RESEARCH ARTICLE

European Journal of **Soil Science** WILEY

Derivation of physically based soil hydraulic parameters in New Zealand by combining soil physics and hydropedology

J. A. P. Pollacco[1] | **J. Fernández-Gálvez**[2] | **T. Webb**[1] | **S. Vickers**[1] | **B. Robertson**[1] | **S. McNeill**[1] | **L. Lilburne**[1] | **C. Rajanayaka**[3] | **H. W. Chau**[4]

[1]Manaaki Whenua—Landcare Research, Lincoln, New Zealand

[2]Department of Regional Geographic Analysis and Physical Geography, University of Granada, Granada, Spain

[3]National Institute of Water and Atmospheric Research, Christchurch, New Zealand

[4]Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada, Lethbridge, Alberta, Canada

Correspondence
J. Fernández-Gálvez, Department of Regional Geographic Analysis and Physical Geography, University of Granada, Granada, Spain.
Email: jesusfg@ugr.es

Abstract

Field-characterised soil morphological data (to 1 m depth) and modelled soil water release characteristics are recorded in the S-map database for soils covering approximately 40% of New Zealand's soil area. This paper shows the development of the Smap-Hydro database that estimates hydraulic parameters by synergising soil morphologic data recorded in S-map and soil physics. The Smap-Hydro parameters were derived using the bi-modal Kosugi hydraulic function. The validity of the Smap-Hydro parameters was tested by applying them within an uncalibrated physically based hydrological model (HyPix) and comparing results with soil water content, θ, measured with Aquaflex soil moisture probes (0–40 cm deep) at 24 sites across New Zealand. The HyPix model provided an excellent fit with observed soil water content for 25% of the sites, a good fit for 33% of the sites and a poor fit for 42% of the sites. Applying the model to all soils in the S-map database required adjustments for the occurrence of rock fragments, hydraulic discontinuities caused by soil pans and required the addition of boundary conditions for water tables and the occurrence of impermeable rock. A discussion on how we can further synergise the development of pedotransfer functions with knowledge of soil physics is provided.

KEYWORDS
hydraulic parameters, HyPix model, physical hydrological model, Richards' equation, rock fragments, S-map, soil mapping, soil water retention, unsaturated hydraulic conductivity

1 | INTRODUCTION

Land and water quality degradation in New Zealand is caused by land intensification and climate change which has highlighted the need for more effective land management practices and nutrient discharge regulations, that can be achieved through the use of national catchment-based hydrological models (e.g., Cao et al., 2009; Srinivasan et al., 2021). These models require knowledge of the spatial variability of physical soil

hydraulic properties describing soil water retention, $\theta(\psi)$, and hydraulic conductivity, $K(\psi)$, functions. In New Zealand, hydraulic parameters are provided by the spatially distributed soil maps, Smap-hydro, which contains soil hydraulic parameters profiles up to 1 m depth for about 6000 soil types. Smap-hydro is linked to a soil map covering 40% of New Zealand's soil area in 2023, which makes up most of the highly productive land. The Smap-hydro is derived from the S-map database (https://smap.landcareresearch.co.nz). S-map provides, for every soil type, the soil water content, θ [L^3 L^{-3}] derived at the following tensions, ψ [L] = [0, 5, 10, 20, 40, 100, 1500] kPa by using empirical pedotransfer functions (PTFs) (PTFs were developed to derive soil hydraulic parameters from more easy to measure soil properties), which uses as explanatory variables *soil order classification, rock class, drainage class, texture* (Lilburne et al., 2004; McNeill et al., 2018).

Recently, there has been tremendous progress in digital soil mapping which has revolutionised the product of fine-resolution gridded soil data (Arrouays et al., 2020). Nevertheless, in the development of PTFs, there has been a lack of introducing knowledge of soil physics accumulated in the last decade. This paper addresses this research gap by deriving physical hydraulic parameters through Smap-Hydro which synergises empirical PTFs with soil physics knowledge. This enables the derivation of more robust plausible hydraulic parameters by assuring that the estimated hydraulic parameter sets are physically constrained from principles of soil physics (Fernández-Gálvez et al., 2021). The saturated hydraulic conductivity, K_s, [L T^{-1}] is derived from principles of soil physics by using as input data hydraulic parameters describing the $\theta(\psi)$ function (Pollacco et al., 2023). This assures that K_s is correlated to the pore size distribution described by $\theta(\psi)$. The advantage of using physical approaches to derive K_s is that it does not require a large observed K_s database compared to, for example, using a random forest algorithm as used in the European Pedotransfer functions (Tóth et al., 2015; Szabó et al., 2021) or by using eXtreme Gradient Boosting in Austria (Zeitfogel et al., 2023). The limited number of samples in New Zealand as well as the very large spatial variability of K_s due to the very young, alluvial and volcanic soils (Hewitt et al., 2021), makes it very problematic to rely on statistical approaches.

This research proposes and tests the following approach (illustrated in Figure 1) to combining empirical models and theoretical physics knowledge to generate soil hydraulic parameters:

1. Collate pedological data (from the S-map database) of soil morphological descriptors relevant to soil water storage and permeability such as *soil order, functional horizons* and *texture*.

Highlights

- Smap-Hydro derives physical soil hydraulic properties down to 1 m deep with a 40% coverage of New Zealand's soil area.
- Smap-Hydro is derived with pedotransfer functions and soil physics.
- Hydraulic parameters are corrected for rock fragments by considering fragment-to-fragment contact.
- Hydraulic parameters derived from Smap-Hydro are validated by a physically based hydrological model.
- The hydrological model gives an excellent fit simulating soil water dynamics for 25% of the sites and a good fit for 33%.

2. Derive, for each soil horizon, data points describing the relationship $\theta(\psi)$ of fines using empirical PTFs.

3. Ensure that the derived hydraulic parameters are physically and dynamically constrained (Fernández-Gálvez et al., 2021).

4. Correct the $\theta(\psi)$ of fines for rock fragments considering the water retention characteristics of the rock fragments, $\theta(\psi)_{rf}$.

5. From the corrected $\theta(\psi)$, derive physically plausible hydraulic parameters by using a lognormal model.

6. Derive the saturated hydraulic conductivity, K_s [L T^{-1}], from hydraulic parameters that describe $\theta(\psi)$ and the percentage of *clay*.

7. Test the effectiveness of the Smap-Hydro hydraulic parameters by introducing Smap-Hydro hydraulic parameters into the uncalibrated physically based HyPix model (Pollacco, Fernandez-Galvez, Ackerer, et al., 2022; Pollacco, Fernández-Gálvez, Rajanayaka, et al., 2022) and comparing modelled soil water content, θ_{model}, outputs with measured soil water content, $\theta_{measured}$, at specific locations.

The K_s model derived in step 6 was developed by Pollacco et al. (2013) from the unimodal Kosugi $\theta(\psi)$ by using the Hagen–Poiseuille equation which was later extended to bi-modal $\theta(\psi)$ (Pollacco et al., 2017). Nevertheless, the Hagen–Poiseuille equation does not give satisfactory results for fine-textured soils because it does not model the processes of clay–water interactions. Pollacco et al. (2023) improved the K_s model to make predictions of K_s for fine-textured soils by introducing a novel clay–water interaction function, which accounts for the strong bonding of water molecules to clay particles.

In Smap-Hydro, the $\theta(\psi)$ and $K(\psi)$ need to be corrected for *rock fragments* which is one of the most

FIGURE 1 A schematic illustrating the processes of deriving Smap-Hydro physical hydrological parameters.
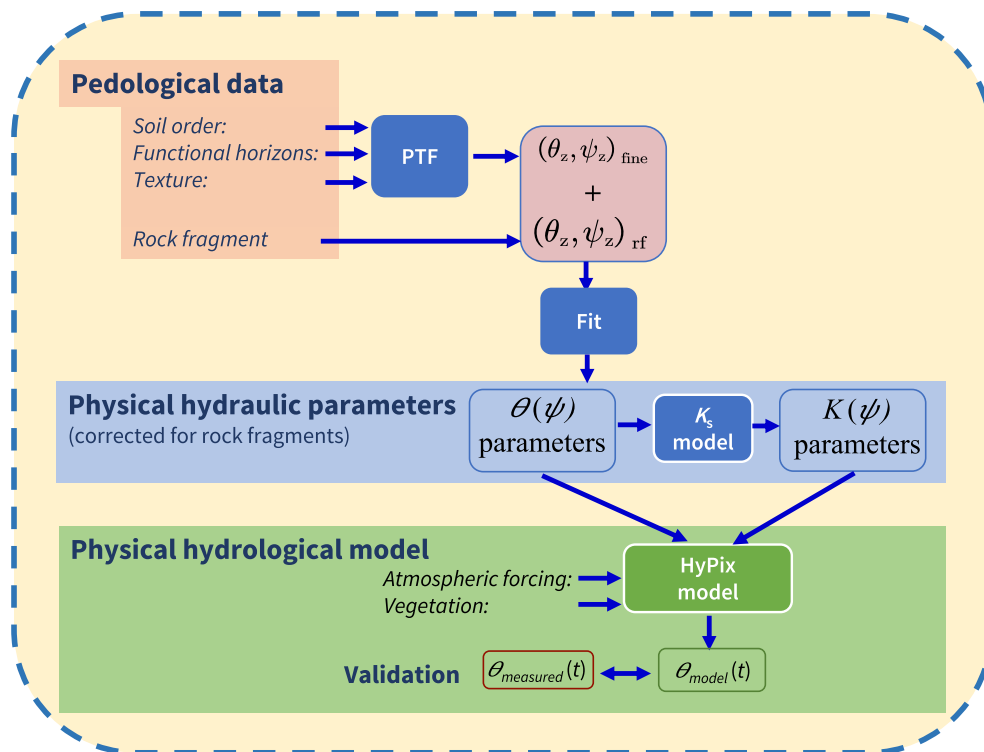
significant knowledge gaps in soil hydrology (Zhang et al., 2016). This is mainly because of the cost-intensive measurements necessary to quantify in situ water flow in stony soils (e.g., Fernández-Gálvez et al., 2006; Coppola et al., 2013; Pakparvar et al., 2016). Smap-Hydro corrects $\theta(\psi)$ for *rock fragments* based on research performed internationally (Naseri et al., 2023; Parajuli et al., 2017; Peters & Klavetter, 1988; Robertson et al., 2021).

The correction of K_s for *rock fragments* is more complicated than correcting $\theta(\psi)$ because *rock fragments* reduce the number of pores and increase the travel time of water travelling downwards. This is because *rock fragments* cause water molecules to meander between the fragments, causing an increase in tortuosity and a decrease in K_s. Nevertheless, when *rock fragment* content becomes high, fragment-to-fragment contact increases and K_s gradually increases again. This explains why different studies (e.g., Fernández-Gálvez et al., 2006; Hlaváčiková et al., 2016; Ma et al., 2010; Ma & Shao, 2008; Mehuys et al., 1975; Novák et al., 2011; Shao et al., 2009; Yang et al., 2013) have shown either positive or negative relationships between *rock fragments* and soil hydrological processes (e.g., *soil water storage, groundwater recharge, runoff generation, solute transport, preferential flow*). K_s starts to increase when *rock fragment* content reaches a certain threshold, which depends on many factors, such as the shape and orientation of

the *rock fragments*. No models are corrected for *rock fragments* greater than this threshold. Therefore, we introduce a novel K_s model for correcting *rock fragments* by accounting for water flow processes in soils with high *rock fragment content* that is no longer dominated by matrix flow but by film flow due to fragment-to-fragment contact.

The spatial 3D hydraulic parameters derived from Smap-Hydro are currently implemented into a wide range of physical and conceptual models in New Zealand such as: JULES (https://jules.jchmr.org/) (Sood et al., 2021); OVERSEER (https://www.overseer.org.nz/) (Giltrap et al., 2022); NZ Water Model – Hydrology (https://niwa.co.nz/freshwater/research-projects/nz-water-model-hydrology-nzwam-hydro) (Rajanayaka et al., 2022); SWAT (https://swat.tamu.edu/) (Pollacco et al. 2022); and HyPix physically-based hydrological models (Pollacco, Fernández-Gálvez, Ackerer et al., 2022; Pollacco, Fernández-Gálvez, Rajanayaka et al., 2022).

This paper is organised as follows: Section 2 describes the theory deriving the algorithms used to generate Smap-Hydro parameters; Section 3 describes the data and model used in the development and validation process; Section 4 presents and analyses the model outputs using Smap-Hydro parameters and compares the simulated soil water content with the point scale measurements from the field; Section 5 discusses how we can further synergise the development of PTFs with soil physics and Section 6 summarises the main conclusions.

2 | THEORY

2.1 | Deriving physical soil hydraulic parameters

2.1.1 | Lognormal bi-modal $\theta(\psi)$ and $K(\psi)$ functions

The Smap-Hydro parameters are derived using Kosugi'(1994, 1996) soil hydraulic functions. The choice of the Kosugi model is based on the physical interpretation of the parameters in relation to the soil pore size distribution and the fact that these parameters can be constrained by exploiting the relationship between them (Fernández-Gálvez et al., 2021). Moreover, the selection of bi-modal functions is based on the prevalence of soils with a bi-modal pore system (Jarvis, 2007; McLeod et al., 2008; Zhang et al., 2022), where macropores and micropores lead to a two-stage drainage. Fast flow (macropore flow) can occur when the water pressure head exceeds the threshold needed to activate the macropore network, adding to the matrix flow. Below this threshold, only the matrix contributes to the flow (Fernández-Gálvez et al., 2021). The representation of the $\theta(\psi)$ and $K(\psi)$ functions uses the dual porosity model from Pollacco et al. (2017) based on the Kosugi (1999; 1996) lognormal distribution. This sums up the *soil water retention* of the matrix domain, $\theta_{Mat}(\psi)$ and the *soil water retention* of the macropore domain, $\theta_{Mac}(\psi)$:

$$\begin{cases} \theta(\psi) = \theta_{Mat}(\psi) + \theta_{Mac}(\psi) \\ \theta_{Mat}(\psi) = \frac{1}{2}[\theta_{sMacMat} - \theta_r]\,\mathrm{erfc}\left[\frac{\ln(\psi/\psi_m)}{\sqrt{2}\sigma}\right] + \theta_r \\ \theta_{Mac}(\psi) = \frac{1}{2}[\theta_s - \theta_{sMacMat}]\,\mathrm{erfc}\left[\frac{\ln(\psi/\psi_{mMac})}{\sqrt{2}\sigma_{Mac}}\right] \end{cases} \quad (1)$$

where ψ [L] is the absolute value of the soil water pressure head (i.e., *matrix suction*, $\psi \geq 0$); θ_s [$L^3\,L^{-3}$] and θ_r [$L^3\,L^{-3}$] are the *saturated* and *residual* volumetric soil water content, respectively; $\theta_{sMacMat}$ [$L^3\,L^{-3}$] is the *volumetric soil water content* when the matrix domain is saturated with water and the macropore domain is filled with air; $\ln\psi_m$ and σ [L] denote the mean and standard deviation of $\ln\psi$, respectively, in the soil matrix domain; and $\ln\psi_{mMac}$ and σ_{Mac} [L] indicate the mean and standard deviation of $\ln\psi$, respectively, in the soil macropore domain.

The Kosugi (1999; 1996) log-normal unsaturated hydraulic conductivity is described by Pollacco et al. (2017) in its bi-modal form as follows:

$$\begin{cases} S_e(\psi) = \frac{1}{\theta_s - \theta_r}\frac{1}{2}\left\{(\theta_{sMacMat} - \theta_r)\mathrm{erfc}\left[\frac{\ln(\psi/\psi_m)}{\sqrt{2}\sigma}\right] \right. \\ \qquad\qquad \left. + (\theta_s - \theta_{sMacMat})\mathrm{erfc}\left[\frac{\ln(\psi/\psi_{mMac})}{\sqrt{2}\sigma_{Mac}}\right]\right\} \\ K_{sMac} = K_s\frac{\theta_s - \theta_{sMacMat}}{\theta_s - \theta_r} \\ K_{sMat} = K_s\frac{\theta_{sMacMat} - \theta_r}{\theta_s - \theta_r} \\ K(\psi) = \frac{\sqrt{S_e}}{2}\left\{K_{sMat}\left[\mathrm{erfc}\left[\frac{\ln(\psi/\psi_m)}{\sqrt{2}\sigma} + \frac{\sigma}{\sqrt{2}}\right]\right]^2 \right. \\ \qquad\qquad \left. + K_{sMac}\left[\mathrm{erfc}\left[\frac{\ln(\psi/\psi_{mMac})}{\sqrt{2}\sigma_{Mac}} + \frac{\sigma}{\sqrt{2}}\right]\right]^2\right\} \\ K_s = K_{sMac} + K_{sMat} \end{cases} \quad (2)$$

where $S_e = (\theta(\psi) - \theta_r)/(\theta_s - \theta_r)$ is the effective soil water content and K_s [$L\,T^{-1}$] is the saturated hydraulic conductivity.

The Smap-Hydro software can also generate the hydraulic parameters of several $\theta(\psi)$ and $K(\psi)$ functions, for example: the van Genuchten (1980) with the Mualem (1976) condition, the Brooks and Corey (1964), the Clapp and Hornberger (1978) and the Green and Ampt (1911) functions.

2.2 | Deriving unique physical soil hydraulic parameters describing $\theta(\psi)$

2.2.1 | Physically constraining the Kosugi hydraulic parameters

The hydraulic parameters are obtained by fitting the modelled points of the $\theta(\psi)$ derived from PTF, to the corresponding Kosugi $\theta(\psi)$ function. However, a severe drawback of deriving the hydraulic parameters by inverse modelling is that they suffer from equifinality or non-uniqueness (Fernández-Gálvez et al., 2021; Pollacco et al., 2008). Equifinality occurs when more than one set of parameters gives acceptable simulations relative to a given measure of goodness-of-fit between simulated and measured values (Pollacco et al., 2008; Pollacco & Angulo-Jaramilo, 2009). Pollacco et al. (2008) found that to eliminate the *equifinality* and obtain a unique set of hydraulic parameters, it is necessary to invert the hydraulic parameters simultaneously from observations of both $\theta(\psi)$ and $K(\psi)$ and that the measurements cover the full range of θ, from fully saturated to oven dry.

TABLE 1 Dynamically constraining the bi-modal hydraulic parameters. $\psi_{\mathrm{MacMat}} = 100$ mm is the water pressure head boundary between the macropores and matrix.

	θs [m³ m⁻³]	θ_{r} [m³ m⁻³]	σ [−]	ψ_{m} [mm]	$\theta_{\mathrm{sMacMat}}$ [m³ m⁻³]	σ_{Mac} [−]	ψ_{mMac} [mm]	K_{s} [mm s⁻¹]
Opt.	–	✓	✓	✓	✓	–	–	–
Ref.	–	–	–	Equation (3)	(Fernández-Gálvez et al., 2021)			Equation (4)
Min.	–	0	0.75	$\sqrt{\psi_{\mathrm{MacMat}}}\,e^{\sigma P_\sigma}$	$0.75\,\theta_{\mathrm{s}}$	$\frac{\ln\left(\sqrt{\psi_{\mathrm{MacMat}}}\right)}{P_{\sigma\mathrm{Mac}}}$	$\sqrt{\psi_{\mathrm{MacMat}}}$	$K_{\mathrm{sModel}}[\theta(\psi), clay]$
Max.	–	0.2	3.75	$\psi_{\mathrm{MacMat}}\,e^{\sigma P_\sigma}$	θ_{s}			

Note: $P_\sigma = 3$ and $P_{\sigma\mathrm{Mac}} = 2$ are the multiplier of the standard deviation for matrix and macropore domains respectively; $\alpha = 0.95$ [−] is a parameter that considers the air entrapment and the fact that not all pores are connected (Pollacco et al., 2013, 2020). The four parameters that are optimised, *Opt*, are indicated by ✓.

However, in S-map, there are only limited observed $\theta(\psi)$ data points, and therefore, it is not possible to obtain a unique estimation of the hydraulic parameters (θ_{s}, θ_{r}, σ, ψ_{m}, $\theta_{\mathrm{sMacMat}}$, σ_{Mac}, ψ_{mMac}). Therefore, it is necessary to physically constrain the hydraulic parameters using relationships among the Kosugi parameters, as described in Table 1.

Pollacco et al. (2013) showed that it is possible to reduce the non-uniqueness of the Kosugi hydraulic parameters and obtain physical parameters, by using the positive linear correlation between ln ψ_{m} and σ to constrain the relationship between ψ_{m} and σ. Such correlation exists because of the large median pore size (small ψ_{m}), characteristic of coarsely textured soils, which is linked to a smaller dispersion in pore size (small σ). This relationship can be explained by the fact that when ψ_{m} is small, the soil tends to have single-grain texture (monodisperse), so σ tends to be small. On the other hand, when ψ_{m} increases, it corresponds to a smaller median pore size, which is characteristic of aggregated soils with finer material and a range of grain sizes (polydisperse) and so σ tends to be larger.

Fernández-Gálvez et al. (2021) physically constrained the relationship between ψ_{m} and σ by assuming that $\theta(\psi)$ has a bi-modal log-normal probability density function, in which, the feasible range of ψ_{m} is derived from σ as follows:

$$\sqrt{\psi_{\mathrm{MacMat}}}\,e^{\sigma P_\sigma} \le \psi_{\mathrm{m}} \le \psi_{\mathrm{MacMat}}\,e^{\sigma P_\sigma} \tag{3}$$

where ψ_{MacMat} [L] is the water pressure head boundary between the macropores and matrix. Thus, when the pores are saturated ($\psi < \psi_{\mathrm{MacMat}}$), the flow is considered macropore flow and when the soil is desaturated ($\psi > \psi_{\mathrm{MacMat}}$), the flow is considered matrix flow. Jarvis (2007) found an average value of $\psi_{\mathrm{MacMat}} = 100$ mm. $P_\sigma = 3$ gives ln $\psi_{\mathrm{MacMat}} >$ ln ψ_{mMac} by three times the standard deviation of the logarithmic macropore size distribution.

The robustness of constraining ψ_{m} and σ has been demonstrated by Pollacco et al. (2013) and Vogeler et al. (2021), which derives satisfactory physical predictions of bi-modal Kosugi hydraulic parameters exclusively from (ψ) in the absence of $K(\psi)$, as in the S-map database.

It is to be noted that the algorithm for reducing the non-uniqueness of the hydraulic parameters is specific to the Kosugi hydraulic functions due to the physical meaning of its parameters.

2.3 | Deriving K_{s} from $\theta(\psi)$ parameters

S-map does not provide direct estimates of K_{s} and therefore, we derive K_{s} from the hydraulic parameters that describe the bi-modal Kosugi $\theta(\psi)$, defined in Equation (1) according to Pollacco et al. (2023) as follows:

$$\begin{cases} K_{\mathrm{sMat}} = 10^{\frac{\tau_1}{\tau_1-1}}\,C\,\pi\left\{(\theta_{\mathrm{sMacMat}} - \theta_{\mathrm{r}})^{\frac{T\mathrm{clay}(clay)}{4(1-\tau_3)}}\left(\frac{Y}{\psi_{\mathrm{m}}}\right)^{3(1-\tau_2)}\exp\left(\frac{[3([1-\tau_2])]^2\,\sigma^2}{2}\right)\right\}^{4(1-\tau_3)} \\[2mm] K_{\mathrm{sMac}} = 10^{\frac{\tau_1}{\tau_1-1}}\,C\,\pi\left\{(\theta_{\mathrm{s}} - \theta_{\mathrm{sMacMat}})\left(\frac{Y}{\psi_{\mathrm{mMac}}}\right)^{3(1-\tau_{2Mac})}\exp\left(\frac{[3(1-\tau_{2Mac})]^2\,\sigma_{\mathrm{Mac}}^2}{2}\right)\right\}^{4(1-\tau_{3Mac})} \\[2mm] K_{\mathrm{sModel}} = K_{\mathrm{sMac}} + K_{\mathrm{sMat}} \end{cases} \tag{4}$$

$K(\psi)_{model}$	τ_1	τ_2	τ_3	τ_{2Mac}	τ_{3Mac}	$Clay_0$	$\tau_{Clay_{\Delta\theta_{sr}}}$	$\tau_{Clay_{max}}$
	[0–1]							[1.1–100]
	0.763	0.604	0.589	0.962	0.00256	0.14	0.34	99.8

TABLE 2 Values of the tortuosity parameters for $K(\psi)_{model}$ [Equation (4)–(5)] with a feasible range [0–1].

where C is a constant ($C = 3.4542 \times 10^7$ mm d^{-1}) and τ_1, τ_2, τ_3, τ_{2Mac} and τ_{3Mac} [0–1] are tortuosity parameters; Y is the Young–Laplace capillary parameter, with $Y = 14.9$ mm^2 for pure water at 25°C and a contact angle of 0°.

The computation of the T_{clay} function is as follows:

$$
\begin{cases}
\Delta\theta_{sr} = \theta_{sMacMat} - \theta_r \\
\Delta\theta_{sr}^\eta = \max\left[\dfrac{\Delta\theta_{sr} - \tau_{Clay_{\Delta\theta_{sr}}}}{1 - \tau_{Clay_{\Delta\theta_{sr}}}}, 0\right] \\
Tclay_{max}\left(\Delta\theta_{sr}^\eta\right) = 1 + \Delta\theta_{sr}^\eta\left(\tau_{Clay_{max}} - 1\right) \\
Clay^\eta = \dfrac{\max\left(Clay - \tau_{Clay_0}, 0\right)}{1 - Clay_0} \\
T_{clay} = Tclay_{max}\left(\Delta\theta_{sr}^\eta\right) - \left[Tclay_{max}\left(\Delta\theta_{sr}^\eta\right) - 1\right]\cos\left(\dfrac{\pi}{2}Clay^\eta\right)
\end{cases}
\tag{5}
$$

where $Clay^\eta$ [0–1] is the normalised $Clay$; τ_{Clay_0} [0–1] is a fitted threshold parameter describing the minimum percentage of clay for which T_{clay} function starts to reduce; $\tau_{Clay_{\Delta\theta_{sr}}}$ [0–1] is another threshold of $\theta_{sMacMat} - \theta_r$; and $\tau_{Clay_{max}}$ is a fitting parameter which indicates the maximum value of T_{clay}.

When the *tortuosity parameter* values are close to 0, then the soil can be theoretically described as a bundle of 'perfect' circular and 'straight' capillary tubes. The values of the tortuosity parameters are described in Table 2.

2.4 | Correction for rock fragments

2.4.1 | Correction of $\theta(\psi)$ for rock fragments

S-map provides (a) data points of $\theta(\psi)$ for the soil fine fraction (<2 mm); (b) field estimate of the percentage volume of rock fragments, R_f [%]; and (c) data points for the water retention curve of rock fragments, $\theta(\psi)_{rf}$, representing the most common rocks found in New Zealand, as described in Table 3 and detailed in the next section (Robertson et al., 2021). The entire description of the soil water retention curve, $\theta(\psi)_{tot}$, is performed by combining the fine and the coarse (>2 mm) fractions. The following algorithm for correcting $\theta(\psi)$ was shown by several researchers (Naseri et al., 2023; Parajuli et al., 2017; Peters & Klavetter, 1988; Robertson et al., 2021) to match reasonably well with laboratory data. This is performed by partitioning $\theta(\psi)$ based on the volume taken by rock

fragments and that of the fines by using volume averaging (the composite-porosity model), as follows:

$$
\theta(\psi)_{tot} = \left(1 - R_f\right)\theta(\psi) + R_f\,\theta(\psi)_{rf}
\tag{6}
$$

The correction of $K(\psi)$ for R_f will be addressed in Section 2.4.2, as the saturated hydraulic conductivity is derived from $\theta(\psi)_{tot}$.

Water retention curve of rock fragments: $\theta(\psi)_{rf}$
The $\theta(\psi)_{rf}$ was measured using a novel methodology described in Robertson et al. (2021) that uses repacked soil cores comprising clasts (rock and glass) and fine earth. This is performed by incorporating a suction plate-core containment system that can be weighed as a unit to overcome typical core size restrictions. The method analyses the relationship between total core volumetric water content and R_f. Cores were packed with a mixture of inert glass and rock fragments to maintain a uniform volume of clasts between treatments, which allowed R_f [Equation (6)] to vary between cores, but the total clast proportion was held at 30%. A constant total clast volume improves accuracy and precision by ensuring the water-holding characteristics of the matrix vary as little as possible among cores. The derived $\theta(\psi)_{rf}$ is input into Equation (6). An estimate of $\theta(\psi)_{rf}$ for a range of rock types is described in Table 3.

2.4.2 | Correction of K_s for rock fragments

The correction of K_s to account for R_f is more complicated because rock fragments reduce the number of pores and increase the travel time of water travelling downwards. This is because rock fragments cause the water molecules to meander between the rock fragments, causing an increase in tortuosity and a decrease in K_s. However, when R_f becomes high, fragment-to-fragment contact increases and K_s gradually increases again. The threshold R_f when K_s starts to increase, $R_{f\_Contact}$, generally occurs when R_f exceeds 40% (e.g., Beibei et al., 2009; Urbanek Shakesby, 2009). This threshold depends on numerous factors, such as the shape and orientation of the rock fragments. There are no models to correct $K(\psi)$ for $R_f > R_{f\_Contact}$, so we propose an empirical model to correct K_s for $R_f > R_{f\_Contact}$.

T A B L E 3 Description of the $\theta(\psi)_{rf}$ modelled for a range of rock lithology implemented into Smap-Hydro.

Rock lithology	Rock class	$\theta(\psi)_{rf}$					Notes
		0	5	10	100	1500	
Hard quartzitic Marble	Hard metamorphic	0.04	0.03	0.02	0.01	0.01	Low porous, hard, coarse grain, does not weather.
Diorite Granite & Gneiss Greywacke Limestone Schist Tuffaceous sandstone	Hard coarse grain	0.1	0.07	0.05	0.03	0.01	Based on Robertson et al. (2021) volumetric quantification of rock fragments across Canterbury.
Hard mudstone Tuffaceous mudstone	Hard fine grain	0.1	0.09	0.08	0.05	0.03	Based on *Hard coarse grain*, but as per Maff (1988) there is a slight increase to values as they are likely to weather faster.
Andesite Basalt Gabbro Ignimbrite Rhyolite	Volcanic	0.12	0.08	0.06	0.03	0.01	Based on Maff (1988), which shows volcanic rocks tend to be more porous than sedimentary hard rocks.
Soft mudstone	Soft fine grain	0.13	0.12	0.11	0.06	0.03	Schoeman et al. (1997)
Carbonaceous Soft calcareous Soft quartzitic sediments Soft sandstone	Soft coarse grain	0.15	0.12	0.1	0.07	0.04	Schoeman et al. (1997)

Note: The $\psi = [0, 5, 10, 100, 1500]$ are in units of kPa.

Correcting K_s for $R_f \leq R_{f\_Contact}$

For $R_f < R_{f\_Contact}$, there is no need to perform further corrections because K_{sModel} [Equation (4)] uses hydraulic parameters that describe $\theta(\psi)_{tot}$, which was previously corrected for R_f by using Equation (6). K_{sModel} can be conceptualised as follows:

$$K_{s\_rf} \approx K_s (1 - R_f)^{1.64} \tag{7}$$

A comparison of this model with that of Naseri et al. (2023), who tested different models (Peck & Watson, 1979; Novák et al., 2011) for how R_f impacts $K(\psi)$ with different prepacked stony soils, found that the Naseri et al., (2022) model performed the best:

$$K_{s\_rf} \approx K_s \left(1 - \frac{R_f}{0.982}\right)^{1.26} \tag{8}$$

Due to the similarities of the proposed model with that of Naseri et al. (2022), we do not modify our model further for $R_f \leq R_{f\_Contact}$. Note, Naseri et al. (2022) did not indicate the range of R_f for which their model is valid.

Correcting K_s for $R_f > R_{f\_Contact}$

The processes for water flow for high R_f are no longer dominated by the matrix flow but by film flow due to the fragment-to-fragment contact. Because there was no available data to test the proposed K_{sModel} model, we obtained the $R_{f\_Contact}$ (40%) from the literature (Beibei et al., 2009, Figure 3; Khetdan et al., 2017, Figure 5; Beckers et al., 2016, Figure 2). The increase in K_s with increasing R_f is gradual and the maximum increase of K_s with increasing R_f is similar to when $R_f = 0$. Therefore, we propose the following model, which increases linearly from $K_{s\_rf}(R_f = R_{f\_Contact}) = K_{sModel}(R_f = R_{f\_Contact})$ to $K_s = K_{sModel}(R_f = R_{f\_Max})$, where $R_{f\_Max} = 90\%$ is the maximum allowed R_f.

For that, we modified θ_s by applying a linear relationship with R_f as follows:

$$\begin{bmatrix} R_{f\_Max} \\ R_{f\_Contact} \end{bmatrix} \begin{bmatrix} R_f \\ R_f \end{bmatrix} = \begin{bmatrix} \theta_s \\ \theta_s(1 - R_f) \end{bmatrix} \tag{9}$$

We modified θ_r by applying a linear relationship with R_f as follows:
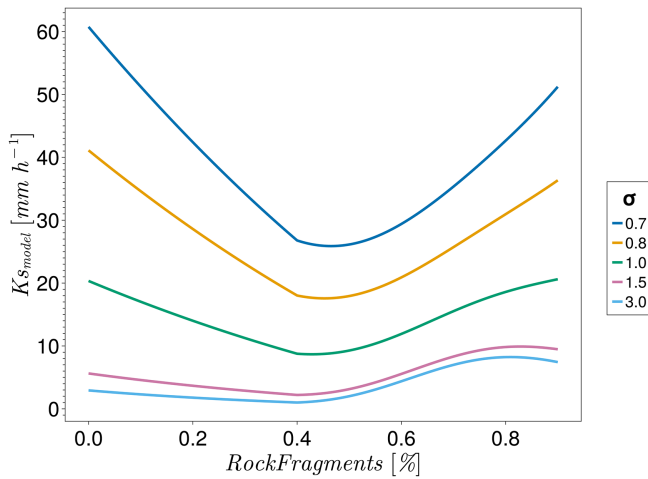
FIGURE 2   Relationship for K_{sModel} and R_f [0 to 0.9] for soils with small σ representing sandy soils and large σ representing clay soils.

$$\begin{bmatrix} R_{f\_Max} \\ R_{f\_Contact} \end{bmatrix} = \begin{bmatrix} R_f \\ R_f \end{bmatrix} \begin{bmatrix} \theta_r \\ \theta_r(1 - R_f) \end{bmatrix} \quad (10)$$

We assume that when fragment-to-fragment contact occurs, macroporosity is described by $\theta_{sMacMat}$ [Equation (1)], which increases with increasing R_f. So we corrected $\theta_{sMacMat}$ as follows:

$$\begin{bmatrix} R_{f\_Max} \\ R_{f\_Contact} \end{bmatrix} = \begin{bmatrix} R_f \\ R_f \end{bmatrix} \begin{bmatrix} 0.7(\theta_s - \theta_r) + \theta_r \\ \theta_{sMacMat}(1 - R_f) \end{bmatrix} \quad (11)$$

Figure 2 illustrates the influence of R_f on K_{sModel}, where low σ is representative of sandy soils and high σ is representative of clay soils, assuming that rocks are not wettable and K_{sModel} is not corrected for high *Clay* [Equation (5)]. There was no data to validate the model. Figure 2 shows that the K_{sModel} behaves as expected, with K_{sModel} decreasing with increasing R_f until $R_{f\_Contact}$ is reached, then K_{sModel} increases for increasing R_f until $R_f = R_{f\_Max}$. As expected, the increase and decrease in K_{sModel} with R_f is more pronounced for sandy soils than for soils containing more clay (high σ) because soils with a higher percentage of fine material obstruct large voids (Poesen & Lavee, 1994).

2.5 | Soils with impervious layers

There are soils with impervious layers (pans or rock interface) requiring default values of $\theta(\psi)$ and $K(\psi)$, as shown in Table 4.

The distinction between different pans is based on their New Zealand Soil Classification (Hewitt, 1992).

A *permeable rock* horizon is generated for soils with rock at the bottom of the soil profile, where that rock is not classified as being impermeable, based on root barrier type and the origin of the parent material.

2.6 | Physically based HyPix model

The physically based HyPix model (Pollacco, Fernandez-Galvez, Ackerer, et al., 2022; Pollacco, Fernández-Gálvez, Rajanayaka, et al., 2022) has been used to test the accuracy of the Smap-Hydro parameters by comparing the time series changes in θ computed by HyPix using Smap-Hydro parameters, θ_{hypix}, with the observed θ data, θ_{obs}, through time.

HyPix can (a) process a large number of soil layers, (b) simulate unimodal and bi-modal Kosugi hydraulic parameters for each soil layer, (c) simulate realistic water *ponding* at the soil surface by using a novel approach for the computation of *sorptivity* (Lassabatere et al., 2021; 2023), (d) compute θ at different depths, (e) compute rainfall *interception* from leaf area index, (f) derive *transpiration* from root water uptake with a compensation mechanism for deeper layers where root density is limited, (g) compute *evaporation* and (h) compute *drainage* through the bottom of the soil profile under different boundary conditions.

HyPix solves the mixed form of the Richardson–Richard's equation (RRE) using the Newton–Raphson method. The non-linear RRE is solved with an efficient heuristic and physical time-stepping strategy using a reduced number of control parameters with a maximum time step of 1 hour. HyPix also incorporates a novel algorithm to avoid 'overshooting' by controlling the Newton–Raphson step. The RRE partial differential equation is solved using a *cell-centred finite-volume* (implicit finite differences) scheme for the spatial discretisation, with an *implicit Euler scheme* for the temporal discretisation by using the *weighted average* inter-cell hydraulic conductivity. Assuming a rigid solid matrix, the mixed form of the RRE is written as:

$$\frac{\theta_i(\psi_i^t) - \theta_i(\psi_i^{t-1})}{\Delta T^t} - S_o \frac{\theta_i(\psi_i^t)}{\theta_{s_i}} \frac{\psi_i^{t-1} - \psi_i^t}{\Delta T^t}$$
$$= \frac{Q_{i-1/2}^t - Q_{i+1/2}^t}{\Delta Z_i} - Sink_i(\psi_i^{t-1}) \quad (12)$$

where ΔT^t [T] is the *time-step* at time t; ΔZ_i [L] is the *mesh size* of cell i, with the vertical coordinate positive downwards; θ_i [L³ L⁻³] is the *volumetric soil water content* of cell i; θ_{s_i} [L³ L⁻³] is the *saturated volumetric soil water content* of cell i; $S_0 = 10^{-6}$ is a parameter that accounts for *fluid compressibility*, which is assumed to be

TABLE 4 Default $\theta(\psi)$ and $K(\psi)$ hydraulic parameters for unconventional soils. The ψ [0, 5, 10, 100, 1500] are presented in units of kPa.

Horizons	K_s [mm s^{-1}]	$\theta(\psi)$							Comments
		0	5	10	20	40	100	1500	
Continuous soil pan	2.31×10^{-5}	0.1	0.098	0.096	0.095	0.092	0.088	0.068	$\theta(\psi)$ derived for soils with high clay content
Discontinuous soil pan	4.63×10^{-5}	0.1	0.079	0.073	0.067	0.060	0.050	0.023	$\theta(\psi)$ derived for soils with high clay content
Permeable rock	2.22×10^{-2}	0.134	0.037	0.019	0.016	0.016	0.016	0.016	$\theta(\psi)$ of (Dann et al., 2009, Lowcliffe sea cliff)

TABLE 5 Features describing boundary conditions.

Boundary cond.	Information
Free drainage	When no other boundary conditions are known
Impermeable	When a non-permeable rock is detected
Water table	When a permanent water table is detected, identified as a 'fluid' horizon

constant with depth; ψ_i [L] is the *soil water pressure* of cell i, considering $\psi < 0$ for unsaturated soils; Q [L T^{-1}] is the *soil water flux* based on the extended Darcy–Buckingham's law, which is positive downward and negative when water moves upwards; $Q_{i-1/2}^t$ [L T^{-1}] is the flux entering cell i and $Q_{i+1/2}^t$ [L T^{-1}] is the flux exiting cell i; and $Sink_i$ [L^3 L^{-3} T^{-1}], taken as positive, is the *sink term* defined as the volume of water per unit time removed from cell i by soil *evaporation* and *root water uptake*. Additional details of the HyPix model can be found in Pollacco, Fernandez-Galvez, Ackerer, et al. (2022) and Pollacco, Fernández-Gálvez, Rajanayaka et al., (2022). The input, outputs and parameters required by the Hypix model are described in Section 3.4.

2.6.1 | Boundary conditions derived from the pedological soil profile

Hydrological models require knowledge of the bottom boundary conditions, which is notoriously difficult to provide accurately. S-map provides information only to a maximum depth of 1 m, so if no pedological/geological feature is reported <1 m, then it is assumed that there is a free boundary condition. A *water table* boundary condition is set where a 'fluid' horizon is present in S-map and an *impermeable layer* boundary condition is set if the soil has 'massive' rock below that is determined to be impermeable, based on the origin of its parent material. The following boundary conditions (summarised in Table 5) are implemented into the physically based HyPix model, which is implemented into the AquaPore-ToolKit (Pollacco, Fernandez-Galvez, Ackerer, et al., 2022; Pollacco, Fernández-Gálvez, Rajanayaka, et al., 2022).

Free drainage boundary conditions
The free drainage boundary condition at the bottom of the soil profile occurs when the gradient of pressure, ψ, with depth, Z [L], is unity. The water flux Q [L T-1] is computed as follows:

$$Q_N = Ks_N(\psi_N^t) \frac{\partial \psi_N}{\partial Z_N} = K_N \qquad (13)$$

where N represents the index of the last cell in the soil profile.

Water table boundary conditions
The prescribed water table bottom water pressure boundary condition, $\psi_{bot} = 0$ [L], is set at the bottom of cell N of the soil profile and results in the following flux, which results in upward capillary flux.

$$Q_N = -Ks_N \left[\frac{0 - \psi_N}{\Delta Z_{N/2}} - 1 \right] \qquad (14)$$

Impermeable boundary conditions
This can be performed by setting $K_s = 0$ in the free drainage boundary condition [Equation (13)].

3 | MATERIALS AND METHODS

3.1 | Workflow of introducing Smap-Hydro into the physically based HyPix model

The hydraulic parameters of Smap-Hydro, describing $\theta(\psi)$ and $K(\psi)$ are derived using the following workflow:

1. Correct $\theta(\psi)$ for rock fragments by using Equation (6).
2. Optimise the bi-modal Kosugi hydraulic parameters by matching observed $\theta(\psi)_{Smap}$ data points with the bi-modal Kosugi function $\theta(\psi)$ [Equation (1)] and by minimising the objective function [Equation (16)].

This step considers constraining physically and dynamically the hydraulic parameters (Table 1).

3. Derive K_s from hydraulic parameters describing $\theta(\psi)$ and the percentage of clay using K_{sModel} [Equation (4)].

4. Correct K_s for R_f greater than 40% by using Equation (9)–(11).

5. Assess the accuracy of Smap-Hydro by introducing the Smap-Hydro hydraulic parameters into the uncalibrated physically based HyPix model and compare the modelled and observed θ averaged for the top 40 cm.

3.2 | Deriving S-map data

3.2.1 | Pedological description of S-map database

S-map is a digital spatial soil information system developed to provide soil morphological information for New Zealand soils (Webb & Lilburne, 2011; McNeill et al., 2018). The definition and descriptions of soil characteristics in S-map are clearly defined and, where appropriate, are quantitative.

At the highest level, S-map soil units are based on the New Zealand Soil Classification (Hewitt, 2010), developed in the 1980s, which organises the soils into 15 soil orders and then into soil groups and sub-groups. The 15 soil orders, three of them split according to their soil group, were found to have valuable relationships with $\theta(\psi)$ (and, indirectly, infiltration rates). The classification of soil orders is based on significant differences in soil forming factors, as follows:

- *Semi-arid, Pallic, Brown* and *Podzol* orders form a climatic sequence of soils with increasing rainfall and leaching.
- *Pumice* (rhyolitic pumice), *Allophanic* (volcanic ash) and *Melanic* (calcareous and basaltic) orders contain soils formed from distinctive parent materials.
- *Ultic, Granular* and *Oxidic* orders contain soils formed on old, weathered landscapes.
- *Raw* and *Recent orders* contain soils related to very young soil development.

- The *Gley* order contains poorly drained soils with an underlying water table.
- The *Organic* order contains soils formed from peat.
- The *Anthropic order* contains soils modified by significant human disturbance.

After soil classification, soils are grouped into soil families based on the soil profile form, *rock type*, *three texture classes* and *three permeability classes*. S-map contains information to a maximum depth of 1 m. The soil profile identifies four depth classes indicating depth to rock, to horizons with >35% R_f, or to a soil pan. The classes are very shallow (<20 cm), shallow (20–45 cm), moderately deep (45–100 cm) and deep (>100 cm).

Soil families are sub-divided into soil siblings (equivalent to soil types), each corresponding to a unique combination of up to six functional horizons. The functional horizons were designed to characterise soil horizons according to physical features related to water-holding capacity and permeability is summarised in Table 6 (Griffiths et al., 1999; Webb, 2003).

In the S-map, functional horizons are defined in terms of:

- *Topsoil/sub-soil*: topsoil generally has greater water content (and usually greater porosity) than sub-soil horizons with the same particle size;
- *Tephra/non-tephra*: soils formed from volcanic ejecta typically have higher total porosity and microporosity than other soils and commonly have greater plant-available water content;
- *Soil strength*: loose, compact and dense for soil with >35% R_f and weak, slightly firm and firm for soil with <35% R_f soil;
- *Ped size*: soil structure is sub-divided into two classes (>2 cm or <2 cm).

Functional horizons are also identified into four stoniness classes and three soil texture classes, but these are not used for the water retention model because every horizon in S-map is assigned a percentage of *rock fragments, sand, silt and clay* content and these are used.

The S-map database contains over 4000 siblings (soil types) with over 16,000 soil horizon descriptions.

Stone content	Texture of fines	Structure size	Consistency
Non-stony	Sandy (A)	Coarse (C)	Weak (w)
Stony (S)	Loamy (L)	Fine (F)	Slightly firm (s)
	Clayey (Y)		Firm (f)
Very stony (V)			Loose (l)
Extremely stony (X)			Compact (c)
			Dense (d)

TABLE 6 Attributes defining functional horizons (Webb, 2003).

3.2.2 | PTFs to derive $\theta(\psi)$ from soil information

Empirical PTFs for $\theta(\psi)$ have been developed by McNeill et al. (2018) to estimate the θ at different ψ, using explanatory variables derived in Section 3.2.1. The model described by McNeill et al. (2018) has recently been extended to incorporate more sample data, as well as additional explanatory variables. The explanatory variables include the *soil order classification, rock class* of the fines, *sample mid-depth, drainage class, texture* (sand, silt and clay) and parsed information from the S-map functional horizons. The parsed S-map information yields qualitative class information, such as the attributes described in Table 6.

The most influential parameters for predicting θ at different ψ were soil *order, texture* (sand, silt and clay), *topsoil* (true, or not) and *strength* (consistency), although the relative influence of these parameters depended on ψ. For instance, the S-map *tephra* class is more influential at ψ values closer to *total porosity* and less influential near the *wilting point*. The model is not causal, so the effect of a variable operates in association with the other variables, often in a manner that is not obvious. For instance, sample depth is not important per se, but sample depth is a surrogate for soil carbon concentration and the latter is influential in predicting θ. Similarly, *soil order* is associated with *soil density*, which most likely explains its strong association within the PTF. The *tephra* class is also influential for some ψ and its variable effect for different ψ is most likely due to its association with different soil orders. As a result, disentangling these effects to determine causal relationships is challenging.

The PTF used in McNeill et al. (2018) uses a generalised additive model (GAM, see Wood, 2017) with θ as the response, modelled as a Beta distribution to constrain the value in the range [0,1]. For ψ equals 1500 kPa, the scaling is direct. For other ψ, the scaling of the response is arranged so that the response is between the maximum at the next highest tension and unity. In this way, the predicted θ estimates are physically constrained to be non-decreasing for increasing ψ.

3.3 | Soil characterisation and soil water content measurements

3.3.1 | S-map descriptions of the monitoring sites

Validation of the Smap-Hydro model predictions was performed by using information from a separate set of soil sampling locations using Aquaflex sensors (www.aquaflex. co.nz) to characterise soil moisture. Soil descriptions were

available for each of the monitoring sites and the soils were described within 50 cm of the soil water content sensor. Soil descriptions by pedologists included estimates of *soil colour, texture, structure, compaction, mottle pattern, root abundance* and the presence of *water tables*. Fifteen of the 24 sites had particle size analysis derived in the laboratory for the top 40 cm of soil material. Site soil descriptions were then translated into the S-Map classification and soil horizon characteristics, which included estimates of R_f, *sand, silt* and *clay* percentages. For the remaining 9 sites, the percentage of the particle size estimates are based on general pedological descriptions of *texture*, such as *stony silt loam* and therefore have significant uncertainty. This data was entered into the S-map database and $\theta(\psi)$ relationships were derived from the water release model.

3.3.2 | Soil water content measurements using uncalibrated Aquaflex

We selected a total of 24 soil water content monitoring sites (Table 8) out of 30 available sites, the location of which is illustrated in Figure 3. The reason why we excluded the data from 6 sites is either because of mismatches of climate data and θ data, uncertainty of soil descriptions, or soil disturbance at the site.

θ measurements were recorded hourly using Aquaflex sensors. The Aquaflex sensors is a time domain transmissivity (TDT) soil moisture tape that measures θ integrated over a 3 m sensor length, installed from 5 to 35 cm depth. The sensing volume around the sensor is up to a 5 cm radius, equivalent to an overall soil volume of 6 L, corresponding to the top 40 cm of the soil profile. Accuracy is estimated at $\pm 2\%$ by the manufacturer. The dates of the simulation depended on the available data and we let the soil settle for 12 months before using the data.

The drawback of this dataset is that the sensors were not calibrated for local soil conditions, which implies that only differences of θ with time can be trusted and not the actual absolute θ value. Therefore, the mean adjustment correction to the observations, $\theta_{\text{MeanAdjObs}}$ [L^3 L^{-3}] and HyPix simulation, $\theta_{\text{MeanAdjHypix}}$ [L^3 L^{-3}], was computed as follows:

$$
\begin{cases}
\theta_{\text{MeanAdjObs}} = \theta_{\text{obs}} - \overline{\theta_{\text{obs}}} \\[2ex]
\theta_{\text{MeanAdjHypix}} = \dfrac{\sum\limits_{i=1}^{L_i} \Delta Z_i \theta_{\text{Hypix}_i}}{\sum\limits_{i=1}^{L_i} \Delta Z_i} - \overline{\dfrac{\sum\limits_{i=1}^{L_i} \Delta Z_i \theta_{\text{Hypix}_i}}{\sum\limits_{i=1}^{L_i} \Delta Z_i}}
\end{cases} \quad (15)
$$

where L_i is the number of soil layers in the top 40 cm; θ_{Hypix} is the simulated θ by HyPix model at

FIGURE 3   Location of the 24-soil water content monitoring sites.

different soil layers i, ΔZ_i is the thickness of the soil layer; θ_{obs} is observed θ derived from aquaflex. For each soil layer, $\bar{\theta}$, is the mean soil water content with time; $\sum_{i=1}^{N_i} \Delta Z_i = 40 cm$.

The number of Smap-Hydro horizons in the top 40 cm is site-specific, but ranges between 1 to 4 horizons.

Smap-Hydro: Optimisation of the soil hydraulic parameters

The fitting of the estimates of Smap-Hydro hydraulic parameters describing the bi-modal Kosugi hydraulic parameters of $\theta(\psi)$ [Equation (1)] is performed by minimising the following objective function (OF):

$$OF = \frac{\sum_{i=1}^{N_i} \left[\theta(\psi_i)_{Smap} - \theta(\psi_i)_{sim}\right]^2}{\sum_{i=1}^{N_i} \left[\theta(\psi_i)_{Smap} - \overline{\theta(\psi_i)_{Smap}}\right]^2} \quad (16)$$

where N_i for each soil sample is the number of experimentally measured S-map data points describing $\theta(\psi_i)_{Smap}$, which is corrected for rock fragments. The simulated values of the bi-modal lognormal model, $\theta(\psi_i)_{sim}$, are derived from Equation (1).

The Kosugi hydraulic parameters are dynamically constrained using the algorithms described in Table 1. The optimisation is performed by using the global adaptive differential evolution algorithm (Qin et al., 2009; Zhang & Sanderson, 2009) derived from the BlackBoxOptim.jl library (https://github.com/robertfeldt/BlackBoxOptim.jl), written in the Julia programming language (https://julialang.org).

3.4 | Parameterisation of the HyPix model

3.4.1 | Precipitation and potential evapotranspiration input data

The precipitation data is measured on-site, for which the missing values are corrected with the Virtual Climate Stations Network (VCSN) (Tait et al., 2006). The Penman–Monteith potential evapotranspiration is computed from VCSN which is recorded in Table 8.

3.4.2 | Discretisation of the Smap-Hydro parameters

Smap-Hydro provides the values of the hydraulic parameters for up to six functional horizons (depending on the heterogeneity of the soil) to a maximum depth of 1 m, depending on the boundary conditions. In HyPix, the vertical discretisation of the soil profile is performed automatically to honour the depth of the different soil horizons which vary in depth. To assure numerical stability, the discretisation has a cell of mesh size $\Delta Z \leq 20$ mm.

The soil types of the study sites are briefly described in Table 8 and include a wide range of soil *texture, soil depth, soil density* and *drainage*. These characteristics are expected to have significant effects on the temporal dynamics of θ.

3.4.3 | Vegetation parameters

Experimental sites are non-irrigated mixed pasture grass in New Zealand. The vegetation parameters introduced into the HyPix model are described in Table 7 for which the algorithms are detailed in (Pollacco, Fernandez-Galvez, Ackerer, et al., 2022; Pollacco, Fernández-Gálvez, Rajanayaka, et al., 2022).

$\psi_{Feddes1}$, $\psi_{Feddes2}$, $\psi_{Feddes3}$, $\psi_{Feddes4}$ are the trapezoidal (Feddes et al., 1978) water stress response functions used to represent parameters for mixed pasture grass in New Zealand and are derived from Wesseling (1991). The maximum root depth, Z_{root}, was derived by Vogeler & Cichota (2019) and ΔRdf_{top} is the percentage of roots in the top 30 cm and was taken from Evans (1978). The crop coefficient, K_C, was taken from rotated grazing pasture according to an FAO irrigation paper 56 (Allen et al., 1998) and the maximum saturated storage capacity of a wet canopy (S_{intsat}) is derived from Pollacco, Fernandez-Galvez, Ackerer, et al. (2022).

3.5 | Goodness-of-fit of the observed and simulated θ

We compare the goodness-of-fit of $\theta_{MeanAdjHypix}$ and $\theta_{MeanAdjObs}$ averaged for the top 40 cm by using the unbiased Nash–Sutcliffe model efficiency coefficient:

$$Nse_{unbias} = \frac{\sum_{t=1}^{T_t} \left[\theta_{MeanAdjHypix_t} - \theta_{MeanAdjObs}\right]^2}{\sum_{t=1}^{T_t} \left[\theta_{MeanAdjObs_t} - \overline{\theta_{MeanAdjObs}}\right]^2} \quad (17)$$

And the goodness-of-fit of θ derived by HyPix, θ_{Hpix} and θ derived from observations (Aquaflex), θ_{obs}, averaged for the top 40 cm by using the bias Nash–Sutcliffe model efficiency coefficient:

$$Nse_{bias} = \frac{\sum_{t=1}^{T_t} \left[\theta_{Hypix_t} - \theta_{Obs}\right]^2}{\sum_{t=1}^{T_t} \left[\theta_{Obs_t} - \overline{\theta_{Obs}}\right]^2} \quad (18)$$

where T_t is the number of time steps.

TABLE 7 Vegetation parameters representing non-irrigated pasture grass.

Z_{root}	ΔRdf_{top}	K_C	S_{intsat}	$\psi_{Feddes1}$	$\psi_{Feddes2}$	$\psi_{Feddes3}$	$\psi_{Feddes4}$
[mm]	[%]	[−]	[mm]	[mm]	[mm]	[mm]	[mm]
800	80%	0.86	1.28	100	250	5000	80,000

4 | RESULTS AND DISCUSSION

4.1 | Water balance from the HyPix model with hydraulic parameters derived from Smap-Hydro

To assess the effectiveness of Smap-Hydro, the Kosugi hydraulic parameters derived from Smap-Hydro were incorporated into the physically based HyPix model. A comparison was established between HyPix simulated θ for the top 40 cm and the point scale average θ measured by Aquaflex for the 24 monitoring sites across New Zealand described in Table 8. Figure 4 shows 9 of the measurement sites representative of different texture and soil order. For each of the 9 monitoring sites, Figure 4 shows the time series of *precipitation*, *drainage*, derived $\theta_{\text{MeanAdjHypix}}$ and $\theta_{\text{MeanAdjObs}}$ [Equation (15)]. For the observed $\theta_{\text{MeanAdjObs}}$ and simulated $\theta_{\text{MeanAdjHypix}}$, positive values of $\theta_{\text{MeanAdjObs}}$ or $\theta_{\text{MeanAdjHypix}}$ indicate that the value of θ is greater than the $\overline{\theta}$ and negative when the values of θ are below the $\overline{\theta}$.

For most sites shown in Figure 4, $\theta_{\text{MeanAdjHypix}}$ adequately match the dynamics of $\theta_{\text{MeanAdjObs}}$ and follow the general trend according to the forcing input data. *Rainfall* events are followed by a sudden increase in θ. After these episodes, θ decreases progressively, with a simultaneous increase in both simulated *drainage* and *evapotranspiration*. A sudden increase in *evapotranspiration* values occurs because, after each *rainfall* event, the amount of water available next to the soil surface is greater. *Drainage* does not occur immediately after a *rainfall* event. There is a delay in the *drainage* pulse, with differences related to the amount of water reaching the surface during each *rainfall* event and the hydraulic characteristics of each monitoring site.

The introduction of the *water table boundary condition* improved, as expected, of 8% for the *Rangiora* site but did not improve the *Turangi* site, which is an organic soil. It is known that S-map has to date some limitations with organic soils.

4.2 | Comparison between the observed and modelled data

The goodness-of-fit between the HyPix outputs, $\theta_{\text{MeanAdjHypix}}$ and θ_{Hpix} and the corresponding observations, $\theta_{\text{MeanAdjObs}}$ and θ_{obs}, was assessed using the unbiased and biased Nash–Sutcliffe model efficiency coefficients described by Nse_{unbias} [Equation (17)] and Nse_{bias} [Equation (18)], respectively, in Table 8. It is to be noted that for every site, Nse_{unbias} adds a constant value to θ_{obs} such that $\overline{\theta_{\text{obsAdj}}} = \overline{\theta_{\text{hypix}}}$, but as shown in Figure 4

it does not make corrections to match the variance between θ_{obsAdj} & θ_{hypix}. The $N_{\text{se\_unbias}}$ shows that the performance of six sites was rated as 'Excellent' ($N_{\text{se\_unbias}}$ [0.8–1]); eight sites were rated 'Good', (N_{se} [0.6–0.7]); and 10 sites performed *poorly* ($N_{\text{se\_unbias}} < 0.6$). Understanding why some sites do not perform well presents opportunities to improve soil descriptions, hydraulic functions, or the models and to provide explanations for why the observations are giving unexpected results. However, there are several challenges to elucidating the issues of the mismatch.

- Installation of the Aquaflex sensors required disturbing and repacking the soils and stones needed to be removed to ensure good contact between the sensor and the soil particles. Poesen & Lavee (1994) showed that soils with high R_{f} are more likely to have large voids between the rock fragments not entirely filled by fine material, which prevents the surrounding soil matrix from compacting (Nasri et al., 2015). This fragile network of voids may be destroyed when stony soils are repacked or reconstituted (Poesen & Lavee, 1994), causing disturbances in the hydraulic properties,
- The Aquaflex sensors give an average value of $\overline{\theta}$ for the top 40 cm and therefore do not provide information on the dynamics of θ at different depths,
- The soil water content sensors were not calibrated and therefore, only the differences in θ are accurate.

4.2.1 | Goodness-of-fit of the unbiased Smap-Hydro

For every monitoring site, pedological descriptions from S-map are provided in Table 8. The pedological information consists of *soil order*, *drainage class*, *texture*, *soil depth*, *water table* description and the R_{f} in the top 40 cm. Nevertheless, the descriptions may not reflect the actual soil description of the monitoring sites. Finding general trends with soil descriptions and the fit of simulated θ by using the HyPix model (Poor [Grey]/Good [Blue]/ Excellent [Red]) for the unbiased estimates is not straightforward. The following observations can be deduced from Table 8.

- *Very stony soils*: seven of the sites described in Table 8, have soil horizons with >60% rock fragments within the top 40 cm. These sites present a major challenge for good aquaflex contact and even more, the soil moisture probes have not been adjusted to account for the stone-probe contacts.

TABLE 8 S-map pedological description of every monitoring site: *order, drainage class, texture, depth, water table,* depth of roots (depending on whether there is a root barrier) and the highest percentage of rock fragments (R_f) within the top 40 cm together with average yearly precipitation (Pr); Aridity Index (AI) computed as Pr/P_{et}; water balance components, computed with HyPix model and Smap-Hydro parameters, simulated drainage (D_{rain}); evapotranspiration (S_{ink}); and root-zone soil water content (θ_{root}) which is the θ for the average top 40 cm. The goodness-of-fit between model outputs and observations, normalised to the average of the top 40 cm [Equation (15)], is evaluated with the Nash–Sutcliffe efficiency coefficient (N_{se}) for biased and unbiased estimates.

Site	Plot.	Soil order	Drainage class	Texture	Water table	R_f [%]	Pr [mm year⁻¹]	P_{et} [mm year⁻¹]	AI [%]	S_{ink} [mm year⁻¹]	D_{rain} [mm year⁻¹]	θ_{root} [%]	N_{se_unbias} [0-1]	ΔN_{se_Rf} [0-1]	N_{se_bias} [0-1]
Hamilton	A	Allophanic	imperfect	loamy	No	0	128	146	88	84	40	31	0.8	-	0.7
Stratford	–	Allophanic	well drained	loamy	No	13	1894	958	198	557	1151	35	0.7	-0.01	-0.1
Dannevirke	B	Brown	well drained	loamy	No	12	926	986	94	494	302	39	0.7	0	-0.2
Dargaville	–	Brown	well drained	loamy	No	0	1026	1132	91	604	271	32	0.6	-	0.1
LakeTekapo	–	Brown	well drained	sandy	No	60	580	1081	54	274	224	14	0.2	-0.01	0.2
Wallaceville	–	Brown	well drained	loamy	No	65	1450	1036	140	373	902	31	0.1	-0.09	-2.9
Darfield	C	Pallic	well drained	loamy	No	63	585	868	67	388	112	32	0.8	0.2	0.1
Winchmore	–	Pallic	well drained	loamy	No	65	759	959	79	457	198	34	0.8	0.4	0.4
Awatere	–	Pallic	well drained	loamy	No	0	559	1134	49	425	41	23	0.8	-	0.0
Middlemarch	–	Pallic	well drained	loamy	No	63	491	930	53	311	101	20	0.4	-0.03	-1.4
Lincoln	–	Pallic	well drained	loamy	No	0	603	1032	58	351	171	11	0.5	-	0.5
Dunedin	D	Pallic	imperfect	loamy	No	25	730	862	85	459	161	18	0.6	0.01	-5.0
Martinborough	–	Pallic	imperfect	loamy	No	0	749	1096	68	327	313	10	0.2	-	0.1
Windsor	–	Pallic	imperfect	loamy	No	0	553	894	62	405	72	15	0.6	-	0.2
Balclutha	–	Pallic	poor	loamy	No	0	677	853	79	443	120	36	0.6	-	-0.5
Waipawa	E	Recent	imperfect	loamy	No	7	748	1081	69	435	216	25	0.8	0	0.1
Hanmer	–	Recent	well drained	loamy	No	55	950	907	105	295	541	15	0.1	-0.05	-1.6
Cromwell	F	Semi-arid	well drained	sandy	No	0	390	1054	37	283	39	5	0.2	-	0.1
FranzJosef	–	Recent	well drained	sandy	No	68	1111	248	449	95	910	23	0.4	-0.09	-10.
Greymouth	–	Recent	well drained	sandy	No	85	2397	847	283	274	1904	21	0.2	-0.11	-31.
Paraparaumu	–	Recent	well drained	sandy	No	0	1029	1016	101	361	537	6	0.3	-	0.1
Pukekohe	G	Granular	well drained	clayey	No	15	366	292	125	170	154	54	0.9	0.01	0.6
Warkworth	H	Ultic	poor	clayey	No	0	386	313	123	185	169	57	0.7	-	0.4
Rangiora	I	Gley	poor	clayey	Yes	45	564	1056	53	386	99	38	0.7	0.14	0.3

Note: The fit is classified as Poor [Grey]/Good [Blue]/Good [Blue]/Excellent [Red]. ΔN_{se_R} is the improvement made when the hydraulic parameters are corrected for R_f. The N_{se_unbias} for simulations without correcting for rock fragments would be N_{se_unbias}-ΔN_{se_Rf}. *Plot* represents the selected monitoring sites which are plotted in Figure 4.
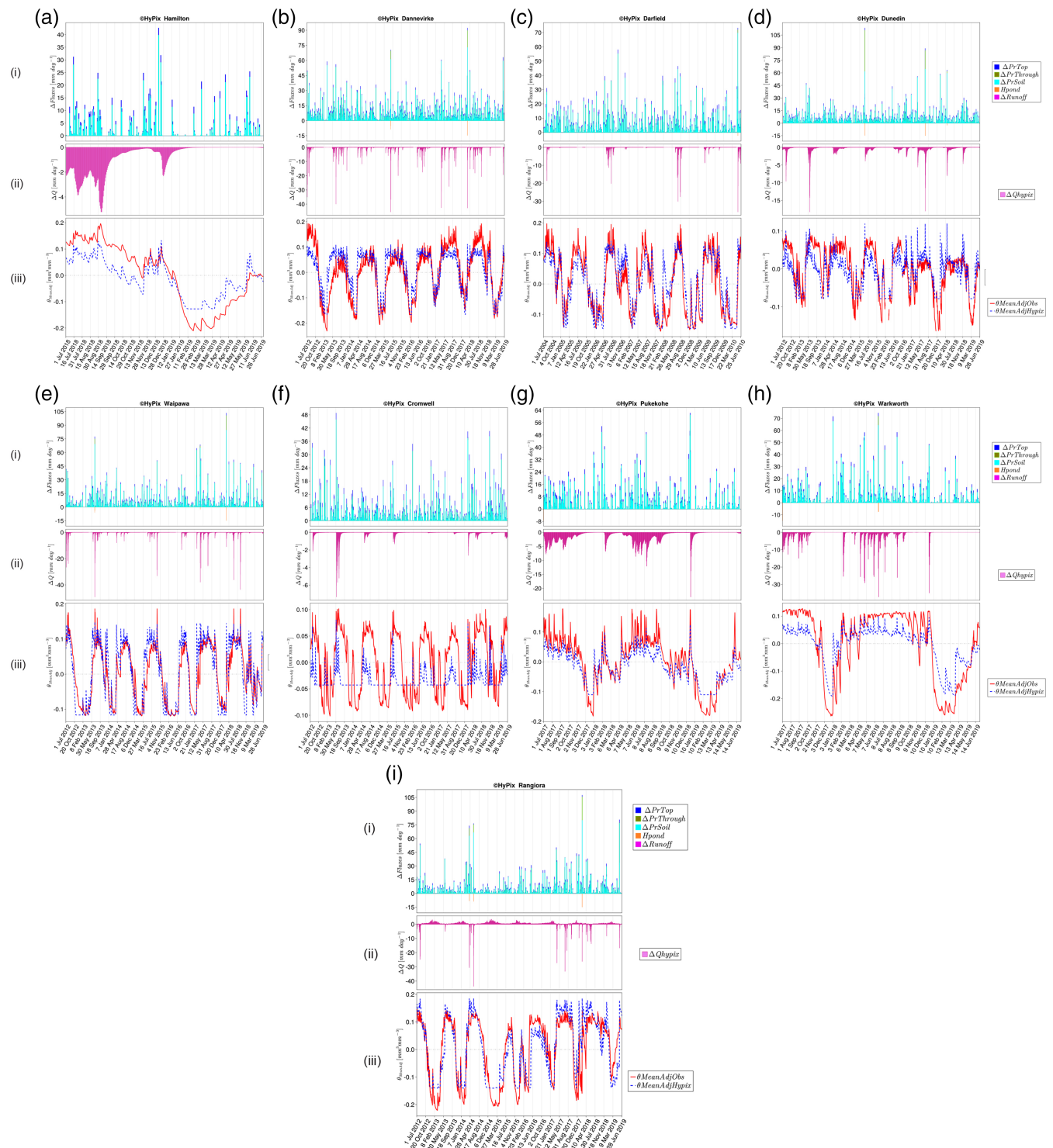
FIGURE 4   For selected monitoring sites representing different texture and soil orders described in Table 8, time series plots of: (i) precipitation at the top of the vegetation ($\Delta PrTop$); throughfall precipitation ($\Delta PrThrough$); precipitation infiltrating into the soil ($\Delta PrSoil$); ponding ($\Delta Hpond$); and runoff ($\Delta Runoff$); (ii) drainage at the bottom of the root zone ($\Delta Q_{\mathrm{Hypix}}$); and (iii) corrected mean values of the soil water content observations ($\theta_{\mathrm{MeanAdjObs}}$), and HyPix simulations ($\theta_{\mathrm{MeanAdjHypix}}$).

- *Sandy soils* do not perform as well as *Loamy* and *Clay* textured soils. This result is plausible because these soils are somewhat less common in New Zealand and so when deriving θ and ψ points from PTFs (Section 3.2.2), there

is a limited dataset for *Sandy soils* from which to develop the PTFs.

- There seems to be no straightforward relationship between $N_{\mathrm{se\_unbias}}$ and the *soil order*, except that *Granular*

(clayey soils formed from strongly weathered volcanic rocks or ash), *Allophanic* (dominated by allophane [and imogolite or ferrihydrite] minerals), *Semi-arid*, *Ultic* (strongly weathered soils developed on quartz-rich rocks with a well-structured, clay-enriched sub-soil horizon) and *Gley* soils seem to give a better performance. Nevertheless, there was a mixed performance for *Brown soils* (thin coatings of iron oxides weathered from the parent material cause the brown colour) and bad performance for *Pallic soils* (low content of iron oxides, weak soil structure and high density in sub-surface horizons).

- There are no obvious relationships in terms of performance between *drainage class* and *depth*.

4.2.2 | Implication of validating Smap-Hydro with uncalibrated θ

The drawback that the sensors were not calibrated for the local soil conditions is demonstrated by the poor performance of N_{se_Bias} described in Table 8. This implies that only differences of θ and fluxes with time can be trusted and not the actual absolute θ_{obs} value. This means that $\overline{\theta_{Hypix}}$ and $\overline{\theta_{Obs}}$ suffers from a random offset. The question is how this offset impacts the validation of Smap-Hydro.

For soils that HyPix can model accurately, the change of θ with time implies that the slope of $\theta(\psi)$ and $K(\psi)$ is correctly modelled by σ and ψ_m and also by the magnitude of K_s which is a sensitive parameter to regulate the flux rate. Nevertheless, we are unable to validate Smap-Hydro effective porosity, θ_{sr}, which is $\theta_s - \theta_r$. This raises a question for sites that HyPix models the dynamics of θ *good* and *excellent*, can we speculate that $\overline{\theta_{Hypix}}$ derived from Smap-Hydro is more accurate than uncalibrated $\overline{\theta_{Obs}}$. This hypothesis is put forward because in Smap-Hydro, K_s is derived from σ and particularly from θ_{sr} through K_{sModel} [Equation (4)]. Therefore, if θ_{sr} is uncertain then K_s is uncertain and therefore the θ dynamics cannot be modelled accurately. Resolution of this interesting question is beyond the scope of the current study but is the subject of future research.

4.2.3 | Modelling soils with high content of rock fragments

The R_f reported in Table 8 represents the top 40 cm viewed by S-map but may not represent the actual R_f at the monitoring sites for which there was no information. The correction of $\theta(\psi)$ for R_f is described in Section 2.4 and the correction of $K(\psi)$ for R_f is described in

Section 2.4.2. The N_{se_unbias} improvements for correcting for R_f are described in Table 8 by ΔN_{se_Rf}. Therefore, the N_{se_unbias} for simulations without correcting for R_f would be N_{se_unbias}- ΔN_{se_Rf}. Only soils with $R_f > 15\%$ seem to have a hydrological impact. The improvements for correcting for R_f are particularly notable for sites with $N_{se\_unbias} > 0.7$ in Table 8. This indicates that the correction for $R_f > 40\%$ of K_s (Figure 1), which accounts for the altered flow of water for soils with high R_f (which are no longer dominated by matrix flow but by film flow due to fragment-to-fragment contact), makes improvements but would require further validation. Sites with $N_{se\_unbias} < 0.6$ that do not have a good match between observed and simulated θ did not benefit from the correction for R_f.

5 | FUTURE IMPROVEMENT OF COMBINING PTFS WITH SOIL PHYSICS

The following research and implementation could be undertaken to improve the predictions of Smap-Hydro through further enhancement of the synergy between soil physics and existing empirical approaches.

McNeill et al. (2018) have derived spatial $\theta(\psi)$ data points for a large part of New Zealand using explanatory variables from the S-map soil information system of New Zealand using empirical PTFs. The explanatory variables include the soil *order classification*, *sample depth*, *texture* (*sand, silt and clay estimates*) and parsed information from the S-map functional horizon description. While successful for predictions of the response $\theta(\psi)$ *and* especially for derived quantities such as the available water (the difference in $\theta(\psi)$ for two adopted tensions ψ), this approach does not encapsulate the causal response that would be available in a physical model. In addition, it is typical in empirical approaches to use the explanatory variables available from local soil descriptions, which means that superfluous variables can be included, while important physical explanatory variables are not available. For instance, the inclusion of sample depth in the empirical model does not mean that sample depth per se is causally relevant in modelling $\theta(\psi)$; most likely the beneficial effect of this explanatory variable is due to the strong association with carbon concentration, which is associated with changes in $\theta(\psi)$.

Replacing the present empirical model for $\theta(\psi)$ with a physical model requires a modelling framework that includes the same S-map explanatory variables that are currently used. Since these explanatory variables are

not included in physically-based explanations of $\theta(\psi)$, the most successful approach is likely to be a Bayesian framework, where the available explanatory variables and known physical constraints of the hydraulic parameters can be used, as described in Table 1. The Bayesian approach permits this prior knowledge to be incorporated into the modelling, either by use of special priors, or by using an explicit modelling language (e.g., Stan, 2023).

6 | CONCLUSIONS

This paper shows that the physical hydraulic parameters describing $\theta(\psi)$ and $K(\psi)$ for up to six functional horizons to a maximum depth of 1 m can be derived by combining empirical PTF and principles of soil physics. This digital database, termed Smap-Hydro, which covers 40% of New Zealand's soil area in 2023. The steps involved in the development of this database are described in detail and the validation using a set of monitoring sites is presented.

Smap-Hydro performs the correction of $\theta(\psi)$ for R_f and ensures the derived hydraulic parameters are physically and dynamically constraining by using algorithms described in Table 1. The derived K_s is produced from principles of soil physics using explanatory hydraulic parameters describing $\theta(\psi)$ and from the percentage of *clay*. This model for K_s introduces a novel correction for R_f. Water-flow processes in the presence of high R_f are no longer dominated by matrix flow but by film flow due to fragment-to-fragment contact, as depicted in Figure 2. It was shown that this model improves the predictions for soils with high R_f. Nevertheless, this model would require further validation.

The accuracy of Smap-Hydro was assessed by introducing derived Smap-Hydro hydraulic parameters into the physically based HyPix model. Modelled and observed soil water content averaged for the top 40 cm for 24 sites across New Zealand was compared. An excellent fit was shown for 25% of the sites and a good fit for 33% of the sites. In general, fine-textured soils performed better than the coarser-textured soils, while organic soils did not perform well.

Improving Smap-Hydro is challenging, mainly because S-map information is not collected precisely at the monitoring sites. Moreover, installation of the measurement devices required disturbing the soils and, in many cases, the removal of rock fragments to ensure better contact with the sensors. Therefore, future validation and improvements would require a larger number of sites for which calibrated time series θ data are available at different depths in the vicinity of sites where S-map information is collected.

This paper highlights the need for the development of novel physico-PTFs by synergising empirical relationships with knowledge gained from soil physics in the last decade.

AUTHOR CONTRIBUTIONS

J. A. P. Pollacco: Software; writing – original draft; writing – review and editing; visualization; investigation; conceptualization; methodology; validation; formal analysis; project administration; supervision. **J. Fernández-Gálvez:** Writing – review and editing; writing – original draft; software; conceptualization; investigation; methodology; validation; formal analysis. **T. Webb:** Writing – original draft; methodology; data curation. **S. Vickers:** Writing – original draft; methodology; software. **B. Robertson:** Writing – original draft. **S. McNeill:** Writing – original draft; methodology. **L. Lilburne:** Methodology; writing – original draft. **C. Rajanayaka:** Resources; data curation. **H. W. Chau:** Writing – review and editing.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY STATEMENT

Data are available from the corresponding author upon reasonable request. The free version of S-map can be obtained from https://smap.landcareresearch.co.nz/.

ORCID

J. A. P. Pollacco https://orcid.org/0000-0003-3622-6135

J. Fernández-Gálvez https://orcid.org/0000-0003-2607-8896

S. Vickers https://orcid.org/0009-0006-0092-9604

B. Robertson https://orcid.org/0000-0001-9861-1130

S. McNeill https://orcid.org/0000-0003-1288-4387

L. Lilburne https://orcid.org/0000-0001-8886-6754

C. Rajanayaka https://orcid.org/0000-0002-3318-2264

H. W. Chau https://orcid.org/0000-0002-9411-9816

REFERENCES

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. FAO, Rome, 300, D05109.

Arrouays, D., McBratney, A., Bouma, J., Libohova, Z., Richer-de-Forges, A. C., Morgan, C. L. S., Roudier, P., Poggio, L., & Mulder, V. L. (2020). Impressions of digital soil maps: The good, the not so good, and making them ever better. *Geoderma Regional*, *20*, e00255.

Beckers, E., Pichault, M., Pansak, W., Degré, A., & Garré, S. (2016). Characterization of stony soils' hydraulic conductivity using laboratory and numerical experiments. *The Soil*, *2*, 421–431.

Beibei, Z., Ming'an, S., & Hongbo, S. (2009). Effects of rock fragments on water movement and solute transport in a Loess Plateau soil. *Comptes Rendus Geoscience*, *341*, 462–472.

Brooks, R. H., & Corey, A. T. (1964). Hydraulic properties of porous media. In *Hydrology Papers 3* (Vol. 27, p. 3). Colorado State University.

Cao, W., Bowden, W. B., Davie, T., & Fenemor, A. (2009). Modelling impacts of land cover change on critical water resources in the Motueka River catchment, New Zealand. *Water Resources Management*, *23*, 137–151.

Clapp, R. B., & Hornberger, G. M. (1978). Empirical equations for some soil hydraulic properties. *Water Resources Research*, *14*, 601–604. https://doi.org/10.1029/WR014i004p00601/abstract

Coppola, A., Dragonetti, G., Comegna, A., Lamaddalena, N., Caushi, B., Haikal, M. A., & Basile, A. (2013). Measuring and modeling water content in stony soils. *Soil and Tillage Research*, *128*, 9–22.

Dann, R., Close, M., Flintoft, M., Hector, R., Barlow, H., Thomas, S., & Francis, G. (2009). Characterization and estimation of hydraulic properties in an alluvial gravel Vadose zone. *Vadose Zone Journal*, *8*, 651–663.

Evans, P. S. (1978). Plant root distribution and water use patterns of some pasture and crop species. *New Zealand Journal of Agricultural Research*, *21*, 261–265.

Feddes, R. A., Kowalik, P. J., & Zaradny, H. (1978). Simulation of field water use and crop yield. In *Simulation Monographs* (p. 189). Wageningen.

Fernández-Gálvez, J., Pollacco, J. A. P., Lilburne, L., McNeill, S., Carrick, S., Lassabatere, L., & Angulo-Jaramillo, R. (2021). Deriving physical and unique bimodal soil Kosugi hydraulic parameters from inverse modelling. *Advances in Water Resources*, *153*, 103933.

Fernández-Gálvez, J., Simmonds, L. P., & Barahona, E. (2006). Estimating detailed soil water profile records from point measurements. *European Journal of Soil Science*, *57*, 708–718.

Giltrap, D. L., Pollacco, J. A. P., Graham, S., Carrick, S., & Lilburne, L. (2022). Multi-layer soil hydrology modelling in OVERSEER.

Green, W. H., & Ampt, G. (1911). Studies on soil physics, 1. The flow of air and water through soils. *Journal of Agricultural Science*, *4*, 1–24.

Griffiths, E., Webb, T. H., Watt, J. P. C., & Singleton, P. L. (1999). Development of soil morphological descriptors to improve field estimation of hydraulic conductivity. *Australian Journal of Soil Research*, *37*, 971–982.

Hewitt, A. (1992). *New Zealand soil classification*. DSIR.

Hewitt, A. E. (2010). New Zealand soil classification.

Hewitt, A. E., Balks, M. R., & Lowe, D. J. (2021). *The soils of Aotearoa New Zealand*. Springer.

Hlaváčiková, H., Novák, V., & Šimůnek, J. (2016). The effects of rock fragment shapes and positions on modeled hydraulic conductivities of stony soils. *Geoderma*, *281*, 39–48.

Jarvis, N. J. (2007). A review of non-equilibrium water flow and solute transport in soil macropores: Principles, controlling factors and consequences for water quality. *European Journal of Soil Science*, *58*, 523–546.

Khetdan, C., Chittamart, N., Tawornpruek, S., Kongkaew, T., Onsamrarn, W., & Garré, S. (2017). Influence of rock fragments on hydraulic properties of Ultisols in Ratchaburi Province, Thailand. *Geoderma Regional*, *10*, 21–28.

Kosugi, K. (1994). Three-parameter lognormal distribution model for soil water retention. *Water Resources Research*, *30*, 891–901.

Kosugi, K. (1996). Lognormal distribution model for unsaturated soil hydraulic properties. *Water Resources Research*, *32*, 2697–2703.

Kosugi, K. (1999). General model for unsaturated hydraulic conductivity for soils with lognormal pore-size distribution. *Soil Science Society of America Journal*, *63*, 270–277.

Lassabatere, L., Peyneau, P.-E., Yilmaz, D., Pollacco, J., Fernández-Gálvez, J., Latorre, B., Moret-Fernández, D., Di Prima, S., Rahmati, M., Stewart, R. D., Abou Najm, M., Hammecker, C., & Angulo-Jaramillo, R. (2021). A scaling procedure for straightforward computation of sorptivity. *Hydrology and Earth System Sciences*, *25*, 5083–5104.

Lassabatere, L., Peyneau, P.-E., Yilmaz, D., Pollacco, J., Fernández-Gálvez, J., Latorre, B., Moret-Fernández, D., Di Prima, S., Rahmati, M., Stewart, R. D., Abou Najm, M., Hammecker, C., & Angulo-Jaramillo, R. (2023). Mixed formulation for an easy and robust numerical computation of sorptivity. *Hydrology and Earth System Sciences*, *27*, 895–915.

Lilburne, L., Hewitt, A., Webb, T. H., & Carrick, S. (2004). S-map: a new soil database for New Zealand. In: *SuperSoil 2004: Proceedings of the 3rd Australian New Zealand Soils Conference, Sydney, Australia, 5-9 Dec 2004*.

Ma, D., & Shao, M. (2008). Simulating infiltration into stony soils with a dual-porosity model. *European Journal of Soil Science*, *59*, 950–959. https://doi.org/10.1111/j.1365-2389.2008.01055.x

Ma, D., Shao, M., Zhang, J., & Wang, Q. (2010). Validation of an analytical method for determining soil hydraulic properties of stony soils using experimental data. *Geoderma*, *159*, 262–269.

Maff. (1988). *Agricultural land classification of England and Wales*.

McLeod, M., Aislabie, J., Ryburn, J., & McGill, A. (2008). Regionalizing potential for microbial bypass flow through New Zealand soils. *Journal of Environmental Quality*, *37*, 1959–1967.

McNeill, S. J., Lilburne, L. R., Carrick, S., Webb, T., & Cuthill, T. (2018). Pedotransfer functions for the soil water characteristics of New Zealand soils using S-map information. *Geoderma*, *326*, 96–110.

Mehuys, G. R., Stolzy, L. H., Letey, J., & Weeks, L. V. (1975). Effect of stones on the hydraulic conductivity of relatively Dry Desert Soils. *Soil Science Society of America Journal*, *39*, 37–42.

Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, *12*, 513–522.

Naseri, M., Iden, S. C., & Durner, W. (2022). Effective hydraulic properties of 3D virtual stony soils identified by inverse modeling. *The Soil*, *8*, 99–112.

Naseri, M., Joshi, D. C., Iden, S. C., & Durner, W. (2023). Rock fragments influence the water retention and hydraulic conductivity of soils. *Vadose Zone Journal*, *22*, e20243.

Nasri, B., Fouché, O., & Torri, D. (2015). Coupling published pedotransfer functions for the estimation of bulk density and saturated hydraulic conductivity in stony soils. *Catena*, *131*, 99–108. http://www.sciencedirect.com/science/article/pii/S0341816215001034

Novák, V., Kňava, K., & Šimůnek, J. (2011). Determining the influence of stones on hydraulic conductivity of saturated soils using numerical method. *Geoderma*, *161*, 177–181.

Pakparvar, M., Cornelis, W., Gabriels, D., Mansouri, Z., & Kowsar, S. A. (2016). Enhancing modelled water content by dielectric permittivity in stony soils. *Soil Research*, *54*, 360.

Parajuli, K., Sadeghi, M., & Jones, S. B. (2017). A binary mixing model for characterizing stony-soil water retention. *Agricultural and Forest Meteorology*, *244–245*, 1–8.

Peck, A. J., & Watson, J. D. (1979). CSIRO Research Publications Repository—Hydraulic conductivity and flow in non-uniform soil. https://publications.csiro.au/rpr/pub?list=BRO&pid=procite:f4500654-39c9-4e75-a12c-37c8ea9e4508

Peters, R. R., & Klavetter, E. A. (1988). A continuum model for water movement in an unsaturated fractured rock mass. *Water Resources Research*, *24*, 416–430.

Poesen, J., & Lavee, H. (1994). Rock fragments in top soils: Significance and processes. *Catena*, *23*, 1–28.

Pollacco, J. A. P., & Angulo-Jaramilo, R. (2009). A linking test that investigates the feasibility of inverse modelling: Application to a simple rainfall interception model for Mt Gambier, southeast South Australia. *Hydrological Processes*, *23*, 2023–2032.

Pollacco, J. A. P., Braud, I., Angulo-Jaramillo, R., & Saugier, B. (2008). A linking test that establishes if groundwater recharge can be determined by optimising vegetation parameters against soil moisture. *Annals of Forest Science*, *65*, 702.

Pollacco, J. A. P., Fernández-Gálvez, J., Ackerer, P., Belfort, B., Lassabatere, L., Angulo-Jaramillo, R., Rajanayaka, C., Lilburne, L., Carrick, S., & Peltzer, D. A. (2022). HyPix: 1D physically based hydrological model with novel adaptive time-stepping management and smoothing dynamic criterion for controlling Newton–Raphson step. *Environmental Modelling & Software*, *154*, 105386.

Pollacco, J. A. P., Fernandez-Galvez, J., & Carrick, S. (2020). Improved prediction of water retention curves for fine texture soils using an intergranular mixing particle size distribution model. *Journal of Hydrology*, *584*, 124597.

Pollacco, J. A. P., Fernández-Gálvez, J., & de Jong van Lier, Q. (2023). Bimodal unsaturated hydraulic conductivity derived from water retention parameters by accounting for clay-water interactions: Deriving a plausible set of hydraulic parameters. *Journal of Hydrology*, *626*, 130227.

Pollacco, J. A. P., Fernández-Gálvez, J., Rajanayaka, C., Zammit, S. C., Ackerer, P., Belfort, B., Lassabatere, L., Angulo-Jaramillo, R., Lilburne, L., Carrick, S., & Peltzer, D. A. (2022). Multistep optimization of HyPix model for flexible vertical scaling of soil hydraulic parameters. *Environmental Modelling & Software*, *156*, 105472.

Pollacco, J. A. P., Nasta, P., Soria-Ugalde, J. M., Angulo-Jaramillo, R., Lassabatere, L., Mohanty, B. P., & Romano, N. (2013). Reduction of feasible parameter space of the inverted soil hydraulic parameter sets for Kosugi model. *Soil Science*, *178*, 267–280.

Pollacco, J. A. P., Nasta, P., Ugalde, J. M. S., Angulo-Jaramillo, R., Lassabatere, L., Mohanty, B. P., & Romano, N. (2013). Reduction of feasible parameter space of the inverted soil hydraulic parameters sets for Kosugi model. *Soil Science*, *178*(6), 267–280.

Pollacco, J. A. P., Ugalde, J. M. S., Angulo-Jaramillo, R., Braud, I., & Saugier, B. (2008). A linking test to reduce the number of hydraulic parameters necessary to simulate groundwater recharge in unsaturated soils. *Advances in Water Resources*, *31*, 355–369.

Pollacco, J. A. P., Webb, T., Lilburne, L., & Vickers, S. (2022). S-map soil parameters for Soil & Water Assessment Tool catchment model. In *Contract Report: LC4156*. Landcare Research.

Pollacco, J. A. P., Webb, T., McNeill, S., Hu, W., Carrick, S., Hewitt, A., & Lilburne, L. (2017). Saturated hydraulic conductivity model computed from bimodal water retention curves for a range of New Zealand soils. *Hydrology and Earth System Sciences*, *21*, 2725–2737.

Qin, A. K., Huang, V. L., & Suganthan, P. N. (2009). Differential evolution algorithm with strategy adaptation for global numerical optimization. *IEEE Transactions on Evolutionary Computation*, *13*, 398–417.

Rajanayaka, C., Woodward, S. J. R., Lilburne, L., Carrick, S., Griffiths, J., Srinivasan, M. S., Zammit, C., & Fernández-Gálvez, J. (2022). Upscaling point-scale soil hydraulic properties for application in a catchment model using Bayesian calibration: An application in two agricultural regions of New Zealand. *Frontiers in Water*, *4*, 986496.

Robertson, B., Gillespie, J., Carrick, S., Almond, P., Payne, J., Chau, H., & Smith, C. (2021). Measuring the water retention curve of rock fragments: A novel repacked core methodology. *European Journal of Soil Science*, *73*(1), e13181.

Robertson, B. B., Almond, P. C., Carrick, S. T., Penny, V., Eger, A., Chau, H. W., & Smith, C. M. S. (2021). The influence of rock fragments on field capacity water content in stony soils from hard sandstone alluvium. *Geoderma*, *389*, 114912.

Schoeman, J. L., Kruger, M. M., & Loock, A. H. (1997). Water-holding capacity of rock fragments in rehabilitated opencast mine soils. *South African Journal of Plant and Soil*, *14*, 98–102.

Shao, M. A., Ma, D. H., Zhu, Y. J., Wang, H. F., & Zhou, B. B. (2009). *Soil water of stony soil on the Chinese Loess Plateau*. Science Press.

Sood, A., Lilburne, L., Carrick, S., Yang, Y., Raghav, S., Pollacco, J., & Porteous, A. (2021). Improving the New Zealand drought index.

Srinivasan, M. S., Muirhead, R. W., Singh, S. K., Monaghan, R. M., Stenger, R., Close, M. E., Manderson, A., Drewry, J. J., Smith, L. C., Selbie, D., & Hodson, R. (2021). Development of a

national-scale framework to characterise transfers of N, P and Escherichia coli from land to water. *New Zealand Journal of Agricultural Research*, *64*, 286–313.

Stan, D. T. (2023). Stan User's Guide, Version 2.32. https://mc-stan.org

Szabó, B., Weynants, M., & Weber, T. K. D. (2021). Updated European hydraulic pedotransfer functions with communicated uncertainties in the predicted variables (euptfv2). *Geoscientific Model Development*, *14*, 151–175.

Tait, A., Henderson, R., Turner, R., & Zheng, X. (2006). Thin plate smoothing spline interpolation of daily rainfall for New Zealand using a climatological rainfall surface: Estimation of daily rainfall over New Zealand. *International Journal of Climatology*, *26*, 2097–2115.

Tóth, B., Weynants, M., Nemes, A., Makó, A., Bilas, G., & Tóth, G. (2015). New generation of hydraulic pedotransfer functions for Europe. *European Journal of Soil Science*, *66*, 226–238.

Urbanek, E., & Shakesby, R. A. (2009). Impact of stone content on water movement in water-repellent sand. *European Journal of Soil Science*, *60*, 412–419.

Van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, *44*, 892–898.

Vogeler, I., Carrick, S., Lilburne, L., Cichota, R., Pollacco, J. A. P., & Fernandez-Galvez, J. (2021). How important is the description of soil unsaturated hydraulic conductivity values for simulating soil saturation level, drainage and pasture yield? *Journal of Hydrology*, *598*, 126257.

Vogeler, I., & Cichota, R. (2019). Compilation of a reference dataset of flux data from various lysimeter sites across New Zealand and APSIM modelling.

Webb, T. H. (2003). Identification of functional horizons to predict physical properties for soils from alluvium in Canterbury, New Zealand. *Soil Research*, *41*, 1005.

Webb, T. H., & Lilburne, L. R. (2011). Criteria for defining the soil family and soil sibling: The fourth and fifth categories of the New Zealand soil classification. *Landcare Research Science Series*. https://www.cabdirect.org/cabdirect/abstract/20113361839

Wesseling, J. G. (1991). CAPSEV: Steady state moisture flow theory, program description and user manual.

Wood, S. N. (2017). *Generalized additive models: An introduction with R* (2nd ed.). Chapman and Hall/CRC. https://www.taylorfrancis.com/books/9781498728348

Yang, Y., Wang, Q., & Zhuang, J. (2013). Estimating hydraulic parameters of stony soils on the basis of one-dimensional water absorption properties. *Acta Agriculturae Scandinavica, Section B—Soil & Plant Science*, *63*, 304–313.

Zeitfogel, H., Feigl, M., & Schulz, K. (2023). Soil information on a regional scale: Two machine learning based approaches for predicting saturated hydraulic conductivity. *Geoderma*, *433*, 116418.

Zhang, J., & Sanderson, A. C. (2009). JADE: Adaptive differential evolution with optional external archive. *IEEE Transactions on Evolutionary Computation*, *13*, 945–958.

Zhang, Y., Weihermüller, L., Toth, B., Noman, M., & Vereecken, H. (2022). Analyzing dual porosity in soil hydraulic properties using soil databases for pedotransfer function development. *Vadose Zone Journal*, *21*, e20227.

Zhang, Y., Zhang, M., Niu, J., Li, H., Xiao, R., Zheng, H., & Bech, J. (2016). Rock fragments and soil hydrological processes: Significance and progress. *Catena*, *147*, 153–166.