitic feldspar content of the feldspar A). We evaluated the mean square error (MSE) for 63 HNNs, each using a different combination of training data (i.e., soil profiles) and environmental variables. The HNNs trained to four or five soil profiles that used a subset of t, T, Q, E, and A as predictor variables yielded lower MSEs than the PBM, and showed globally convergence. At least two variables are needed to achieve an MSE within 1% of the corresponding PBM. The HNNs generally predicted the slope better than the depth of the front because the PBM was not used to predict depth. HNN results identify t and P as the most and least useful variable in predicting the reaction front, respectively. This is the first time a NN was hybridized to a PBM to simulate reactions in soils. As part of this effort, we developed a tool to identify cases which have converged to a global solution, and cases which present local solutions. The approach shows promise for future efforts but should be applied to larger sets of soil profile data and PBMs that predict both the depth and slope of reaction fronts.

Keywords

Soil reaction front, neural network, hybrid neural network, physics-based model

Highlights

1. We developed a physics-informed neural network to predict slopes of reaction fronts
2. We developed a tool to identify neural network models presenting local solutions
3. Residence time is the most useful variable in predicting the slope of a soil reaction front

1. Introduction

Recent improvements in data-driven modeling have opened up new avenues to assess Earth and environmental science data (Bergen et al., 2019; Shen et al., 2018; Shen, 2018). However, this approach has yet to show utility in identifying natural laws from observational data (Schmidt and Lipson, 2009) and is generally not used in hypothesis testing (Shen, 2018). One promising new avenue is the integration of data-driven techniques with physics-based models (PBM).

In soils, dissolution reaction fronts are localized zones of weathering in areas of generally downward-flowing water that appear as depth intervals where one mineral dissolves and is removed from the soil as a solute while another mineral may be precipitated (e.g., Brantley and White, 2009; Lichtner, 1988). Reaction fronts can reveal the flow of meteoric water in the subsurface over geologic time periods (Brantley et al., 2017) and may indicate geological rates of CO_2 removal from the atmosphere by natural long-term processes (Godderis et al., 2019). Researchers have developed and adopted a variety of PBMs, e.g., reactive transport models (RTM), to reconstruct reaction fronts in soils (e.g., Godderis et al., 2019; Li et al., 2017; Maher and Navarre-Sitchler, 2019; White et al., 2001). These models require RTM codes that treat transport of solutes in flowing water (advection), transport of solutes through the water in the soil and rock pores (diffusion), and chemical reactions between solutes and the soil grains (e.g., Lebedeva et al., 2007; Li et al., 2017; Lichtner et al., 1996). Although RTMs are becoming increasingly useful to predict changes in aqueous and solid-phase chemistry over space and time, they are often difficult to parameterize because of the lack of data for and prior knowledge of the environmental, thermodynamic, and kinetic conditions of the associated system (e.g., Moore et al., 2012). Researchers therefore generally rely on uncertainty propagation, sensitivity testing (both local and global) of parameters, and calibration to evaluate the importance of the many

84 required but generally unconstrained or loosely constrained parameters (Laloy and Jacques,
85 2019). To ascertain the uncertainty of the model prediction accurately can require thousands to
86 tens of thousands of simulations.

87 Several published papers have explored the use of neural networks (NN) to simulate aspects
88 of soil-water evolution. One such recent effort was to determine dispersivity and retardation
89 factors of solutes in waters infiltrating soils (Mojid et al., 2019). Another team trained a NN to a
90 reactive transport simulation of a geochemical system considered at microscopic scale and then
91 successfully used it within a model of a macroscopic system (Prasianakis et al., 2020). Another
92 research effort attempted to use a NN as an emulator model in comparison to a complete reactive
93 transport model (Laloy and Jacques, 2019).

94 Here, we focus on solid-phase chemical and mineralogical profiles in regolith over space
95 and time. These solid-phase datasets document the long-time interaction of soil materials and
96 meteoric waters. Generally when using reactive transport models to simulate the changes in soil
97 materials during weathering, we often start with a known depth profile of soil/rock composition,
98 the change of which over long time periods due to weathering is dictated by infiltration and
99 percolation of recharge water through the soil profile (e.g., Moore et al., 2012). In this study, we
100 did not seek to simulate this full complexity of processes. Rather, we tried to take a small first
101 step in developing a NN for use in predicting reaction fronts in soils. We sought specifically to
102 minimize model discrepancy between a PBM that had been previously proposed to describe a
103 single reaction front in a soil column (Brantley et al., 2008) and predictions from the use of NNs.
104 Note, this simplified PBM is ultimately a sigmoidal function fitted to the sodium concentration
105 in solid-phase soil across depth. We attempted to develop a hybrid neural network (HNN) that
106 incorporates the NN and this PBM to explore whether NNs can be used to improve the accuracy

of physics-based modeling results and whether we can extrapolate findings from one location to another. We trained the HNN with soil data included in the training data and then tested the trained model against data from other soil profiles contained in the testing data. We sought to determine if the HNN was able to (1) simulate soil profiles that were not part of the training data but had formed in a similar climate and erosional regime like the soils in the training data; (2) simulate soil profiles that were not part of the training data and had formed in different climates; and (3) simulate soil profiles that were not part of the training data and had formed in different erosional regimes. Our approach differs from the previous NN efforts in that we focus on the solid-phase chemistry in water-rock reaction systems which represent the integrated results of weathering processes over a longer timescale compared to the solute chemistry. In addition, unlike the PBM which requires domain knowledge of every single simulated soil profile for the model parameterization, our proposed HNN can potentially extrapolate the knowledge (i.e., trained HNN) learned from one or more soil profiles to the other (i.e., predictions using the trained HNN). We anticipate that such HNNs could eventually become of great help to simulate soil profiles without prior domain knowledge.

2. Theoretical Background

2.1 Soil profiles

We first compiled soil chemistry data from a few soil profiles that have formed as meteoric water infiltrates parent material over geologically long periods of weathering (e.g., Brantley and Lebedeva, 2011). We model the soils as one-dimensional systems where the parent material (rock or sediment) is exposed and can be affected by erosional processes that may remove material from the land surface while meteoric water infiltrates, on net, unidirectionally

downward. Over time, minerals dissolve and precipitate in the soil column. Therefore, the abundances of different minerals vary with depth over time. Multiple minerals that have different solubility exist in soil and rock. For example, rocks commonly contain quartz (mostly insoluble) plus feldspar (soluble). Thus, in many soils, feldspar weathers to precipitate a clay such as kaolinite (insoluble) while the quartz remains unweathered, propping open the porosity and holding the overall volume of the weathering material constant (isovolumetric weathering). Feldspar contains the common element sodium (Na), but kaolinite does not contain Na. Therefore, weathering of a column of a rock (that includes feldspar + quartz) yields a column with decreasing concentrations of Na near the land surface. Na is removed from the soil layers as kaolinite is precipitated until the Na concentration (and by inference, feldspar) reaches zero at the land surface. At that point, the soil at the land surface is strictly quartz + kaolinite. The depth interval over which Na concentration varies from zero to the abundance in the parent material is called the interval of the reaction front. In some soils, the reaction front can be overlain by a depth interval that lacks feldspar, if the reaction front has advanced and no longer truncates the land surface.

As rock material weathers to soil, it can lose mass not only by (bio)geochemical weathering (e.g., dissolution) but also at the land surface, by physical losses known as erosion (Riebe et al., 2016). While weathering of a column of a rock (that includes feldspar + quartz) yields a column with decreasing concentrations of Na towards the land surface, erosion can remove the topsoil. Regardless of whether a soil is eroding or not eroding, the depth profile of Na concentration can generally be used to indicate the feldspar reaction front. More details of soil weathering can be found in the supplementary text (“Theoretical background of Soil Weathering”).

153 2.2 Physics-based modeling of soil profiles

154 A full implementation of a reactive transport model to simulate soil weathering requires
155 many parameters that are often unknown. Moreover, the choice of those parameter values to
156 ensure a successful model simulation is often non-unique. This property of equifinality is
157 common to a wide variety of Earth and environmental science models (Shen, 2018) including
158 hydrologic models (Beven and Freer, 2001), models of global biogeochemical cycling (Tang and
159 Zhuang, 2008), reaction kinetics (Bandstra and Tratnyek, 2004), and water quality models
160 (Schulz et al., 1999). In many of these cases it is possible to formulate relatively simple
161 empirical relations that describe biogeochemical processes on an average basis over sufficiently
162 large spatial or temporal scales (Savenije, 2001). Such models have the advantage of being
163 parsimonious and capable of describing the relevant phenomena under a broad range of
164 environmental conditions but with the drawback of parameter values that are *sui generis* and that
165 must be estimated from data (e.g., by regression analysis). This modeling approach is referred to
166 as physics-based modeling or PBM as illustrated in Figure 1a.

167 Brantley et al. (2008) proposed a parsimonious PBM for soil profiles that accounts for the
168 opening of reactive surface area as the parent material first begins to weather leading to eventual
169 depletion of non-conservative elements (such as Na) as the weathering process moves toward
170 completion. Brantley et al. (2008) showed that under a generalized set of assumptions, the soil
171 profile can be modeled as the reaction front of an autocatalytic process, i.e., a sigmoidal curve of
172 the form:

173 $$C = \frac{C_0}{\frac{C_0 - C_{x=0}}{C_{x=0}} \exp(\Gamma_{ini} \cdot \hat{k} \cdot x) + 1} \quad (1)$$

174 where C is the concentration, or abundance (mol/m^3), of an element such as Na in a soil profile at
175 a given depth x below the land surface, Γ_{ini} describes the roughness of the surface area of the

176 dissolving mineral, and $\hat{k}$ is a lumped kinetic parameter that describes the reactivity of the
177 mineral and its initial specific surface area. $C_{x=0}$ is the concentration (abundance) of the element
178 at the land surface (x = 0), or, for a reaction front that has advanced into the subsurface, at the
179 top of the front. C_0 is the concentration of the element in the underlying (unweathered) parent
180 material.

181 Eqn. (1), when applied as a non-linear regression equation, suitably describes soil profiles
182 developed on several different parent materials, under different climate conditions, and with
183 different geomorphic time-scales (Brantley et al., 2008). Each soil profile, however, requires a
184 unique set of parameter values and, therefore, Eqn. (1) lacks predictive power for soil profiles
185 developed under novel conditions. To overcome this limitation, we sought to leverage the power
186 of NN to develop an HNN that could similarly simulate the depth profile of Na content in the soil
187 after long duration weathering. Although this PBM is not nearly as powerful as an RTM, our
188 attempt to develop an HNN using this PBM represents a first step in the direction of ultimately
189 hybridizing with RTMs. Furthermore, the equation implicitly reveals the reaction front thickness
190 – an observable that can provide information related to the advective Damköhler number for the
191 reactive transport system in some soils (Brantley and Lebedeva, 2021).

192

193 2.3 Hybridization of the physics-based model with an artificial neural network

194 Artificial neural networks (ANNs) represent a group of widely-used machine learning
195 models (Figure 1b) that can show superior performance in regression fitting problems
196 (Goodfellow et al., 2016; LeCun et al., 2015). Typically, these models are described as a stack of
197 layers consisting of an input layer, at least one hidden layer of neurons, and an output layer. The

input layer corresponds to the independent (or predictor) variable(s) in the regression analysis while the output layer corresponds to the dependent (or target) variable.

Each hidden layer in ANNs consists of several neurons and the input to each neuron is a linear combination of the outputs from the previous layer as specified by a vector of weights W and a bias b, i.e., activation function [Eqn. (2)]. Note that the activation function could be set as non-linear (e.g., sigmoid).

$$h_{i,j} = \sigma\left(\overrightarrow{W}_{i,j} \cdot \vec{h}_{i-1} + b_{i,j}\right) \quad (2)$$

Here, i denotes the *i-th* layer and j denotes the *j-th* neuron within the layer. For the *j-th* neuron in the *i-th* layer ($\vec{h}_{i,j}$), the output of this neuron is the bias scalar $b_{i,j}$ plus the weighted sum of input neurons from the previous layer, $\vec{h}_{i-1}$ (the vector of weights is $\overrightarrow{W}_{i,j}$). The vector length in Eqn. (2) is determined by the previous layer while the maximum of indices i and j are hyperparameters defined by the user. ANN containing multiple hidden layers is able to approximate arbitrarily complicated functions (Cybenko, 1989; Hornik, 1991), a property known as universality.

Although universality is a useful property when it comes to fitting variable phenomena, it is a limiting factor for their direct use in scientific discovery (Brouwer et al., 2014; Livingstone et al., 1997). In a manner analogous to equifinality in PBMs, universality implies that many ANNs can be constructed to adequately capture the state of knowledge about a natural phenomenon. While this does not limit the use of ANNs in forecasting, it does limit the use of ANNs for extracting principles or testing hypotheses. To overcome this limitation, NNs are now being hybridized with PBMs to learn new scientific concepts from observational data (Karpatne et al., 2017).

An HNN (Figure 1c) combines a PBM and a NN. Different from the NN itself, the output of the HNN is still governed by the physical formula. In the HNN, a NN is used to find the optimal

values of parameters (e.g., environmental variables) in the formula by modeling and minimizing the mismatch between prediction results from the PBM and the observations. For the example here, the HNN predicts the term, $\Gamma_{ini} \cdot \hat{k}$, in Eqn. (1) by using site climate characteristics, geomorphic rates, and parent mineral composition as the input layer to an ANN. Specifically, we assess how the term, $\Gamma_{ini} \cdot \hat{k}$, varies with the residence time and quartz content of the soil, the erosion rate, temperature and precipitation rate of the soil site, and the composition of the feldspar observed in the targeted reaction front.

The PBM was derived (Brantley et al., 2008) to describe a generic feldspar reaction front in soil (e.g., Brantley and White, 2009), in which the proposed formula, Eqn. (1), has only been used to fit individual soils. The parameter for surface area and kinetic reactivity [$\Gamma_{ini} \cdot \hat{k}$ in Eqn. (1)] is a function of not only soil depth x but also the residence time of particles in the soil and other factors as further discussed later. In this work, we explore how the value of $\Gamma_{ini} \cdot \hat{k}$, which is implicitly related to the advective Damköhler number for reactive transport in some soils (Brantley and Lebedeva, 2021), varies with environmental conditions and parent material characteristics. Specifically, previous works have shown that soil/rock weathering is affected by a set of soil-forming factors related to climate, biota, relief, parent material, and time (Dokuchaev, 1883; Jenny, 1941; Merrill, 1906). Thus, variations in the value of $\Gamma_{ini} \cdot \hat{k}$ could be affected by a combination of variables such as soil residence time or exposure time (t), mean annual temperature (T), erosion rate (E), and/or precipitation rate (P), and so on. Given that $\hat{k}$ includes the rate constant of feldspar dissolution, and this is known to vary with the albite content (A) of the dominant feldspar in the parent material (Blum and Stillings, 1995), we also sought to understand if the sodium (Na) content of the feldspar (equivalent to albite content) controls the characteristics of the reaction front. Finally, we also aimed to test the hypothesis that

the quartz content (Q) of the starting material might also affect the depth of the reaction front as suggested by reactive transport models (Brantley et al., 2017). These are all specific tests that we chose to pursue within our overall goal of exploring how to use a NN with a PBM for soil modelling.

To formulate the HNN, we defined $f(z) = \Gamma_{ini} \cdot \hat{k}$, where z represents any combination of predictor variables, e.g., $f(t)$, $f(t,T)$, $f(P,E)$. $f(z)$ is a NN consisting of one input layer, two fully connected hidden layers [Eqn. (2)], and one output layer [Eqn. (3)]. The activation function of the first hidden layer is a sigmoidal function [Eqn. (4)] while the activation function of the second hidden layer is the simple linear combination of all input neuron results. The output of the second fully connected layer is not determined by an activation function but rather by the PBM as shown in the equation below that is modified from Eqn. (1).

$$C' = \frac{1}{\frac{C_0 - C_{x=0}}{C_{x=0}} exp(f(z) \cdot x) + 1} \quad (3)$$

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad (4)$$

where we have normalized Eqn. (1) by the parent concentration C_0, i.e., $C' = \frac{C}{C_0}$. Eqn. (3) represents the HNN which incorporates a data-driven model into a physics model. The HNN is guided by the physics law that governs the PBM, while the lumped rate parameter, $\Gamma_{ini} \cdot \hat{k}$, is inferred from data using a NN.

3. Computational Methods

3.1 Soil datasets

Seven soil profiles were compiled from four sites in the U.S.: Santa Cruz (California; four soil profiles: SCT1,2,3,5; Figure 2), Davis Run (Virginia), Panola (Georgia), and Jughandle State

Natural Reserve (California) (Table 1 and S1). Except for the Jughandle soil profile, all other soil data were previously investigated (Brantley et al., 2017, 2008; Brimhall and Dietrich, 1987; Eckert et al., 2012; Maher et al., 2009; Masiello et al., 2004; Merritts and Bull, 1989; Merritts et al., 1991; Moore et al., 2012; Northup et al., 1995; Uroz et al., 2014; White et al., 2009, 2008, 2001). Detailed description of these soil profiles can be found in the supplementary text "Detailed Descriptions of Soil Datasets". Sodium concentrations in the solid phases of these soils from these seven profiles were normalized by the sodium concentration in the corresponding parent material (underlying unweathered material) before being fed into the neural network. Normalized sodium concentration mostly varied from 0 to 1. Predictor variables were not rescaled due to two primary reasons: 1) our NN is a two-layer dense neural network that should be capable of learning to scale the variables itself without the need of manual preprocessing (Hornik et al., 1989); 2) we also sought to derive the mathematical formula from the HNN that we can apply to calculate the lumped parameters of the PBM. The data preprocessing without the rescaling allows the direct derivation of such mathematical formula from the HNN.

3.2 Model setup and model training

Based on knowledge from soils research and modeling (Brantley et al., 2017; Lebedeva et al., 2010), we hypothesized that residence time (t), temperature (T), precipitation (P), erosion rate (E), quartz content (Q), and the albite content of the feldspar in the parent material (A) could explain differences in the reaction fronts in different soils. We thus used t, T, P, E, Q, and A as predictor variables in the HNN. Measurements of predictor variables were not scaled while Na concentrations were normalized by the Na concentration in the parent material. In total, an HNN was trained and constructed for each of the 63 different combinations of predictor variables. In

289 every case, C_x=0 and C_o (Table 1) were treated as known quantities. Full results of the training
290 and test phases for each of the 63 HNNs are summarized in Table S2. Although we did not
291 perform feature selection in this study since the selection of all the six features was already
292 parsimonious and the computational cost was not an issue for the proposed simple HNN, it might
293 be necessary to perform feature selection if the proposed HNN would be used to simulate other
294 soil profiles using a lot more features from the perspective of computational efficiency.

295 Models were trained with three different sets of soil profiles: (SCT1, SCT 2, SCT 3, SCT 5,
296 Panola) or (SCT1, SCT 2, SCT 3, SCT 5, Davis Run) or (SCT1, SCT 2, SCT 3, SCT 5) when
297 possible (Table S2). Models using these training sets were labelled a, b, and c, respectively. For
298 variables such as residence time t where every soil was characterized by a different value, i.e.,
299 the investigated soils all differed in estimated residence times, the model (e.g., in this case f(t))
300 was trained with all three sets. But for other variables such as T where subsets of soils displayed
301 the same value(s) for predictor variable(s) (e.g., SCT1, 2, 3, 5 all have the same temperature T),
302 models were not trained with all three sets of soils (a, b, c). For this example, f(T) was only
303 trained with a and b sets. Given this, only a subset models of f(t), f(Q), f(A), f(t,Q), f(t,A), f(Q,A),
304 and f(t,Q,A) were trained with training set c and then used to predict the other three non-
305 chronosequence (non-SCT) soils.

306 Optimized weights and biases were found for the NN component of the HNN by minimizing
307 mean square error (MSE) between the HNN output and the normalized measured Na
308 concentration data. MSE was chosen to measure the model performance because it reflects the
309 discrepancy between model prediction and observational results. The Adam method (Kingma
310 and Ba, 2014) was used for stochastic optimization with a learning rate of 0.001. The

optimization was terminated after 15,000 steps, upon which all our models converged (i.e., MSE stopped decreasing significantly).

The first hidden layer of 16 neurons was used with a sigmoid activation function. The second hidden layer has 1 neuron with linear activation. We performed a grid search to determine the best number of neurons based on MSE. For the number of neurons in the first layer, we tested 8, 16, 32 while 1 and 2 were tested for the hyperparameter of the number of neurons in the second layer. Another hyperparameter – the number of hidden layers – was also fined tuned by testing values of 1 and 2. Given the extremely small number of layers (i.e., two) of the NN and the small number of features (i.e., at most six), this simple structured neural network model is very unlikely to overfit. In particular, for those best-performing HNNs (see Section 4.2), each of them considers only two features, which renders it less likely for them to overfit.

3.3 Regression analysis with physics-based models

Parameters in the PBM – C_0, $C_{x=0}$, and $\Gamma_{ini} \cdot \hat{k}$ – were selected to fit each soil profile by minimizing the chi-squared statistic (χ^2). χ^2 is the sum-of-square errors between Eqn. (1) and the measured Na concentrations normalized by the variance of the residuals. For an individual soil profile, χ^2 is directly proportional to the MSE as discussed below. χ^2 was minimized using the Levenberg-Marquardt algorithm (as implemented in Igor Pro 8 from Wavemetrics Inc.) with a termination criterion of nine iterations with no more than a 0.1% decrease in χ^2 (Press et al., 2007). Initial guesses for the parameters were determined by visually estimating (1) the concentration of Na in the parent material, (2) the slope of the soil profile at its inflection point, and (3) the depth of the inflection point. The initial guess for C_0 was taken directly from the estimated parent concentration. Initial guesses for the other two parameters were calculated as:

$$\Gamma_{ini}k \approx \frac{4}{C_0}\left|\frac{dC}{dx}\right|_{IP} \quad (5)$$

$$C_{x=0} \approx \frac{C_0}{1+exp(\Gamma_{ini}k|x|_{IP})} \quad (6)$$

where the subscript *IP* denotes values estimated at the soil profile inflection point.

In several cases, we followed standard practice and excluded a few near-surface concentration data points from the fits owing to apparent exogenous disturbance (i.e., dust inputs). These near-surface soils were perturbed by eolian input and bioturbation, which interfered with the pristine signatures of soil development over time (Brantley et al., 2008). Data excluded from fitting were the first 2 meters of the ~10-meter-deep Panola granite profile, the first 5 meters of the ~22-meter-deep Davis Run profile, and the first meter of the ~6-meter deep Jughandle profile.

4. Results and Discussion

4.1 Training phase of the hybrid neural network

MSE values for the HNN results for all seven soil profiles ranged from 0.0030 to 0.38 with an average MSE of 0.072 (Table S2). Below, we compare the HNN MSE values with PBM MSE values by calculating the percent difference between these two MSE values as follows:

$$\text{Percent Difference in MSE} = \frac{MSE_{\text{HNN}} - MSE_{\text{PBM}}}{MSE_{\text{HNN}}} \quad (7)$$

In some cases, the HNN was able to achieve slightly lower MSE than the PBM (see Table S2) but in all such cases, the difference in MSE was likely due to rounding of the PBM parameters as previously reported (Brantley et al., 2008). For visualization purposes, these negative percent differences were set to 0.1% in the Figures 3, 4, and 6. This value effectively identifies these cases as ones where the HNN was able to precisely reproduce the PBM.

356 In other cases, the HNN underperformed the PBM. This could be due to HNNs that
357 converged to local minima in MSE instead of the global minima or it could be due to an
358 inadequate set of predictor variables. A methodology for identifying cases of local convergence
359 is developed in Section 4.1.1. In Sections 4.1.2 and 4.1.3 we assess how the number of predictor
360 variables controlled the prediction accuracy of HNNs as well as which predictor variable(s)
361 might be the most important environmental factor dictating the development of reaction front in
362 the soil profile.
363
364 *4.1.1 Local convergence in the hybrid neural network*
365 The percent difference in MSE between HNNs and their corresponding PBMs in the training
366 phase was calculated by following Eqn. (7) and is plotted as a histogram in Figure 3. A total of
367 26 HNNs (out of 63) reported a MSE less than 10% of that of the corresponding PBM, indicating
368 a comparable prediction performance between the HNN and PBM in the training phase for these
369 26 HNNs.
370 For any given HNN we define a subset HNN as one that was trained with the same data but
371 using only a subset of the same predictor variables. For instance, HNN 22a was trained on
372 SCT1,2,3,5 and Panola data using t, T, and P as predictor variables. HNN 7a was trained on
373 those same data with t and T as predictor variables. Therefore, in our terminology, HNN 7a is a
374 subset HNN to HNN 22a.
375 If an HNN and all of its subset HNNs have converged each to their global minimum in MSE,
376 then the parent HNN should perform no worse in terms of MSE than any of its subset HNNs. Of
377 course, the parent HNN could outperform one or more of its subset HNNs but the reverse cannot
378 occur. Because the HNN parameterization methods are stochastic, some allowance for slight and

379 insignificant outperformance must be made, but we observed a number of cases where an HNN
380 was outperformed by one of its subset HNNs by more than 10% (i.e., the MSE for the parent was
381 greater than 1.1 times the MSE for the subset HNN), which are defined as locally-converged
382 HNNs. Otherwise, HNNs are named as globally-converged HNNs.

383 In Figure 3, those HNNs that were not significantly outperformed (at a level of 10%) by any
384 subset HNN have been shaded green. These HNNs have likely converged to a global minimum.
385 A characteristic feature of these apparently well converged HNNs is that they tended to use a
386 larger number of predictor variables (see Section 4.1.3). In contrast, we also found many HNNs
387 with MSE 10% larger than that of at least one subset HNN, suggesting convergence to a local
388 minimum in MSE, i.e., local convergence. This latter circumstance could lead to the conclusion
389 that additional predictor variables lowered the MSE of the HNN. In the following discussion, we
390 excluded all these locally-converged HNNs, and only considered those HNNs more likely to
391 have converged to the global minima (shaded green in Figure 3).

392

393 *4.1.2 Hybrid neural network and predictor variables*

394 As discussed in the text above, six predictive factors were considered: climate characteristics
395 (temperature T and precipitation P), geomorphic parameters (soil residence time or exposure
396 time t and erosion rate E), and parent material mineralogy (quartz abundance in the soil Q and
397 albite content in the feldspar A). We explored the relative importance of all the six predictor
398 variables in the HNN by assessing which predictor variable was most frequently included in the
399 best performing HNNs.

400 Figure 4 summarizes the number of HNNs (excluding those likely locally converged) that
401 includes each of the six predictor variables. Among those HNNs within 10% of the MSE

402 reported by the corresponding PBM, *t*, *T*, *P*, *E*, *Q*, and *A* were used in 16, 15, 1, 9, 14, and 12
403 HNNs, respectively. For those HNN with a percent difference in MSE of less than 0.1%, 14, 11,
404 1, 7, 11, and 11 HNNs included *t*, *T*, *P*, *E*, *Q*, and *A*, respectively.

405 These results suggested that soil residence time or exposure time (*t*) is the most useful
406 predictor variable for parameterizing an HNN that can reproduce the behavior of the PBM with
407 respect to slope of the reaction front. Temperature (*T*), erosion rate (*E*), and albite content of the
408 feldspar (*A*) are also helpful predictors while the precipitation (*P*) was the least useful predictor
409 variable.

410 The HNN ultimately predicted the parameter of $\Gamma_{ini} \cdot \hat{k}$ in the PBM (Eqn. 1) in a data-driven
411 approach, and yielded the derived value of $\Gamma_{ini} \cdot \hat{k}$ as the function – f(z). To further illustrate how
412 the derived f(z) varied with the soil residence time or exposure time (*t*), the most useful predictor
413 variable, and how it compared to the $\Gamma_{ini} \cdot \hat{k}$ in the PBM, we inspect the prediction results of
414 HNN model 1c f(*t*) (Figure 5). HNN model 1c was trained to soil profiles SCT1, 2, 3, and 5. As
415 shown in Figure 5, the derived f(z) value from the HNN and $\Gamma_{ini} \cdot \hat{k}$ from the corresponding
416 PBM were very similar. In addition, the derived f(z) varied significantly with the soil residence
417 time or exposure time. As the soil residence time or exposure time increased in the order of
418 SCT1, 2, 3, and 5, f(z) first decreased and then increased before reaching a plateau.

419

420 *4.1.3 Hybrid neural network and number of predictor variables*

421 We then explored how the number of predictor variables included in the HNN contributed to
422 the improvement of HNN performance (i.e., lowering of the MSE value).

423 Figure 6A plots the percent difference in MSE between the HNN and the corresponding
424 PBM as a function of the number of predictor variables included in the HNN. The improvement

425 in the MSE was significant from the HNN using one predictor variable to the HNN using two
426 predictor variables. At best, HNNs with only one predictor variable could only achieve MSE
427 values within 10% of the corresponding PBM. At least two predictor variables were needed to
428 construct an HNN that achieves an MSE within 1% of the corresponding PBM. Adding
429 additional predictor variables beyond two was not significantly advantageous.
430 Among those globally-converged HNNs, a total of 23 HNNs report MSE values within 10%
431 different from that of the corresponding PBM. Among these were seven HNNs that include two
432 predictor variables (Figure 6B and Table 2). These seven HNNs that use only two predictor
433 variables are defined as the best performing HNNs in this study, and are further discussed below.
434
435 4.2 Best performing hybrid neural network models
436 MSE values of the seven best performing HNNs in the training and test phases, and the MSE
437 values of the corresponding PBMs were listed in the Table 2. To reiterate, a best-performing
438 HNN is defined as (1) the HNN attained an MSE no more than 10% worse than any subset HNN
439 in the training phase and (2) the HNN attained an MSE value within 10% of the corresponding
440 PBM on the same training soil profiles.
441 The predictor variable of residence or exposure time of the soil (t) was included in all but
442 two of the best performing HNNs. This emphasizes again that residence or exposure time is the
443 most useful variable in predicting the slope of the reaction front using the HNN (see, for example,
444 Figure 5). Many researchers have previously emphasized the importance of residence time or
445 exposure time on reactivity (Washton et al., 2008; White and Brantley, 2003), and by inference,
446 on the slope or thickness of the reaction front. In contrast, precipitation is the least useful
447 predictor variable. This latter insight from the HNN work was surprising from a physics point of

448 view. The HNN is mostly predicting the slope of the reaction front (see discussion below), and
449 the slope is a function of the infiltration velocity of meteoric water into the soil, the reaction rate
450 constant, the mineral surface area available for dissolution, and the abundance of the reacting
451 mineral (Brantley and White, 2009). Apparently, although the precipitation P at the land surface
452 is a boundary condition for the infiltration velocity, the velocity at the depth of the reaction front
453 is more of a function of residence time or exposure time than the climate variable P, at least for
454 this subset of soils developed in moderately rainy conditions. This could be related to many
455 complicating factors related to evapotranspiration, biotic activity, plant succession, plant
456 physiology, soil nutrient content, and climate. In addition, as a soil profile develops, water that
457 initially infiltrates vertically and reacts with minerals begins to flow laterally, especially at
458 reaction fronts (Brantley and Lebedeva, 2021). With lateral flow, less downward-advecting water
459 percolates through the deepening reaction front as soil layers develop and differentiate. This in
460 turn affects the slope of the reaction front. The finding from the HNN is novel in that it
461 emphasizes the importance of t and de-emphasizes the importance of P in determining the slopes
462 of reaction fronts in field settings.

463 Among these seven HNNs, four models (i.e., 7a, 9a, 11a, 14a) were trained to SCT1,2,3,5
464 and Panola soil profiles ("a" series training set). Two models (i.e., 7b, 14b) were trained to
465 SCT1,2,3,5 and Davis soil profiles ("b" series training set) while one model (i.e., 11c) was
466 trained to only SCT1,2,3,5 soil profiles ("c" series training set). In particular, both f(t,T) and
467 f(T,Q) were able to train successfully with "a" and "b" series training sets (7a and 7b, and 14a
468 and 14b), while f(t,A) was able to train successfully with "a" and "c" series training sets (11a and
469 11c). Model f(t,T) was not successfully trained with the "c" series training set because the
470 temperature (T) did not vary across the soil profiles of SCT1,2,3,5 and f(t,T) therefore could not

471 be calculated for the "c" series training set. Due to the same reason, f(T,Q) was not trained with
472 the "c" series training set.
473 As for f(t,A), it was successfully trained with "a" and "c" series training sets but not with the
474 "b" series training set. A comparison of MSE values of HNN 11b and those HNNs using a subset
475 of predictor variables used by 11b indicated that HNN model 1b f(t) was 82% better than HNN
476 11b in terms of MSE values. Therefore, we inferred that HNN 11b was likely converged locally,
477 and it was excluded from the list of best performing HNNs.
478
479 4.3 Comparison of HNN Models to Validation Data
480 Trained HNNs were also applied to predict the reaction front for those soil profiles not
481 included in the training set. Reaction fronts predicted by the HNN and by the corresponding
482 PBM were plotted for all the seven best performing HNNs in Figures 7, 8, and 9. The depth
483 profiles of normalized measured Na concentrations are also shown for comparison. However,
484 given that HNNs were trained on some of the profiles, not all profiles are plotted for all HNN
485 predictions. For example, for HNNs trained with the "a" series training set, only soil profiles for
486 Davis Run and Jughandle were tested because Panola was used in training (Figure 7). Likewise,
487 for HNNs trained with the "b" series training set, soil profiles of Panola and Jughandle were
488 tested (Figure 8). For HNNs trained with the "c" series training set (trained only against SCT
489 profiles), we tested the trained model with soil profiles of Davis Run, Panola, and Jughandle
490 (Figure 9).
491 In general, trained HNNs can mostly predict the slope of the reaction front very well as
492 compared to the corresponding PBM. We did not expect the models to predict the depth of the
493 soil profile as well as the PBM. From Eqn. (5), we know that the initial guess for $\Gamma_{ini} \cdot \hat{k}$ in the

PBM depends on the slope of the reaction front. The HNN, which ultimately predicts $\Gamma_{ini} \cdot \hat{k}$ in a data-driven way, in essence, is able to find the slope of the soil profile which yields a quantification for the ratio of rate of dissolution of the albitic feldspar and the rate of transport. However, Eqn. (6) shows that the initial guess for $C_{x=0}$ depends on both $\Gamma_{ini} \cdot \hat{k}$ and the depth of the soil profile. So, since the HNN keeps the same $C_{x=0}$ value as the PBM while seeking potentially different values of $\Gamma_{ini} \cdot \hat{k}$, HNN's prediction for the depth of the soil profile is relatively inaccurate.

5. Conclusions

In this study we explored whether we could use hybrid neural networks (HNN), i.e., a neural network guided by a physics-based model (PBM), to predict concentration-depth data for reaction fronts in soils. We tested the idea that an HNN might teach us which variables that control soil formation are the most important in controlling reaction fronts. We specifically tested soil residence time (t), site climate characteristics [mean annual temperature (T), precipitation (P)], site erosion rate (E), quartz content (Q) of the parent material, and/or albite content (A) of the feldspar. For each of the 63 combinations of these environmental predictor variables, we trained and tested an HNN for comparison to a PBM for three different sets of soil training and test datasets. To seek the best performing HNN in terms of mimicking the prediction results from the corresponding PBM, we evaluated the percent difference in MSE (1) between each of the HNNs and any HNNs using the same subset of predictor variables in the training phase, and (2) between HNNs and their corresponding PBMs in the training phase. Among those best performing HNNs, soil residence time (t) was most frequently included as a predictor variable in the HNNs.

517 The HNN thus taught us that soil residence/exposure time is the most useful predictor
518 variable in terms of predicting the slope of a reaction front, an observable that is related in some
519 soils to the advective Damköhler number of the weathering system (e.g., Brantley and Lebedeva,
520 2021; von Blanckenburg et al., 2021). Surprisingly, precipitation was the least useful predictor
521 variable. We also explored how adding additional predictor variables might contribute to the
522 improvement of prediction accuracy of HNNs. The preliminary results showed that at least two
523 predictor variables were needed to achieve an MSE within 1% different of the corresponding
524 PBM. Trained best performing HNNs were also used to predict the reaction front for soil profiles
525 not included in the corresponding training dataset. Prediction results showed that HNNs can
526 generally mimic the PBM-simulated reaction front with respect to the front's slope but not with
527 respect to the depth of the front, a result that is at least partly related to the limitation of the
528 original PBM equation.

529 This is the first time an NN was incorporated into a PBM to develop an HNN to simulate the
530 depth profile of soil geochemistry. Although the proposed HNN consisted of only two layers and
531 a small number of predictor variables, the preliminary results of our work indeed pointed out a
532 promising direction for the application of NN in the even more complex PBM (e.g., high
533 resolution reactive transport models): replacing the sub-model of a PBM with a machine learning
534 model to either simplify the complicated parametrization that might be in high demand of
535 domain knowledge or to improve the computational efficiency as the machine learning models
536 usually run a lot faster than the corresponding physics sub-model (e.g., Prasianakis et al., 2020;
537 Reichstein et al., 2019). This requires that NN/HNN models be generalized to more soil profiles
538 (Mojid et al., 2019) and be improved to incorporate automation of NN/HNN training
539 (Prasianakis et al., 2020).

6. Acknowledgements

Funding for CC, SLB, TW, and ZL derived from National Science Foundation IIS-16-39150 to SLB and Z. Li. We acknowledge unpublished soil data from the Jughandle soil contributed by A.F. White and M. Schulz of the U.S. Geological Survey.

7. Computer Code Availability

All hybrid neural network models in this study are implemented using the TensorFlow 1.13.1 with the Python 3.5. Source codes of model development and application as well as the associated input data can be downloaded from the corresponding GitHub repository via this link: https://github.com/HANDS-Research-Group/HNN_Soil_Reaction_Front

References

Bandstra, J.Z., Tratnyek, P.G., 2004. Applicability of Single-Site Rate Equations for Reactions on Inhomogeneous Surfaces. Ind. Eng. Chem. Res. 43, 1615–1622. https://doi.org/10.1021/ie034250a

Bergen, K.J., Johnson, P.A., Hoop, M.V., Beroza, G.C., 2019. Machine learning for data-driven discovery in solid Earth geoscience. Science 363, 1299,. https://doi.org/10.1126/science.aau0323.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. Journal of Hydrology 249, 11–29.

Bierman, P., Gillespie, A., Caffee, M., Elmore, D., 1995. Estimating erosion rates and exposure ages with 36Cl produced by neutron activation. Geochimica et Cosmochimica Acta 59, 3779–3798. https://doi.org/10.1016/0016-7037(95)00267-4

Blum, A., Stillings, L.L., 1995. Feldspar Dissolution Kinetics, in: White, A.F., Brantley, S.L. (Eds.), Chemical Weathering Rates of Silicate Mineral: An Overview. Mineralogical Society of America. pp. 291–351.

Brantley, S.L., Bandstra, J., Moore, J., White, A.F., 2008. Modelling chemical depletion profiles in regolith. Geoderma 145, 494–504. https://doi.org/10.1016/j.geoderma.2008.02.010

Brantley, S.L., Lebedeva, M., 2011. Learning to read the chemistry of regolith to understand the Critical Zone. Annual Review Earth Planetary Science 39, 387–416.

Brantley, S.L., Lebedeva, M.I., 2021. Relating land surface, water table, and weathering fronts with a conceptual valve model for headwater catchments. Hydrological Processes 35. https://doi.org/10.1002/hyp.14010

Brantley, S.L., Lebedeva, M.I., Balashov, V.N., Singha, K., Sullivan, P.L., Stinchcomb, G., 2017. Toward a conceptual model relating chemical reaction fronts to water flow paths in hills. Geomorphology 277, 100–117. https://doi.org/10.1016/j.geomorph.2016.09.027

Brantley, S.L., White, A.F., 2009. Approaches to Modeling Weathered Regolith. Reviews in Mineralogy and Geochemistry 435–484.

Brimhall, G.H., Dietrich, W.E., 1987. Constitutive mass balance relations between chemical composition, volume, density, porosity, and strain in metasomatic hydrochemical systems: results on weathering and pedogenesis. Geochimica et Cosmochimica Acta 51, 567–587.

Brouwer, W.J., Kubicki, J.D., Sofo, J.O., Giles, C.L., 2014. An Investigation of Machine Learning Methods Applied to Structure Prediction in Condensed Matter.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. Mathematics of control, signals and systems 2, 303–314.

Dokuchaev, V.V., 1883. Russian Chernozem Selected Works of V.V.

Eckert, A.J., Shahi, H., Datwyler, S.L., Neale, D.B., 2012. Spatially variable natural selection and the divergence between parapatric subspecies of lodgepole pine (pinus contorta, pinaceae. American Journal of Botany 99, 1323–1334.

Godderis, Y., Schott, J., Brantley, S.L., 2019. Reactive transport models of weathering. Elements 1, 103-106,. https://doi.org/10.2138/gselements.15.2.103.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2, 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Hornik, K., 1991. Approximation capabilities of multilayer feedforward networks. Neural networks 4, 251–257.
Jenny, H., 1941. The Factors of Soil Formation. McGraw Hill, New York.
Karpatne, A., Atluri, G., Faghmous, J.H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N., Kumar, V., 2017. Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data. IEEE Trans. Knowl. Data Eng. 29, 2318–2331. https://doi.org/10.1109/TKDE.2017.2720168
Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization.
Laloy, E., Jacques, D., 2019. Emulation of CPU-demanding reactive transport models: a comparison of Gaussian processes, polynomial chaos expansion, and deep neural networks. Computational Geosciences 23, 1193–1215.
Lebedeva, M.I., Fletcher, R.C., Balashov, V.N., Brantley, S.L., 2007. A reactive diffusion model describing transformation of bedrock to saprolite. Chemical Geology 244, 624–645. https://doi.org/10.1016/j.chemgeo.2007.07.008
Lebedeva, M.I., Fletcher, R.C., Brantley, S.L., 2010. A mathematical model for steady-state regolith production at constant erosion rate. Earth Surf. Process. Landforms n/a-n/a. https://doi.org/10.1002/esp.1954
LeCun, Y., Bengio, Y., Hinton, G., 2015. . Nature 521, 436–444.
Li, L., Maher, K., Navarre-Sitchler, A., Druhan, J., Meile, C., Lawrence, C., Moore, J., Perdrial, J., Sullivan, P., Thompson, A., Jin, L., Bolton, E.W., Brantley, S.L., Dietrich, W.E., Mayer, K.U., Steefel, C.I., Valocchi, A., Zachara, J., Kocar, B., Mcintosh, J., Tutolo, B.M., Kumar, M., Sonnenthal, E., Bao, C., Beisman, J., 2017. Expanding the role of reactive transport models in critical zone processes. Earth-Science Reviews 165, 280–301. https://doi.org/10.1016/j.earscirev.2016.09.001
Lichtner, P.C., 1988. The quasi-stationary state approximation to coupled mass transport and fluid-rock interaction in a porous medium. Geochimica et Cosmochimica Acta 52, 143–165.
Lichtner, P.C., Steefel, C.I., Oelkers, E.H., 1996. Reactive Transport in Porous Media. Reviews in Mineralogy 34, 438.
Livingstone, D.J., Manallack, D.T., Tetko, I.V., 1997. Data modelling with neural networks: Advantages and limitations. Journal of Computer-Aided Molecular Design 8.
Maher, K., Navarre-Sitchler, A., 2019. Reactive Transport Processes that Drive Chemical Weathering: From Making Space for Water to Dismantling Continents. Reviews in Mineralogy & Geochemistry 349–380.
Maher, K., Steefel, C.I., White, A.F., Stonestrom, D.A., 2009. The role of reaction affinity and secondary minerals in regulating chemical weathering rates at the Santa Cruz Soil Chronosequence, California. Geochimica et Cosmochimica Acta 73, 2804–2831. https://doi.org/10.1016/j.gca.2009.01.030
Masiello, C.A., Chadwick, O.A., Southon, J., Torn, M.S., Harden, J.W., 2004. Weathering controls on mechanisms of carbon storage in grassland soils. Global Biogeochemical Cycles 18, 9.
Merrill, G.P., 1906. A Treatise on Rocks, Rock Weathering and Soils. MacMillian Inc, New York.
Merritts, D., Bull, W.B., 1989. Interpreting quaternary uplift rates at the mendocino triple junction, northern california, from uplifted marine terraces. Geology 17, 1020–1024.

643 Merritts, D.J., Chadwick, O.A., Hendricks, D.M., 1991. Rates and processes of soil evolution on
644 uplifted marine terraces, northern california. Geoderma 51, 241–275.
645 Mojid, M.A., Hossain, A.B.M.Z., Ashraf, M.A., 2019. Artificial neural network model to predict
646 transport parameters of reactive solutes from basic soil properties. Environmental
647 Pollution 255. https://doi.org/10.1016/j.envpol.2019.113355.
648 Moore, J., Lichtner, P.C., White, A.F., Brantley, S.L., 2012. Using a reactive transport model to
649 elucidate differences between laboratory and field dissolution rates in regolith.
650 Geochimica Cosmochimica Acta 93, 235–261.
651 Northup, R.R., Dahlgren, R.A., Yu, Z., 1995. Intraspecific variation of conifer phenolic
652 concentration on a marine terrace soil acidity gradient; a new interpretation. Plant Soil
653 171, 255–262. https://doi.org/10.1007/BF00010279
654 Pavich, M.J., 1986. Processes and Rates of Saprolite Production and Erosion on a Foliated
655 Granitic Rock of the Virginia Piedmont, in: Colman, S.M., Dethier, D.P. (Eds.), Rates of
656 Chemical Weathering of Rocks and Minerals. Academic Press, Orlando, FL, p. 603.
657 Pavich, M.J., Brown, L., Valette-Silver, J.N., Klein, J., Middleton, R., 1985. 10Be analysis of a
658 Quaternary weathering profile in the Virginia Piedmont. Geology 13, 39–41.
659 https://doi.org/10.1130/0091-7613(1985)13<39:BAOAQW>2.0.CO;2
660 Prasianakis, N.I., Haller, R., Mahrous, M., Poonoosamy, J., Pfingsten, W., Churakov, S.V., 2020.
661 Neural network based process coupling and parameter upscaling in reactive transport
662 simulations. Geochimica et Cosmochimica Acta 291, 126–143.
663 https://doi.org/10.1016/j.gca.2020.07.019
664 Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical recipes: The art
665 of scientific computing. Cambridge University Press.
666 Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat,
667 2019. Deep learning and process understanding for data-driven Earth system science.
668 Nature 566, 195–204. https://doi.org/10.1038/s41586-019-0912-1
669 Riebe, C.S., Hahm, W.J., Brantley, S.L., 2016. Controls on deep critical zone architecture: a
670 historical review and four testable hypotheses. Earth Surface Processes and Landforms 42,
671 128-156,. https://doi.org/10.1002/esp.4052.
672 Savenije, H.H.G., 2001. Equifinality, a blessing in disguise? Hydrological Processes 15, 2835–
673 2838.
674 Schmidt, M., Lipson, H., 2009. Distilling free-form natural laws form experimental data. Science
675 324, 81–85.
676 Schulz, K., Beven, K., H., B., 1999. Equifinality and the problem of robust calibration in
677 nitrogen budget simulations. Soil Science Society of America Journal 63, 1934–1941.
678 Shen, C., Laloy, E., Elshorbagy, A., Albert, A., Bales, J., Chang, F.-J., Ganguly, S., Hsu, K.-L.,
679 Kifer, D., Fang, Z., Fang, K., Li, D., Li, X., Tsai, W.-P., 2018. HESS Opinions:
680 Incubating deep-learning-powered hydrologic science advances as a community. Hydrol.
681 Earth Syst. Sci. 22, 5639–5656. https://doi.org/10.5194/hess-22-5639-2018
682 Shen, C.P., 2018. A Transdisciplinary Review of Deep Learning Research and Its Relevance for
683 Water Resources Scientists. Water Resources Research 54, 8558–8593.
684 Tang, J., Zhuang, Q., 2008. Equifinality in parameterization of process-based biogeochemistry
685 models: A significant uncertainty source to the estimation of regional carbon dynamics.
686 Journal of Geophysical Research: Biogeosciences 113.

Uroz, S., Tech, J.J., Sawaya, N.A., Frey-Klett, P., Leveau, J.H.J., 2014. Structure and function of bacterial communities in ageing soils: Insights from the Mendocino ecological staircase. Soil Biology & Biochemistry 69, 265–274.

von Blanckenburg, F., Schuessler, J.A., Bouchez, J., Frings, P.J., Uhlig, D., Oelze, M., Frick, D.A., Hewawasam, T., Dixon, J., Norton, K., 2021. Rock weathering and nutrient cycling along an erodosequence. Am J Sci 321, 1111–1163. https://doi.org/10.2475/08.2021.01

Washton, N.M., Brantley, S.L., Mueller, K.T., 2008. Probing the molecular-level control of aluminosilicate dissolution: A sensitive solid-state NMR proxy for reactive surface area. Geochimica et Cosmochimica Acta 72, 5949–5961. https://doi.org/10.1016/j.gca.2008.09.018

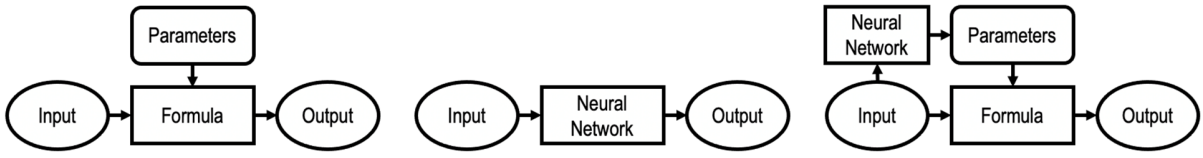
White, A.F., Brantley, S.L., 2003. The effect of time on the weathering of silicate minerals: why do weathering rates differ in the laboratory and field? Chemical Geology 202, 479–506. https://doi.org/10.1016/j.chemgeo.2003.03.001

White, A.F., Bullen, T.D., Schulz, M.S., Blum, A.E., Huntington, T.G., Peters, N.E., 2001. Differential rates of feldspar weathering in granitic regoliths. Geochimica et Cosmochimica Acta 65, 847–869. https://doi.org/10.1016/S0016-7037(00)00577-9

White, A.F., Schulz, M.S., Stonestrom, D.A., Vivit, D.V., Fitzpatrick, J., Bullen, T.D., Maher, K., Blum, A.E., 2009. Chemical weathering of a marine terrace chronosequence, Santa Cruz, California. Part II: Solute profiles, gradients and the comparisons of contemporary and long-term weathering rates. Geochimica et Cosmochimica Acta 73, 2769–2803. https://doi.org/10.1016/j.gca.2009.01.029

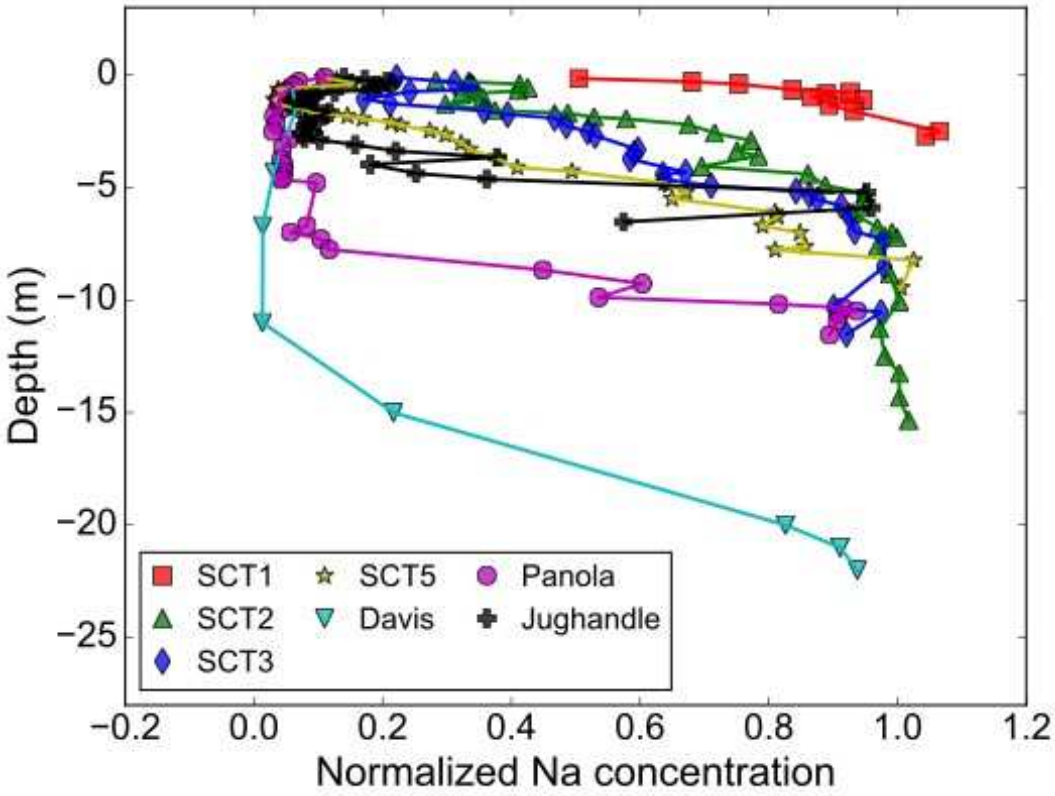
White, A.F., Schulz, M.S., Vivit, D.V., Blum, A.E., Stonestrom, D.A., Anderson, S.P., 2008. Chemical weathering of a marine terrace chronosequence, Santa Cruz, California I: Interpreting rates and controls based on soil concentration–depth profiles. Geochimica et Cosmochimica Acta 72, 36–68. https://doi.org/10.1016/j.gca.2007.08.029

Figures



(A) Physics-based model (PBM) (B) Neural network (NN) (C) Hybrid neural network (HNN)

Figure 1. Conceptual illustration for (A) physics-based model or PBM, (B) neural network or NN, and (C) hybrid neural network or HNN. The PBM develops an equation (central box labelled "formula") to predict output values (y) from input value(s) (x) for physical or chemical processes. This formula is based on physical observations and reflects physics-based laws. Thus, a PBM provides an equation with parameters where the parameters (here, $\Gamma\_ini \cdot \hat{k}$) in the formula are thought to be explained by known observable phenomena. These parameters may change as a function of properties of the environment where the data samples are collected. In contrast, the NN directly learns the function that links the input and output. A HNN uses both the PBM and the NN to learn the appropriate parameters.



726
727 **Figure 2.** Depth profiles of normalized Na concentration in the seven soil profiles discussed in
728 this study. All Na concentrations are divided by that of the corresponding parent material.
729

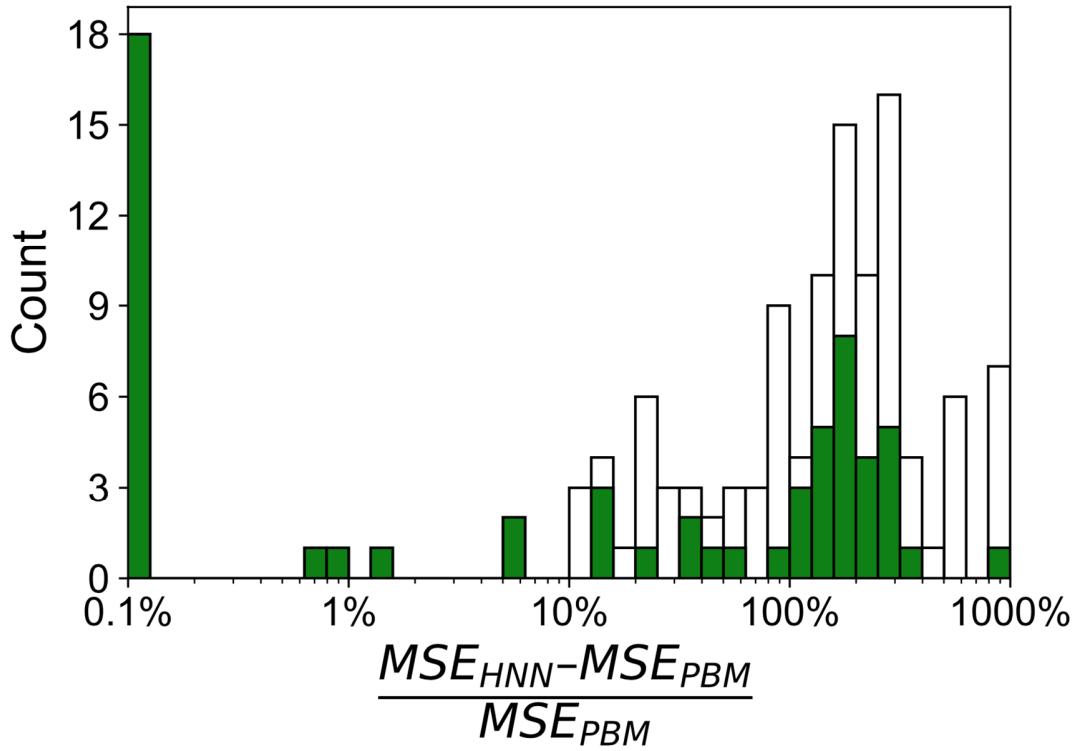


Figure 3. Histogram of percent difference in MSE between HNNs and their corresponding PBMs in the training phase. All negative percent difference values were set to 0.1% for visualization purposes. The unshaded portion of the histogram represents those HNNs that used a subset of predictor variables and yielded MSE values that were 10% larger than that of at least one HNN. We inferred that these HNNs very likely converged to local minima instead of a global optimum. The green shaded portion represents those HNNs with MSE within 10% of that of any HNNs that use a subset of predictor variables. These HNNs were considered to have been more likely to converge to the global minima.

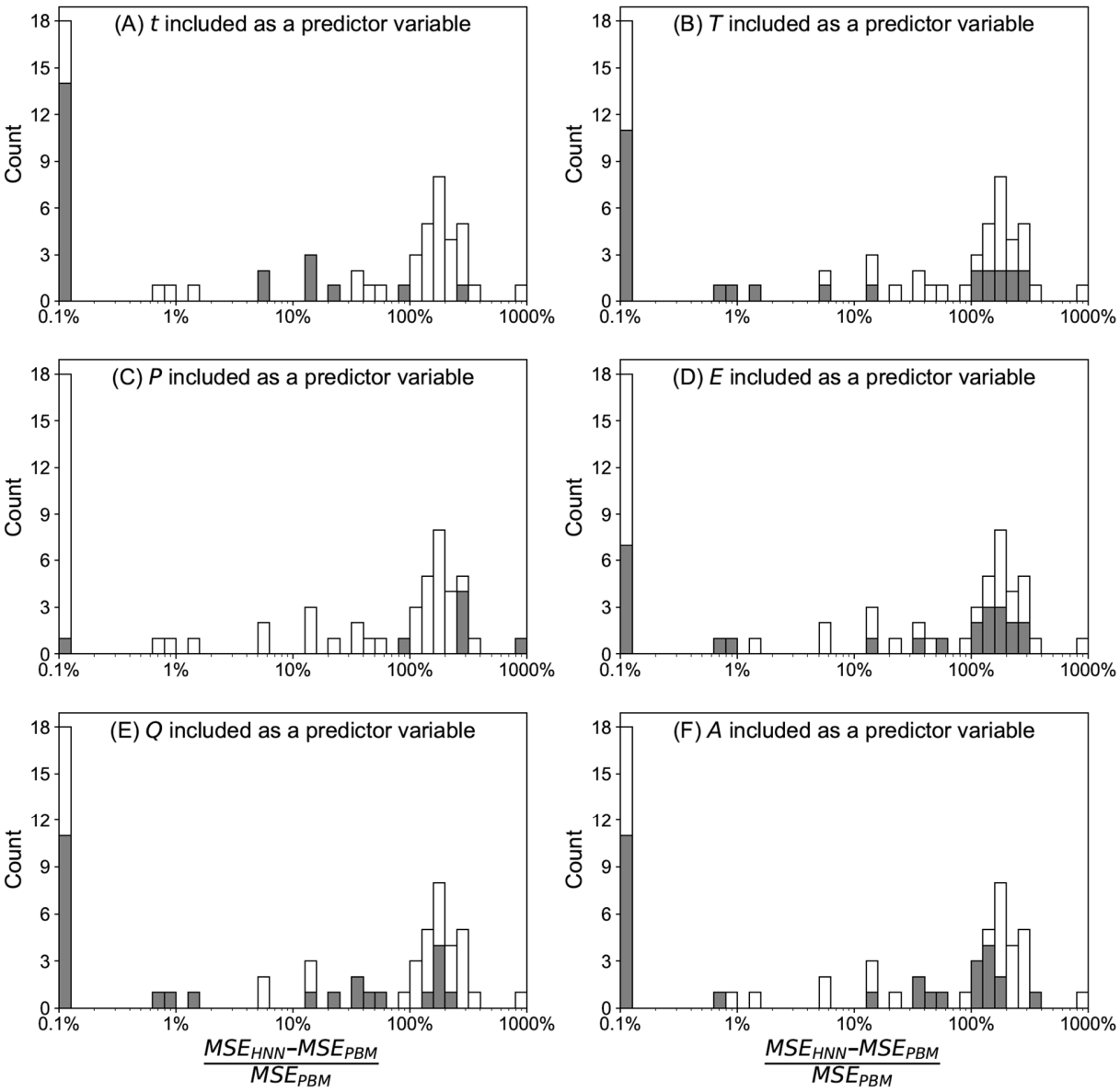
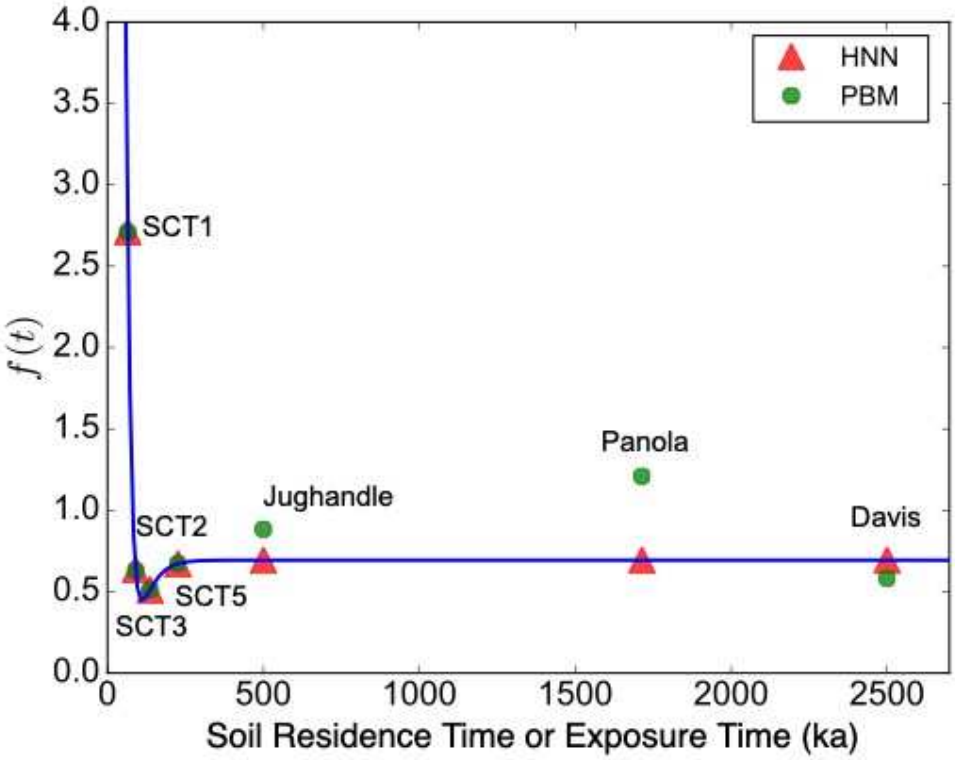


Figure 4. Histogram of percent differences in MSE between HNNs and their corresponding PBMs in the training phase (local-converged HNNs excluded). All negative percent differences were set to 0.1% for visualization purposes. The shaded portion represent those HNNs that include (A) soil residence time / exposure time or t, (B) temperature or T, (C) precipitation or P, (D) erosion rate or E, (E) quartz content or Q, and (F) albite content of the feldspar or A as one of the predictor variables.



748
749 **Figure 5.** Plot showing *f(z)* predicted by HNN 1c *f(t)* (red triangles) as a function of the soil
750 residence time. $\Gamma_{ini} \cdot \hat{k}$ values derived from the corresponding PBM are also shown for each soil
751 profile (green circles). Blue curve calculated from the function *f(t)* given by HNN 1c is plotted to
752 illustrate how the *f(z)* value varies with the soil residence time in the HNN 1c.
753

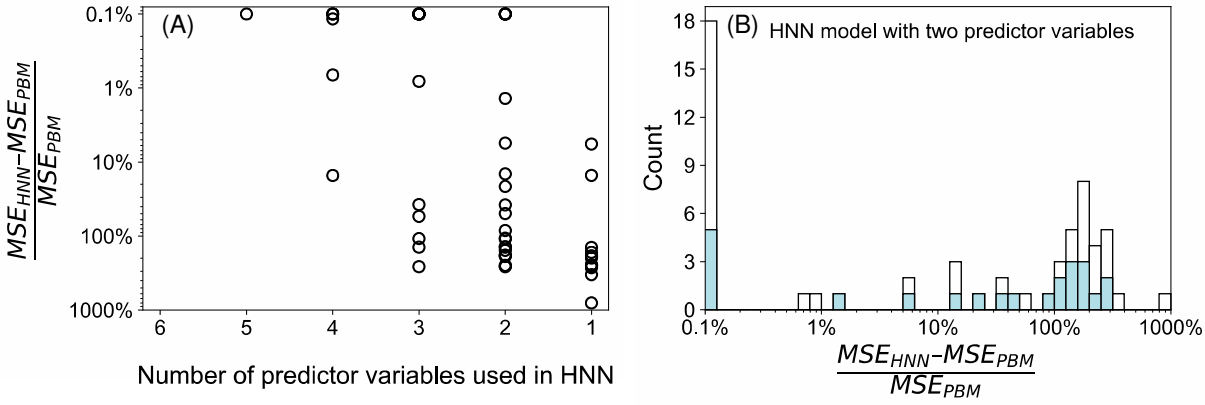


Figure 6. (A) Plot of percent difference in MSE between HNNs and their corresponding PBMs in the training phase as a function of the number of predictor variables used in the HNN. (B) Histogram of percent difference in MSE between HNNs and their corresponding PBMs in the training phase (local-converged HNNs excluded). All negative percent difference were set to 0.1% for visualization purposes. The blue shaded portion represent those HNNs with only two predictor variables.

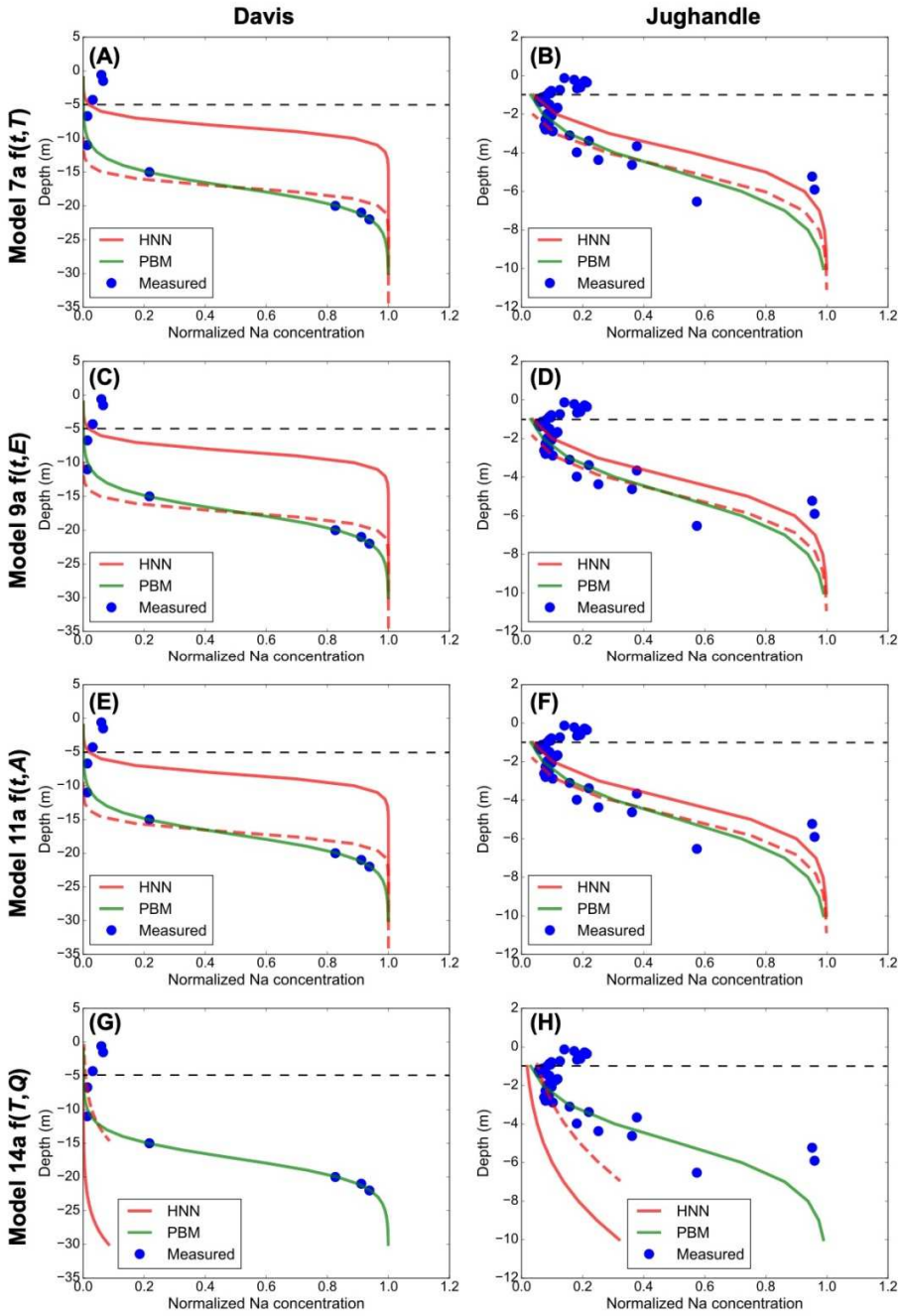


Figure 7. Normalized Na concentration (measured, blue symbols) plotted versus depth with results from the PBM fitted to the individual data points (green line) and the HNN results for each soil profile (red line). Panels (A), (C), (E), and (G) are for the Davis soil profile while panels (B), (D), (F), and (H) are for the Jughandle soil profile. (A) and (B), (C) and (D), (E) and (F), and (G) and (H) show results from HNN models 7a, 9a, 11a, and 14a, respectively. Note: red dashed curves represent HNN prediction results with a vertical offset to facilitate the comparison of the slope of PBM and HNN results. The black dashed line denotes the boundary above which soil samples might include exogenous disturbance (e.g., eolian input and bioturbation).

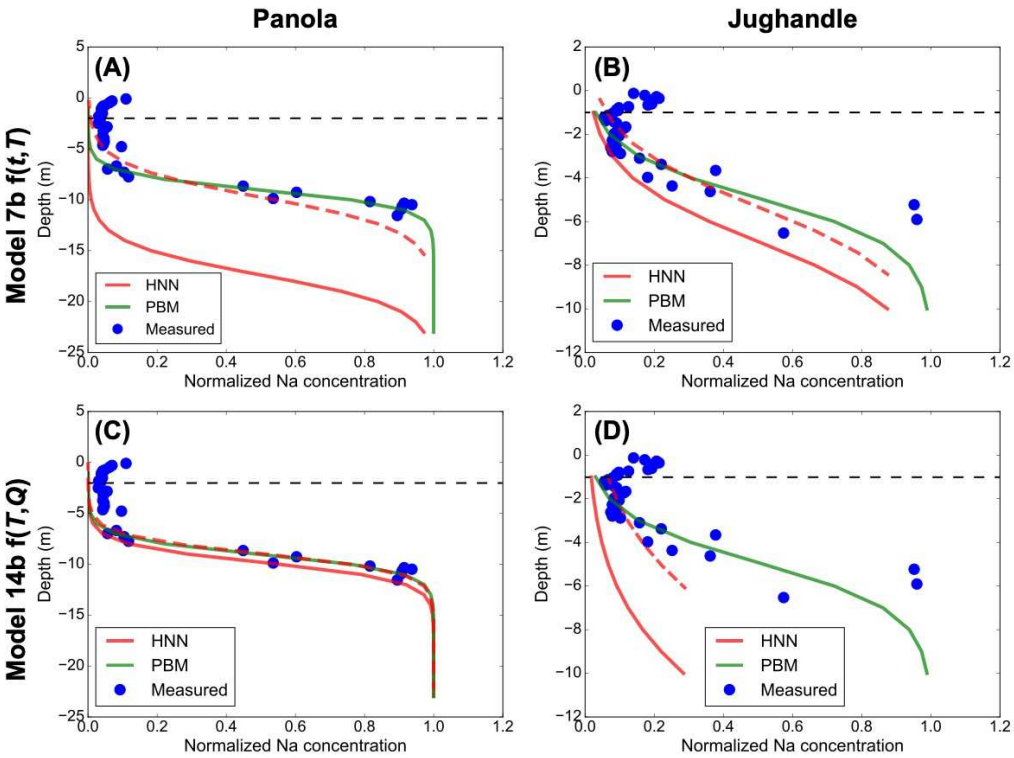


Figure 8. Normalized Na concentration (measured, blue symbols) plotted versus depth with results from the PBM fitted to the individual data points (green line) and the HNN results for each soil profile (red line). Panels (A) and (C) are for the Panola soil profile while panels (B) and (D) are for the Jughandle soil profile. (A) and (B), and (C) and (D) show results from HNN models 7b and 14b, respectively. Note: red dashed curves represent HNN prediction results with a vertical offset to facilitate the comparison of the slope of PBM and HNN results. The black dashed line denotes the boundary above which soil samples might include exogenous disturbance (e.g., eolian input and bioturbation).

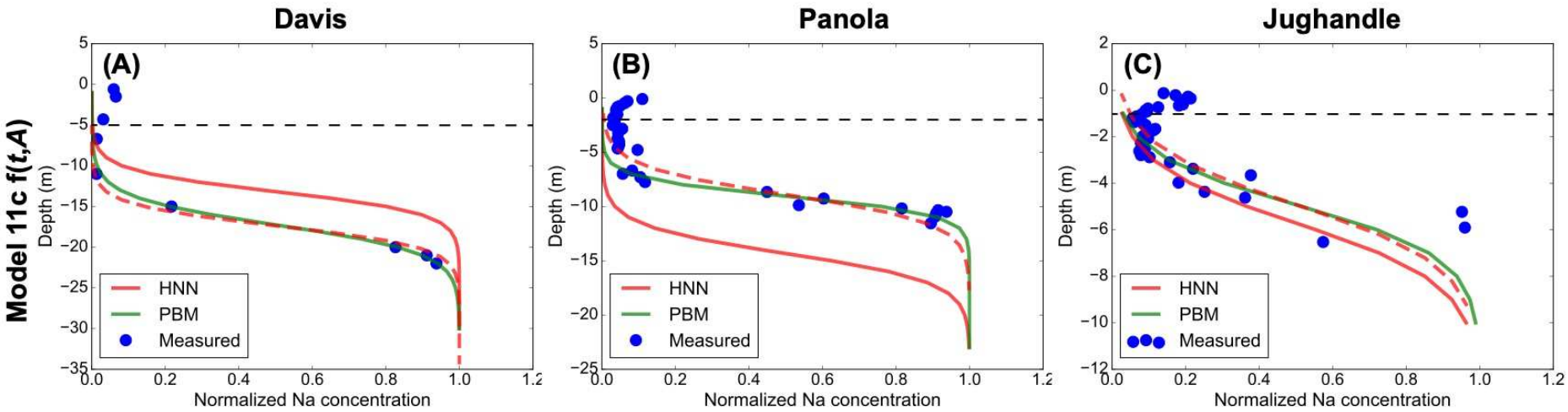


Figure 9. Normalized Na concentration (measured, blue dot) plotted versus depth with results from the PBM fitted to the individual data points (green line) and the HNN 11c results (red line) for soil profiles of (A) Davis, (B) Panola, and (C) Jughandle. Note: red dashed curves represent HNN prediction results with a vertical offset to facilitate the comparison of the slope of PBM and HNN results. The black dashed line denotes the boundary above which soil samples might include exogenous disturbance (e.g., eolian input and bioturbation).

Tables

Table 1. Summary of parameters describing soils[1]

Name	Residence time (t, ky)	Temperature (T, ^{o}C)	Erosion rate (E, m/My)	Precipitation (P, mm/y)	Quartz content (Q, vol %)	Fraction of Na feldspar in parent feldspar (A)	C_0 (mol/m^3)
SCT1	65	13.4	0	727	27.17[5]	0.66	1560[6]
SCT2	90	13.4	0	727	23.25[5]	0.73	1380[6]
SCT3	137	13.4	0	727	25.36[5]	0.72	1250[6]
SCT5	226	13.4	0	727	34.08[5]	0.70	1320[6]
Davis	2500[2]	10	4	1040	36.3	0.94	2500[6]
Panola	1714[2]	17	7	1240	27.5	0.77	3020[6]
Jughandle[4]	500	12.5	0[3]	983	34.08[3]	0.70[3]	656[7]

[1] Values were derived from White et al. (2008) and White et al., (2001) [for summary see also Brantley et al. (2008)] unless noted otherwise

[2] These times were calculated by assuming an erosion rate of 8 m/My or 7m My and the regolith thickness of 20 m and 12 m for Davis Run and Panola soils, respectively, for the residence time of minerals in the regolith depth interval assuming steady-state thickness, i.e., the soil residence time. These erosion rate and regolith thickness values were reported in the literature (Bierman et al., 1995; Pavich, 1986; Pavich et al., 1985; White et al., 2001).

[3] Assumed the same as SCT5

[4] Climate values from (Northup et al., 1995) and residence time estimated from (Merritts and Bull, 1989)

[5] Calculated as vol. % based on data reported in White et al. (2008) and Brantley et al. (2008)

[6] Best model fits from Brantley et al. (2008)

[7] Best PBM model fits performed in this study

803 **Table 2. HNN and PBM results for the best performing HNNs that use only two predictor**
804 **variables**

Predicted soils	MSE, HNN	MSE, PBM	$f(z)$	$\Gamma\_ini \cdot k\hat{}$
Model 7a with f(t,T) trained to SCT1,2,3,5, Panola				
SCT1,2,3,5, Panola (training)	0.0033	0.0033	-	-
Davis	0.18	0.001	1.2	0.58
Jughandle	0.023	0.015	1.1	0.88
Model 7b with f(t,T) trained to SCT1,2,3,5, Davis				
SCT1,2,3,5, Davis (training)	0.0031	0.003	-	-
Panola	0.17	0.0038	0.62	1.2
Jughandle	0.029	0.015	0.68	0.88
Model 9a with f(t,E) trained to SCT1,2,3,5, Panola				
SCT1,2,3,5, Panola (training)	0.0033	0.0033	-	-
Davis	0.18	0.001	1.2	0.58
Jughandle	0.021	0.015	0.88	0.88
Model 11a with f(t,A) trained to SCT1,2,3,5, Panola				
SCT1,2,3,5, Panola (training)	0.0033	0.0033	-	-
Davis	0.18	0.001	1.2	0.58
Jughandle	0.017	0.015	0.98	0.88
Model 11c with f(t,A) trained to SCT1,2,3,5				
SCT1,2,3,5 (training)	0.0031	0.0031	-	-
Panola	0.14	0.0038	0.82	1.2
Davis	0.064	0.001	0.82	0.58
Jughandle	0.019	0.015	0.81	0.88
Model 14a with f(T,Q) trained to SCT1,2,3,5, Panola				
SCT1,2,3,5, Panola (training)	0.0033	0.0033	-	-
Davis	0.27	0.001	-0.47	0.58
Jughandle	0.072	0.015	0.14	0.88
Model 14b with f(T,Q) trained to SCT1,2,3,5, Davis				
SCT1,2,3,5, Davis (training)	0.003	0.003	-	-
Panola	0.085	0.0038	1.7	1.2
Jughandle	0.051	0.015	0.48	0.88

805
806