

Faculté des bioingénieurs

Does reforestation increase or deplete soil water storage?

Quantifying the effect of reforestation on the soil water balance: a HYDRUS-1D case study comparing a regenerated plot and an adjacent pasture in the Atlantic Forest, Brazil.

Auteur : JEANGOUT Lucas

Promoteur·rices : JAVAUX Mathieu (UCLouvain)
BORMA Laura (INPE)

Lecteurs : COUVREUR Valentin (UCLouvain)
VANCLOOSTER Marnik (UCLouvain)

Année académique 2025 - 2026

Mémoire de fin d'études présenté en vue de l'obtention du diplôme de Bioingénieur :
Master [120]: Bioingénieur en sciences et technologies de l'environnement, à
finalité spécialisée, orientation Sustainability engineering.

ACKNOWLEDGMENTS

Je tiens tout d’abord à exprimer ma sincère gratitude à Mathieu Javaux pour son encadrement attentif et ses critiques justes et constructives tout au long de ce mémoire. Ses conseils avisés et sa disponibilité ont grandement contribué à la qualité de ce travail.

Agradeço igualmente a Laura Borma pelo valioso acompanhamento que me ofereceu ao longo desta tese de mestrado, assim como pelo caloroso acolhimento que nos reservou a Charlotte e a mim durante a nossa estadia no Brasil. A sua generosidade tornou esta experiência ainda mais marcante.

Mes remerciements vont également à Marnik Vanclooster et Valentin Couvreur, qui ont accepté d’être les lecteurs de ce mémoire. Je suis convaincu que leur regard et leurs remarques seront d’un apport précieux.

Gostaria também de agradecer a toda a equipa do LabEcoh do INPE. Ao Breno e ao Daniel, pelo apoio essencial no trabalho de campo e pelos muitos momentos de boa disposição que tornaram os dias de terreno tão divertidos, mas também ao Daniel pela sua competência técnica. O meu estágio no Brasil permanecerá para sempre na minha memória como uma experiência inesquecível, sobretudo graças a vocês.

Je remercie chaleureusement Aurélien Lengrand, Basile Delvoie et Louis pour leur relecture attentive et leurs commentaires avisés.

Merci à Charlotte, sans qui ce stage au Brésil n’aurait jamais eu la même saveur.

À mes parents, sœurs mais aussi cousins, tantes et plus, merci pour votre soutien permanent et inconditionnel.

À ma copine, merci pour son soutien sans faille au quotidien, ainsi que tout le reste, dont son vélo.

Merci à la faculté d’avoir mis à notre disposition l’espace mémorant, dont j’ai pu pleinement profiter.

Merci aux copains du troisième pour la motivation et les bêtises, mais aussi sans qui cette fin de mémoire aurait été bien plus difficile.

Les Kotextipotes, merci de m’épauler au quotidien, et ce déjà depuis plus d’un an.

Valentin, Marie-Keza, Sacha, Sarah, Eloïse, Clémence, Tiffany et tous les autres, merci pour votre soutien durant cette dernière année et toutes celles avant.

Merci à tous ceux que je n’ai pas cités mais qui ont une place dans mon cœur.

Et enfin, merci à toutes les personnes qui ont croisé mon chemin au cours de mes études et qui ont contribué à les rendre exceptionnelles.

CRediT

	Lucas Jeangout	Autres contributeur·ice	Intelligence artificielle
Conceptualisation	×	M.J., L.B.	
Méthodologie	×	M.J., L.B.	
Investigation	×	M.J., L.B., LabEcoh, anciens mémorants	Claude, NotebookLM
Analyse formelle	×		Claude
Validation	×	M.J., L.B.	
Rédaction – version originale	×		
Rédaction – révision et édition	×	B.D, A.L.	Claude
Visualisation	×		Claude
Supervision		M.J., L.B.	

Mathieu Javaux (M.J.), Laura Borma (L.B.), Aurélien Lengrand (A.L.), Basile Delvoie (B.D.).

L'intelligence artificielle a été utilisée à plusieurs reprises dans le cadre de ce mémoire. Pour la recherche d'article, l'écriture de différents scripts d'analyse et de visualisation de données mais également dans le cadre de réécriture de texte en vue d'en améliorer la formulation. Pour autant, ce mémoire, que je suis fier de vous proposer, reste entièrement le fruit de mon travail.

Contents

I.	Introduction	1
II.	Context of the study	4
1	Geographical situation	4
2	Climate	4
3	Pedology	5
4	Vegetation	6
III.	State of the art	7
1	The water cycle of a tropical forest soil column	7
2	Soil hydraulic properties	9
2.1	The water retention curve	9
2.2	Unsaturated hydraulic conductivity	10
2.3	The Richards equation	10
2.4	Hysteresis	11
3	The influence of forest vegetation on soil hydraulic properties	12
3.1	Macroporosity	12
3.2	Organic matter, aggregation and soil structure	13
3.3	Water repellency	13
3.4	Implications for this study	13
4	Land use change and its hydrological effects	14
4.1	Deforestation	14
4.2	Reforestation	15
4.3	Consequences for the water balance	15
IV.	Objectives	17
V.	Materials and methods	18
1	Study plots	18
2	Meteorological data	19
2.1	Meteorological Station	19
2.1.1.	Quantile mapping	20
2.1.2.	Residual gaps	21
2.2	Reference and potential evapotranspiration	22

3	Soil properties	23
3.1	Field monitoring	23
3.2	Bulk density	24
3.3	Retention curves	24
3.4	Conductivity curves	26
3.5	Hysteresis	27
4	Vegetation data	27
4.1	Root density	27
4.2	Leaf area index	28
4.3	Actual evapotranspiration	29
4.4	Canopy interception	30
5	Water flow modeling	32
5.1	HYDRUS-1D presentation	32
5.1.1.	Spatial and temporal discretization	32
5.1.2.	Boundary conditions	32
5.1.3.	Initial conditions and simulation periods	33
5.2	Global sensitivity analysis	33
5.3	Scenarios analysis	34
5.3.1.	Root water uptake compensation	35
5.3.2.	Near-saturation retention behavior	35
5.3.3.	Canopy interception	35
5.4	Inverse modeling	35
5.4.1.	Parameters optimized	36
5.4.2.	Likelihood function	37
5.4.3.	Prior distributions	37
5.4.4.	Posterior sampling and propagation	38
5.5	Results expectation	39
VI.	Results and discussion	40
1	Collected data	40
1.1	Meteorological data	40
1.2	Soil data	41
1.2.1.	Volumetric water content	41
1.2.2.	Matric Potential	42

1.2.3.	Soil hydraulic properties	43
1.3	Vegetation parameters	45
1.3.1.	Root density	45
1.3.2.	Leaf area index	46
1.3.3.	Canopy interception	47
2	Sensitivity analysis	50
3	Structural model choices	51
4	Inverse modeling	54
5	Model validation	58
6	Simulated water balance of each plot	62
6.1	Term-by-term analysis	63
6.2	Full-column balance	65
7	Impact of reforestation on the soil water balance	66
7.1	Attribution to individual fluxes	67
7.2	Net storage change over the common window	68
7.3	Synthesis and answer to the research question	69
VII.	Conclusion	71
VIII.	Limits and perspectives	74
1	Limitations	74
2	Perspectives	75
	References	77

Appendix

A	Detailed Penman–Monteith equation for the hourly reference evapotranspiration	I
B	Labros fit parameter’s distribution	III
C	Gap filled meteorological series at São Francisco Xavier for the six measured variables.	IV
D	Leaf area index monthly box plots for both sites	V
E	Canopy interception visualization for both sites	VI
F	Morris sensitivity analysis heat maps for both sites	VIII
G	Bayesian inference post-to-posterior distribution	X
H	Posterior predictive simulations of VWC and matric potential	XIII

List of Figures

Introduction

- 1 Map of the original extent and mature forests contemporary remnants of the Atlantic Forest. 1

Context of the study

- 2 Map of Brazil showing the extent of the Atlantic Forest, the state of São Paulo, the city of São José dos Campos and the study area. 4
- 3 Climatograph of the Campos do Jordão station, as a representative reference for the regional highland climate of the Serra da Mantiqueira. 5

State of the art

- 4 Forest water balance diagram with the different fluxes composing it, each with an arrow indicating the relative direction of that flux. 7
- 5 Graph depicting the hysteresis effect on the retention curve. 11

Materials and methods

- 6 Map of the two study areas, R40 and Pasture, located in the São Francisco Xavier region, with the meteorological station indicated. 18
- 7 Picture of the R40 plot, focused on the canopy and its high density. 19
- 8 Picture of the São Francisco Xavier active meteorological station. 19
- 9 General shape of the water stress response function. 30

Results and discussion

- 10 Graph comparing the predicted and observed values for the six meteorological variables. 40
- 11 Graph of the precipitation and the potential evapotranspiration fluxes evaluated for the whole study period. 41
- 12 Volumetric water content measured at both sites for the study period. 42
- 13 Matric potential measured at both sites for the study period. 43
- 14 Laboratory retention curves by depth groups for the R40 site with the in situ measurements. 44
- 15 Laboratory retention curves by depth groups for the Pasture site with the in situ measurements. 45
- 16 Vertical soil profile of the fine root density g dm^{-3} for R40 and Pasture. 46
- 17 Time series of leaf area index m^2m^{-2} values for the R40 and the Pasture plots. 47

18	Goodness of fit of the calibrated Von Hoyningen-Hüne Braden interception model at R40 over the 18 pluviometer periods.	48
19	Gross rainfall and predicted canopy interception at R40, 2019–2025.	49
20	Posterior marginal distributions and pairwise correlations of the ten calibrated parameters at R40.	56
21	Posterior marginal distributions and pairwise correlations of the ten calibrated parameters at Pasture.	58
22	Simulated versus observed VWC (left) and matric potential in pF (right) for R40, pooling all observation depths over the full simulation period. Points are colored by depth. The dashed line is the 1:1 reference. Tight scatter around the 1:1 line is visible for most depths; the 10 cm layer shows underestimation of peak wet states, and the hysteretic loops in the pF panel confirm that the model captures the wetting–drying asymmetry.	60
23	Simulated versus observed VWC (left) and matric potential in pF (right) for Pasture, pooling all observation depths over the calibration period. Points are colored by depth. The 100, 150 and 200 cm VWC clusters above the 1:1 line illustrate the systematic between-depth offset that dominates the pooled R^2 . The compressed observed pF range (≈ 2.0 – 3.3 pF) reflects the Teros 21 sensor saturation near the wet limit, rendering the shallow pF scores near zero by construction rather than by model failure.	62
24	Daily ET_a/ET_p ratio for R40 and Pasture.	64
25	Cumulative water fluxes for R40 and Pasture.	65
26	Soil water storage of the 0–200 cm profile for R40 and Pasture.	66
27	Seasonal soil water balance and net storage change for R40 and Pasture.	67

Appendix

28	Gap filled meteorological series at São Francisco Xavier for the six measured variables.	IV
29	Box plots of the monthly leaf area index for R40 and Pasture.	V
30	Daily interception against daily precipitation at R40, with the fitted interception model.	VI
31	Gross rainfall and predicted canopy interception at Pasture, 2019–2025.	VI
32	Daily interception against daily precipitation at Pasture, with the fitted interception model.	VII

33	Heat map of the normalized μ^* of the Morris screening for every parameter across the four evaluated outputs for R40.	VIII
34	Heat map of the normalized μ^* of the Morris screening for every parameter across the four evaluated outputs for Pasture.	IX
35	Prior and posterior marginal distributions for the ten calibrated parameters at R40.	X
36	DREAM(ZS) sampling traces for the ten calibrated parameters at R40.	XI
37	DREAM(ZS) sampling traces for the ten calibrated parameters at Pasture	XI
38	Prior and posterior marginal distributions for the ten calibrated parameters at Pasture.	XII
39	Posterior predictive simulation of volumetric water content at R40, by depth. .	XIII
40	Posterior predictive simulation of matric potential (pF) at R40, by depth. . . .	XIV
41	Posterior predictive simulation of volumetric water content at Pasture, by depth.	XV
42	Posterior predictive simulation of matric potential (pF) at Pasture, by depth. . .	XVI

List of Tables

1	Sensor deployment and specifications at the R40 and Pasture sites.	23
2	Number of Hyprop and WP4C measurement pairs per depth at the R40 and Pasture sites.	25
3	Depth groups for the R40 and Pasture sites.	26
4	Morris sampling bounds for each of the ten parameter slots, applied identically to all four material layers.	34
5	Prior specifications for the eight optimized parameters.	38
6	Soil hydraulic parameters by depth group and by sites, obtained through laboratory measurements.	44
7	Overall ranks of the Morris screening of the parameters for both sites.	51
8	Coefficients of determination (R^2) for in situ VWC and pF, with and without the near-saturation air-entry regularization ($h_{\text{air}} = -2$ cm), for the R40 and Pasture sites.	52
9	Coefficients of determination (R^2) for in situ VWC and pF as a function of the critical stress index ω for root water uptake, for the R40 and Pasture sites.	52
10	Coefficients of determination (R^2) for in situ VWC and pF, and the resulting mean annual interception fraction, as a function of the canopy storage capacity parameter a , at the R40 site.	53
11	Posterior estimates of the eight van Genuchten–Mualem parameters and the two measurement-error hyperparameters for R40.	55
12	Posterior estimates of the eight van Genuchten–Mualem parameters and the two measurement-error hyperparameters for Pasture.	57
13	Coefficients of determination R^2 of the posterior-median simulation against the in-situ VWC and MP records.	59
14	Water-balance terms for R40 and Pasture over the common window.	68
15	Labros fit parameter’s distribution for both study sites.	III

List of abbreviations

CDF Cumulative distribution function

DREAM(ZS) Differential Evolution Adaptive Metropolis with Snooker updater

ET₀ Reference evapotranspiration

ET_a Actual evapotranspiration

ET_p Potential evapotranspiration

ESA European Space Agency

GSA Global sensitivity analysis

INMET The Brazilian National Institute of Meteorology

INPE Instituto Nacional de Pesquisas Espaciais, the Brazilian National Institute for Space Research

K_c Crop coefficient

LAI Leaf area index

MAD Median absolute deviation

MCMC Markov Chain Monte Carlo

MP Matric potential

QM Quantile mapping

R40 Central study area of the present work, constituting of a 40 years old regenerated forest area

RH Relative humidity

RTM Radiative transfer modeling

SFC Soil fraction covered

SFX São Francisco Xavier

VGM van Genuchten–Mualem

VWC Volumetric water content

I. Introduction

The Atlantic Forest (*Mata Atlântica*) is a tropical and subtropical forest domain that once extended in an almost continuous belt of about 1.5 million km² along the eastern coast of Brazil and into eastern Paraguay and north-eastern Argentina (Ribeiro et al., 2009). Despite its reduced present extent (Figure 1), it remains one of the most species-rich and most threatened biomes on Earth, and was among the first regions to be designated a global biodiversity hot spot on the joint criteria of exceptional endemism and exceptional habitat loss (Myers et al., 2000). It harbors on the order of 20,000 vascular plant species, roughly 40 % of which occur nowhere else, a level of endemism that reflects its long isolation from the other forest blocks of South America (Joly et al., 2014; Myers et al., 2000). Beyond its intrinsic biological value, the biome underpins a dense network of ecosystem services, from climate and water regulation to provisioning and cultural services (Diaz et al., 2018; Joly et al., 2014). This matters acutely because the Atlantic Forest coincides with the economic and demographic core of Brazil: it encompasses the catchments that supply the country's largest metropolitan regions and sustains roughly two-thirds of the country population and around three-quarters of gross domestic product (Joly et al., 2014; Ribeiro et al., 2009). Its forests are therefore not a peripheral conservation concern but a foundational pillar of the region's economic and social resilience.



Figure 1: Map of the original extent and mature forests contemporary remnants of the Atlantic Forest, produced by IPCI-Canada (2010).

Paradoxically, this strategic importance has coincided with one of the most severe trajectories of deforestation anywhere in the tropics. Five centuries of logging, agricultural expansion,

pasture conversion and urban growth have reduced the domain. Earlier estimates put remaining cover at 11 to 16 percent, including secondary forests and small fragments (Ribeiro et al., 2009), while higher-resolution mapping raised this to about 28 percent of its original cover (Rezende et al., 2018). What remains is highly fragmented, with the great majority of remnants smaller than 50 ha and embedded in an agricultural matrix (Ribeiro et al., 2009). Land-use change in the tropics, principally the replacement of forest by pasture and cropland, is among the largest anthropogenic sources of greenhouse-gas emissions and a major driver of soil carbon loss (Calvin et al., 2023; Don et al., 2011). At the land-surface scale it also reorganizes the partitioning of rainfall, typically lowering interception and evapotranspiration while degrading the soil structure that controls infiltration through compaction (Bruijnzeel, 2004).

Against this backdrop, forest restoration has become a central environmental policy, supported by national legislation, large-scale restoration pacts (Calmon et al., 2011), and decades of applied research on re-establishing diverse forest on abandoned land (Rodrigues et al., 2009). As active restoration and spontaneous regrowth proceed, secondary forests of varying age now make up a growing fraction of the remaining cover (Chazdon et al., 2016; Poorter et al., 2016). Yet the consequences of this regrowth for water resources are neither simple nor settled. Reforesting degraded land restores the macroporosity and structure that govern infiltration through root formation and the recovery of soil macro-fauna activity (Hassler et al., 2011; Ilstedt et al., 2007), but it simultaneously raises the evaporative demand returned to the atmosphere, so the net effect on the water that drains to streams and aquifers depends on which of these opposing influences dominates under the local soil and climate. This is the infiltration-evapotranspiration trade-off articulated by Bruijnzeel (2004). Reviews of the empirical evidence confirm that restoration can either decrease or, where infiltration gains are large, maintain or increase catchment water yield (Filoso et al., 2017; Ilstedt et al., 2007). Because this yield is the stream flow and aquifer recharge that supplies the metropolitan regions described above, whether restoration adds to or subtracts from it is not a secondary concern but a question of regional water security. Whether a given reforested system is a net water consumer or a net contributor therefore cannot be assumed: it has to be quantified.

The present work is a part of a broader collaborative project between the Instituto Nacional de Pesquisas Espaciais, the Brazilian National Institute for Space Research (INPE) and the Earth and Life Institute of UCLouvain. This collaboration is supervised by Laura Borma (INPE) and Mathieu Javaux (UCLouvain). This project addresses the above question at the study area of São Francisco Xavier (state of São Paulo), where adjacent plots differing only in land-use history allow a controlled comparison. Successive studies at this site have characterized individual components of the system: Delvoie (2023) established a first water balance description of the plots, Parmentier (2024) quantified evapotranspiration by remote sensing, Dermauw (2024) measured how the change of cover alters soil carbon, structure and hydraulic properties, and Oufini (2025) furthered the water balance evaluation while taking the root compensation into account. Each advanced the understanding of the ecohydrological behavior of the study area.

This work quantifies the hydrological value of reforestation at the study site by asking, over a complete hydrological year, whether the regenerated forest gains or loses soil water relative to the pasture, and through which fluxes that difference arises. The soil water balance of both plots is simulated in one dimension with the HYDRUS-1D model (Simunek et al., 1999), which solves the Richards equation under the van Genuchten–Mualem description of water retention and hydraulic conductivity curves. The hydraulic parameters are constrained by inverse estimation against in-situ water-content and matric-potential records. The model accounts for the hysteresis imposed by the alternation of wet and dry seasons and is coupled to the vegetation through root distribution, leaf area index, canopy interception and evapotranspiration. A global sensitivity analysis and a set of scenarios then isolate the parameters that most control the annual change in soil water stock. The remainder of this manuscript describes the study area, reviews the physical processes and prior knowledge underpinning the model, states the objectives, details the materials and methods, and finally presents and discusses the results.

II. Context of the study

1 Geographical situation

The study area lies in the rural district of São Francisco Xavier, within the municipality of São José dos Campos, in the north-east of the state of São Paulo, south-eastern Brazil (Figure 2). It is situated on the western flank of the Serra da Mantiqueira, a highland that forms part of the Atlantic Forest domain, and rises to elevations of about 1240 and 1180 meters a.s.l. respectively at the forested and pastured sites used in this study. The relief is markedly dissected, with steep, forested slopes alternating with cleared valley bottoms used for Pasture, a juxtaposition that is itself the expression of the land-use mosaic described above.

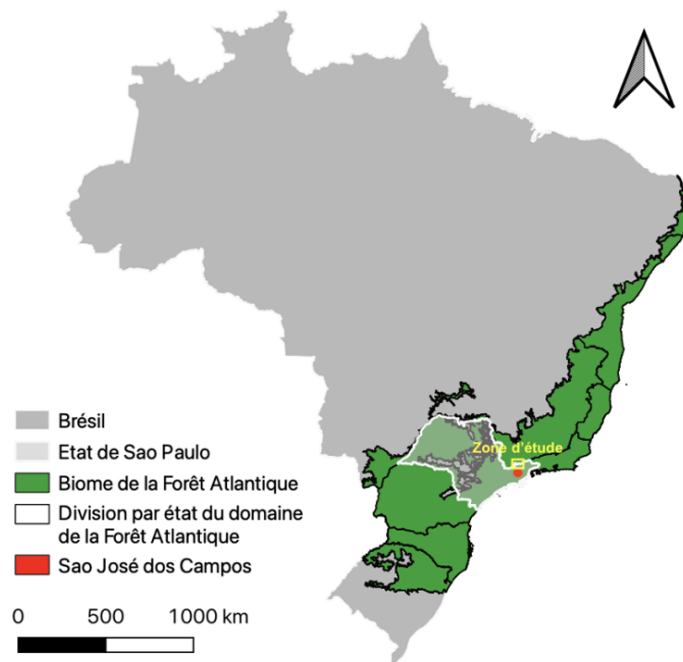


Figure 2: Map of Brazil showing the extent of the Atlantic Forest (green), the state of São Paulo (white outline), the city of São José dos Campos (red) and the study area (yellow). Realized by Delvoie (2023).

2 Climate

At the regional scale, the Serra da Mantiqueira highlands experience a humid subtropical highland climate, classified in the Köppen-Geiger system as Cwb, characterized by mild summers and a pronounced reduction of winter rainfall (Alvares et al., 2013). The defining feature for the present study is the strong seasonality of precipitation. A wet season that is concentrated in the austral summer (approximately October to April) delivers the large majority of annual rainfall, while a dry season in the austral winter (approximately May to September) imposes a sustained atmospheric water demand under comparatively low rainfall. This seasonal pattern is illustrated by the climate normals of the Campos do Jordão station (1642 m a.s.l.), the nearest long-record station of the Brazilian National Institute of Meteorology (INMET) to the study area and used here as a representative reference for the regional highland climate

(Figure 3). The wet season shows monthly precipitation totals exceeding 250 mm in January, while the dry season falls below 40 mm in June and August, imposing a sustained atmospheric water demand. Temperature follows the same seasonal rhythm, with a mean annual value of 14.5 °C, the warmest month in the austral summer (February, 17.7 °C) and the coldest in winter (July, 10.3 °C). The mean annual precipitation total of ≈ 1565 mm confirms the abundant but strongly seasonal water supply of the range. As Campos do Jordão (1642 m a.s.l.) lies above the study sites (1180 m to 1240 m a.s.l.), its normals serve as a regional reference for the seasonal regime rather than a precise local proxy. Absolute temperatures at the plots are expected to be somewhat higher. This pronounced seasonality between a wet summer and a dry winter is the primary climatic forcing shaping the soil water dynamics analysed in the remainder of this work.

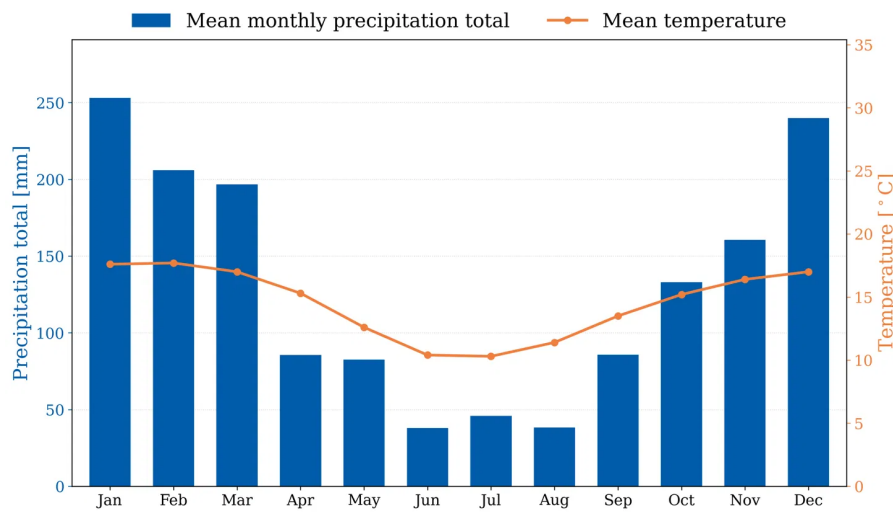


Figure 3: Climatograph of the Campos do Jordão station (Vila Capivari, 1642 m a.s.l.), the nearest long-record INMET station to the study area and a representative reference for the regional highland climate of the Serra da Mantiqueira. Blue bars show mean monthly precipitation totals (left axis) and the orange line shows mean monthly air temperature (right axis). The diagram displays the marked seasonality of the region, with a wet austral summer (October–April) and a dry austral winter (May–September). Mean annual precipitation is ≈ 1565 mm and mean annual temperature is 14.5 °C. This is not intended to represent the weather at the study sites, but rather to characterize the regional climate type. Data from “Instituto Nacional de Meteorologia - INMET” (n.d.), climatological normals (1991–2020), station 83714.

3 Pedology

The highlands of the Serra da Mantiqueira are developed on Precambrian crystalline basement (granites, gneisses and migmatites) deeply altered under a warm, humid climate (Ross & Moroz, 1996). The resulting regional soil cover is dominated by strongly weathered, acidic and dystrophic soils, principally deep, well-drained *Latosolos* (Ferralsols in the FAO-WRB system) on the more stable surfaces and shallower, more youthful *Cambissolos* (Cambisols) on the steeper slopes, both poor in nutrients and rich in kaolinite and in iron and aluminum oxides (Rossi, 2017; Schad et al., 2015). This oxide-rich mineralogy favors stable micro-aggregation and a high structural porosity (Ferreira et al., 1999; Pinheiro-Dick & Schwertmann, 1996)

4 Vegetation

The potential natural vegetation of the region belongs to the Atlantic Forest domain. Across the altitudinal gradient of the Serra da Mantiqueira the cover grades from montane dense ombrophilous forest (*Floresta Ombrófila Densa Montana*) into dense alto-montane forest and, at the highest and coldest summits, mixed *Araucaria*-bearing forest and altitude grasslands, the floristic composition turning over markedly with elevation (Meireles et al., 2008). These are structurally complex, evergreen to semi-evergreen forests with a dense, multi-layered canopy and high floristic diversity, and across the south-eastern highlands their composition shifts systematically with climate and elevation (Joly et al., 2014; Oliveira-Filho & Fontes, 2000). Centuries of clearing for timber, charcoal, agriculture and cattle have, however, replaced much of this cover with managed pasture and early successional vegetation, so that forest at different stages of regeneration and open herbaceous land now coexist across the landscape. This regional contrast between a structurally developed forest canopy and root system on the one hand, and a shallow herbaceous cover on the other, is precisely the difference in interception, transpiration and root water uptake that the plot-scale comparison developed in this work sets out to quantify.

III. State of the art

1 The water cycle of a tropical forest soil column

The hydrological behavior of a forest stand can be described as a one-dimensional vertical column in which incoming precipitation is progressively partitioned, stored, and returned to the atmosphere or to deeper layers. This is the conceptualization adopted by Simunek et al. (1999) in HYDRUS-1D and therefore the frame for the present work. The fluxes resolved in this model are: gross precipitation, canopy interception and the resulting net rainfall consisting of throughfall plus stem flow, evapotranspiration split in plant transpiration and evaporation, infiltration at the surface and drainage at the base of the profile, capillary rise, and surface runoff. They are visually represented in Figure 4. The objective of this section is to define each of these terms and to establish their order of magnitude in seasonally dry tropical forests.

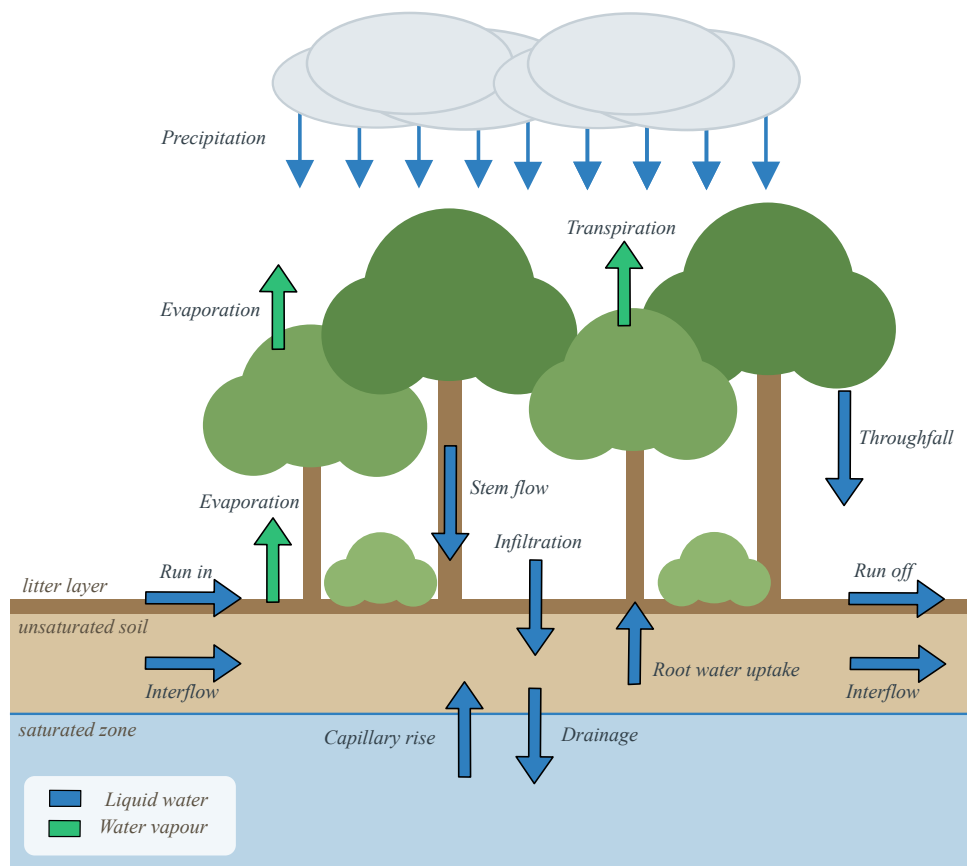


Figure 4: Forest water balance diagram with the different fluxes composing it, each with an arrow indicating the relative direction of that flux.

Interception is the first partition of incoming rainfall and a substantial loss term, in tropical rainforest it can remove roughly 10 to 20 % of gross rainfall before it reaches the soil, the fraction depending on canopy structure and storage capacity as well as on the size and intensity of events (Cuartas et al., 2007). Stem flow returns only a small share to the base of the trunks, about 2 % in a study conducted in a native Atlantic Forest plot (Ribeiro Filho et al., 2019).

Evapotranspiration is the dominant output of the annual water balance in humid and seasonally dry tropical forests, commonly returning half or more of gross precipitation to the atmosphere, the precise fraction varying with the rainfall regime as well as with canopy type and stand age (Bruijnzeel, 2004; Salemi et al., 2013). These evergreen forests are notable for sustaining high transpiration through the dry season, drawing on water held at depth even when rainfall is low (Nepstad et al., 1994). This sustained dry-season demand is what can establish the upward water deficit in the root zone under which capillary rise arises.

Infiltration partitions the net rainfall between water that enters the profile and water shed as surface runoff, and so sets how much reaches the rest of the balance. Its controlling property, infiltration capacity, is fixed chiefly by the near-surface saturated conductivity and by the macroporosity that roots, fauna and aggregation create, so it is a structural rather than a purely textural property (Bruijnzeel, 2004). In well-structured tropical forest soils this capacity usually exceeds rainfall intensity, so infiltration-excess surface runoff is generally a minor term and most water leaves through the subsurface. In undisturbed Amazonian forest, overland flow amounted to about 1 % of incident rainfall, the remainder infiltrating (Germer et al., 2009). Moreover, the flow that does form tends to re-infiltrate a short distance down slope wherever the infiltration capacity still exceeds the local supply, so that for a soil column on a uniform slope the lateral surface terms can cancel, run-in from up slope offsetting run-off down slope, and the net surface exchange is small (Bonell, 1998). Water that infiltrates and is not taken up by roots percolates downward and can leave the profile as drainage, the dominant non-evaporative output, which recharges groundwater.

Capillary rise happens when the root zone is drier than the layers beneath it, typically above a water table or a persistently wetter horizon. Its rate is set by the unsaturated hydraulic conductivity together with the depth to the water table and the dryness of the root zone. The maximum sustainable upward flux falls sharply as the water table deepens, the groundwater contribution vanishing below a soil-dependent depth (Soylu et al., 2011). The contribution of a shallow water table depends on the soil's pore-size distribution, which is itself closely related to texture. For coarse-textured and fine-textured soils, capillary rise is generally negligible once the groundwater is more than one meter below the surface, whereas for medium-textured soils it remains a significant supply at much greater depths, often reaching 1.5 to 2.5 meters below the surface (Raes & Deproost, 2003). Such conditions are common in the humid tropics. Over the Amazon the simulated water table is shallower than two meters across 20 to 40 % of the basin, and where it is shallow the dry-season upward flux from groundwater sustains part of the evapotranspiration (Miguez-Macho & Fan, 2012). This matrix flux is distinct from direct uptake of deep soil water by roots (Nepstad et al., 1994) and from hydraulic redistribution within the root system (Oliveira et al., 2005).

The preceding fluxes can be assembled into a mass balance for the soil column over a chosen period. Writing S for the water stored in the profile, the change in storage ΔS is:

$$\Delta S = \underbrace{(P - I_c)}_{\text{net rainfall}} - ET_a - D - R_{\text{off}} + R_{\text{in}} + C_r \quad (1)$$

where P is gross precipitation, I_c canopy interception loss, ET_a actual evapotranspiration, D drainage, R_{off} and R_{in} surface run-off and run-in, and C_r capillary rise. All terms are expressed as an equivalent depth of water, the present work uses cm. Equation (1) is the relation that this thesis evaluates with HYDRUS-1D, the sign and magnitude of ΔS over the hydrological year determine whether the reforested soil is gaining or losing water stock and how this differs from the pasture system.

2 Soil hydraulic properties

The fluxes of the water balance are characterized at every depth by two constitutive functions of the soil: the water retention curve, which relates water content to matric potential (MP), and the unsaturated hydraulic conductivity function, which relates the ability to transmit water to the same potential. Together they close the Richards equation, the physical core of HYDRUS-1D.

In the one-dimensional configuration adopted here, lateral exchange cannot be resolved: run-in R_{in} is set to zero, and surface runoff R_{off} arises only when the ponding head exceeds a certain value. On a sloped plot this neglect of lateral surface and subsurface flow is a simplifying assumption, justified for a uniform slope by the tendency of overland flow to re-infiltrate downslope (Bonell, 1998), but it remains a limitation of the column approach.

2.1 The water retention curve

The water retention curve $\theta(h)$ expresses the volumetric water content θ retained by a soil at a given MP h (equivalently the suction or pressure head). It represents the capacity of the soil to retain water at a given suction. Physically, it is a cumulative expression of the pore-size distribution. At small suction, only the largest pores drain, whereas progressively higher suction is required to empty progressively smaller pores, following the capillary law that relates the radius of a water-filled pore to the MP (Brooks & Corey, 1964). The shape of the curve is therefore controlled by texture and structure, that is by the size, geometry and connectivity of the pore space.

Several closed-form models describe $\theta(h)$ [$\text{cm}^3 \text{cm}^{-3}$]. This work adopts the model of Van Genuchten (1980), which is the most widely used and best documented parametrization and is applicable to the great majority of soils,

$$\theta(h) = \theta_r + \frac{\theta_s - \theta_r}{\left[1 + |\alpha h|^n\right]^m}, \quad m = 1 - \frac{1}{n} \quad (2)$$

where θ_r and θ_s are the residual and saturated water contents [$\text{cm}^3 \text{cm}^{-3}$], α is related to the inverse of the air-entry pressure [cm^{-1}], and n is a pore-size distribution index [-]. The single modal van Genuchten form is retained here because no constraint specific to the site favored a more complex model.

2.2 Unsaturated hydraulic conductivity

As a soil dries, the larger and better-connected pores empty first, the remaining flow paths are narrower, more tortuous and partly disconnected, so the hydraulic conductivity K can fall by several orders of magnitude as the MP becomes more negative. Predicting $K(h)$ or $K(\theta)$ from the more easily measured retention curve is therefore useful to vadose zone modeling. The statistical pore-bundle model of Mualem (1976) provides this link, and its combination with Equation (2) yields the van Genuchten–Mualem (VGM) expression used in this work,

$$K(S_e) = K_s S_e^l \left[1 - (1 - S_e^{1/m})^m \right]^2, \quad S_e = \frac{\theta - \theta_r}{\theta_s - \theta_r} \quad (3)$$

where $K(S_e)$ is the conductivity relative to the effective saturation [cm d^{-1}], S_e is the effective saturation [-], K_s the saturated hydraulic conductivity [cm d^{-1}], and l a pore-connectivity and tortuosity parameter [-].

2.3 The Richards equation

The vertical movement of water in the unsaturated zone follows the Buckingham-Darcy law, which writes the volumetric water flux q (positive upward) as the product of the unsaturated hydraulic conductivity and the gradient of the total hydraulic head $H = h + z$, the sum of the matric head h and the gravitational head z (with z positive upward):

$$q = -K(h) \frac{\partial H}{\partial z} = -K(h) \left(\frac{\partial h}{\partial z} + 1 \right). \quad (4)$$

In the absence of sources or sinks, conservation of mass for an incompressible fluid in a non-deforming soil relates the change in water content to the divergence of this flux:

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z}. \quad (5)$$

Substituting Equation (4) into Equation (5) yields the one-dimensional Richards equation (Richards, 1931),

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[K(h) \left(\frac{\partial h}{\partial z} + 1 \right) \right] \quad (6)$$

Equation (6) shows explicitly that the retention function (through $\partial \theta / \partial h$) and the conductivity function $K(h)$ jointly determine water movement, including the sign of the vertical flux: downward percolation and drainage when the gradient term is negative, upward capillary rise when it is positive. This equation, solved numerically with prescribed boundary conditions, is the physical engine of HYDRUS-1D (Simunek et al., 1999). The retention and conductivity functions must therefore be characterized before any flux can be simulated.

2.4 Hysteresis

The relationship between water content and MP is not unique but depends on the wetting and drying history of the soil. At a given MP, a soil that is drying retains more water than the same soil that is wetting, so that distinct retention curves are followed in each direction of flow (Topp, 1971). This path dependence is termed hysteresis and was first systematically described for soil water retention by Haines (1930).

Three mechanisms are classically invoked. The first is the ink-bottle effect: pores consist of wide bodies connected by narrow throats, so emptying is governed by the small throat radius (high suction required) whereas filling is governed by the larger body radius (lower suction), producing different water contents at the same potential during drying and wetting. The second is contact-angle hysteresis: the advancing contact angle of a wetting front differs from the receding angle of a drying front, which shifts the capillary pressure for a meniscus of given curvature. The third is air entrapment: on wetting, air is trapped behind advancing menisci and isolated as occluded bubbles, so the water content reached on wetting falls short of the value attained from full saturation on the preceding drying path (Hosseini et al., 2024). Together these mechanisms make the drying branch lie above the wetting branch over the whole curve.

Two boundary curves bound the hysteretic domain. The main drying curve, followed from full saturation, and the main wetting curve, followed from the dry end. Any reversal of the flow direction before a boundary curve is reached produces an intermediate scanning curve that runs inside the loop and connects the two boundaries (Figure 5). Because natural soils under field conditions rarely follow a boundary curve to its end, their state typically moves along a sequence of scanning curves, which is the situation of interest under repeated rainfall and evapotranspiration cycles.

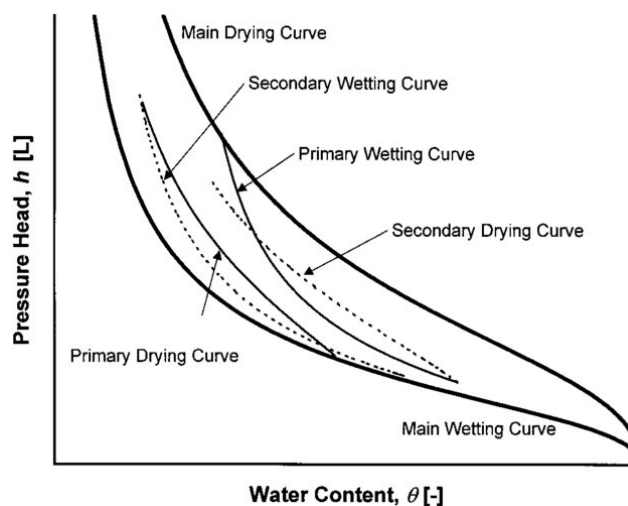


Figure 5: Graph depicting the hysteresis effect on the retention curve. The scanning curves are bounded between the main drying curve and the main wetting curve. Figure originally displayed in Šimůnek et al. (1999).

Most predictive models derive the scanning curves from the measured boundary curves. Do-

main models represent the soil as a population of pore domains, each switching between full and empty states, in the independent-domain formulation the domains do not interact (Philip, 1964). Mualem (1974) and Topp (1971) extended the framework to account for domain dependence and pore interaction. For practical numerical modeling, simpler empirical scaling approaches exist. Kool and Parker (1987) scale drying scanning curves from the main drying curve and wetting scanning curves from the main wetting curve. A known artifact of these scaling models is the so-called pumping effect, by which repeated wetting and drying reversals can drive the simulated state into physically unrealistic, non-closing loops. The model of Lenhard et al. (1991), as implemented in HYDRUS-1D, eliminates this artifact by explicitly tracking the historical sequence of reversal points, which constrains each new scanning curve to remain within the previously visited portion of the retention function, and it represents air entrapment by reducing the effective saturated water content on the wetting branch. This is the formulation adopted in the present work.

Neglecting hysteresis can introduce systematic errors in simulated soil moisture and fluxes, and these errors are largest precisely under the alternating wetting and drying conditions imposed by intermittent rainfall against a continuous evaporative demand. Under such forcing the surface and root-zone layers oscillate between scanning curves rather than tracking a single retention relation, so a non-hysteretic model misrepresents both the stored water at a given potential and the gradients that drive vertical flux. For the present site this is directly determinant. The seasonal alternation of a wet and a dry season subjects the profile to repeated reversals. Accounting for hysteresis is therefore expected to influence the dynamics inside the soil profile.

3 The influence of forest vegetation on soil hydraulic properties

The hydraulic properties are not fixed attributes of the parent material but are continuously modified by the vegetation growing on the soil. A standing forest acts on its own substrate through several coupled biophysical pathways, building and maintaining a topsoil that is generally more porous, better structured and richer in organic matter than that of an unforested site. This section reviews those mechanisms for an established forest.

3.1 Macroporosity

Living and decaying roots, together with the soil fauna they sustain such as earthworms, termites and ants, perforate the soil with biopores and root channels. These biopores are a dominant component of soil macroporosity, the class of large, continuous and interconnected voids through which water moves by preferential flow rather than through the matrix (Beven & Germann, 1982; Jarvis, 2007). Such voids are typically those with an equivalent cylindrical diameter greater than about 0.3 mm (Jarvis, 2007). Their hydrological effect is concentrated near saturation because open and connected macropores sharply raise the saturated and near saturated hydraulic conductivity, and hence the infiltration capacity. Root architecture is itself a control, so that species with contrasting rooting habits generate different macropore

networks and therefore different conductivities at a given texture. The consequence is that the elevated saturated conductivity of a forest topsoil is largely a structural, biologically maintained property rather than a textural one (Jarvis, 2007).

3.2 Organic matter, aggregation and soil structure

Litterfall and root turnover supply organic matter to the soil. The associated microbial and faunal activity decomposes and transforms it, redistributing and stabilizing part of it, with microbial necromass itself becoming a major component of persistent soil organic matter. Accumulation is largest in the surface horizon, where inputs are concentrated. Organic matter is a primary binding agent of soil structure. Tisdall and Oades (1982) showed that roots and fungal hyphae enmesh and stabilize macro-aggregates, while persistent organic compounds stabilize micro-aggregates, so that a well-aggregated, structurally stable pore system develops under sustained organic input. This structure increases total porosity and the proportion of structural inter-aggregate pores, which tends to produce the multimodal or structured pore-size distribution for which the single-modal retention model is only an approximation (Durner, 1994). Organic matter also modifies retention directly. Rawls et al. (2003) found that increasing soil organic carbon generally increases water retention, with the response depending on texture and being strongest in coarse-textured soils. The effect is not uniform across the retention curve as organic matter tends to raise water content more at the wet end, near field capacity, than at the dry end, near the permanent wilting point (Hudson, 1994), so that the saturated water content and, commonly, the plant-available water held by the profile both increase.

3.3 Water repellency

The same organic compounds that build structure can, on drying, coat mineral and aggregate surfaces with hydrophobic films and render the soil water-repellent. In a wide-ranging review, Doerr et al. (2000) established that water repellency is common in forested and other organic-rich soils and is not restricted to coarse textures or to fire-affected sites. Repellency lowers the wettability and the effective infiltration capacity when the soil is dry, promotes uneven wetting and finger or preferential flow, and can generate localized overland flow at the onset of rain. In Brazilian Atlantic Forest soils specifically, Lozano-Baez et al. (2020) identified soil water repellency as a measurable control on infiltration. Repellency is therefore a counteracting process as a forest topsoil can be simultaneously highly conductive when wet, through its macropores, and locally resistant to wetting when dry.

3.4 Implications for this study

Taken together, these mechanisms make the soil under an established forest a biologically engineered medium. More macroporous and structurally stable, richer in organic matter, and higher in near-surface conductivity and retention than a comparable unforested soil, but also

prone to preferential flow and, when dry, to water repellency. Because the root density and organic matter inputs that build this structure are themselves concentrated in the topsoil and decline sharply with depth, with roughly half of the soil organic carbon in the first meter held in the upper 20 cm under forest (Jobbágy & Jackson, 2000), these enhancements are strongest at the surface and attenuate downward. Two points follow for the present work. First, the properties estimated for the reforested site are expected to reflect this structural enhancement, in particular a higher near-surface saturated conductivity than the pasture. Second, because the model represents the macropore and repellency effects implicitly, through the van Genuchten–Mualem parameters rather than as separate processes, their individual contributions cannot be disentangled.

4 Land use change and its hydrological effects

The two plots compared in this work, a 40-year regenerated secondary forest and a degraded pasture, represent two states along the trajectory of land-use change that has reshaped the Atlantic Forest. This section reviews the knowledge on how deforestation alters soil hydraulic properties, how and over what timescale reforestation restores them, and how these changes propagate to the water balance.

4.1 Deforestation

Conversion of forest to pasture degrades the near-surface soil structure. The mechanized clearing compacts and partly strips the topsoil, after which cattle trampling sustains the compaction while the biological agents that build and maintain structure are lost (Bruijnzeel, 2004; Zimmermann et al., 2006). Conversion to cropland shares this initial clearing but differs thereafter, the recurrent plowing and seedbed preparation imposing a mechanical disruption of the aggregate structure, and an acceleration of organic-matter loss, that grazed pasture does not experience (Don et al., 2011). The result is a decline in organic matter together with a higher bulk density and a marked reduction of macroporosity. The organic-matter loss is not merely a parallel symptom but itself a primary driver of the densification, since reduced aggregation increases the soil’s susceptibility to compaction (Tolimir et al., 2020). These changes together lower the saturated hydraulic conductivity and the infiltration capacity (Bruijnzeel, 2004; Zimmermann et al., 2006). Zimmermann et al. (2006) showed for an Amazonian site that the change in soil hydraulic properties accompanying forest-to-pasture conversion is sufficient to alter runoff generation. Germer et al. (2010) documented the associated shift in near-surface hydrological pathways from vertical infiltration toward lateral and overland flow. The reduction in infiltration capacity can be large in heavily used pastures relative to the original forest (Bruijnzeel, 2004). The hydrological signature of deforested pasture is thus an increase in surface runoff and a reduction in the water that infiltrates and percolates.

4.2 Reforestation

Hassler et al. (2011) found that as trees re-establish on former pasture, saturated hydraulic conductivity recovers toward pre-disturbance levels in the uppermost soil layer within roughly a decade. This recovery is driven by root formation and soil macro-fauna activity which increase macroporosity, a process that enhances vertical infiltration and reduces the potential for overland flow even while soil organic carbon levels remain at degraded pasture levels values (Neumann-Cosel et al., 2011). A systematic review and meta-analysis by Ilstedt et al. (2007) found that infiltration capacity increased on average roughly threefold following reforestation or the introduction of trees on degraded tropical land. Hassler et al. (2011) observed recovery of saturated hydraulic conductivity under secondary succession on former pasture in the humid tropics, and Bruijnzeel (2004) reported that hydrological functioning can be partially restored within a few years of regrowth, although full recovery of deep-soil properties is slower.

For the Atlantic Forest specifically, a series of studies on restored and remnant plots provides directly comparable evidence. Lozano-Baez et al. (2018) showed that the recovery of soil hydraulic properties after forest restoration depends strongly on the previous land use. Lozano-Baez et al. (2019) compared assisted, passive and active restoration in relation to historical land use and forest structure, and Lozano-Baez et al. (2020) found that water infiltration was drastically reduced in pasture relative to remnant and restored forest, and identified soil water repellency as an additional control. These results indicate that a 40 year regenerated plot, such as the one used for this study, can be expected to have recovered a substantial part of its infiltration capacity relative to adjacent pasture, while the degree of recovery, and any residual gap with mature forest, remains site dependent. This last point matters here because the cited plots lie in a warmer, lower and flatter part of São Paulo state than the present highland site, and because the Atlantic Forest spans a wide range of climates and terrain; recovery is therefore governed by prior land use, topography and local soil and climate rather than by biome membership alone (Alvares et al., 2013; Lozano-Baez et al., 2018; Lozano-Baez et al., 2019).

4.3 Consequences for the water balance

At the level of the water balance, the contrasting effects of forest and pasture do not act in a single direction, which is the essence of the infiltration-evapotranspiration trade-off hypothesis of Bruijnzeel (2004). Forest cover raises evapotranspiration but also restores infiltration, so the net effect on drainage and groundwater recharge depends on which influence dominates under the local soil and climate. Krishnaswamy et al. (2013) found empirical support for this trade-off in the Western Ghats of India, showing that improved infiltration under reforestation can offset or even reverse the recharge penalty of higher evapotranspiration. Relative to pasture, a forest is expected to transpire more, and to do so even in the dry season because of deep root systems (Nepstad et al., 1994), while admitting more water by infiltration through its recovered structure (Hassler et al., 2011; Ilstedt et al., 2007). Advanced here as a working

hypothesis of this thesis rather than as an established result: a forest that draws the root zone drier in the dry season may establish a stronger upward gradient and hence greater capillary rise where a shallow water table is present, a link the literature on shallow groundwater and forest transpiration makes plausible (Miguez-Macho & Fan, 2012; Soylu et al., 2011) but has not quantified for secondary tropical forest.

These competing shifts frame the comparison at the center of this thesis. Whether the reforested plot gains or loses water stock on an annual basis relative to the degraded pasture depends on the net outcome of higher evapotranspiration, higher infiltration, and a possibly larger capillary rise, that is, on the sign of ΔS in Equation (1). Quantifying this outcome, which the available literature frames but does not resolve for this specific setting, is the contribution that the present modeling study aims to make.

IV. Objectives

The literature reviewed frames a clear trade-off but does not resolve it for secondary tropical forests. Reforestation simultaneously raises evapotranspiration and restores infiltration, so whether a regenerated plot gains or loses soil water relative to a degraded pasture cannot be assumed and must be quantified for the specific soil and climate. Against this background, the general objective of this thesis is to answer the following question:

What is the impact of reforestation on the soil water balance in the Atlantic Forest?

More precisely, over a complete hydrological year, the aim is to determine the sign and magnitude of the change in soil water storage (ΔS in Equation 1) of a specific 40-year-old regenerated plot and how it differs from that of an adjacent pasture. It also aims to identify the fluxes through which any difference between the two land-use states arises.

Reaching this general objective requires three sub-objectives:

- Characterize the key ecohydrological parameters of the two plots, integrating the soil hydraulic properties, the hydro-climatic drivers, and the forest biometrics.
- Model the one dimensional soil water dynamics of each plot by solving the Richards equation under the van Genuchten–Mualem description in HYDRUS-1D, constraining the hydraulic parameters through inverse estimation against the in-situ water-content and matric-potential records, and validating the simulated states against those observations.
- Assess the direction and magnitude of the annual soil water stock variation in the reforested plot, compare it with that of the pasture, attribute the difference to the individual fluxes of the water balance, and identify, through scenario and global sensitivity analyses, the parameters that control the outcome.

V. Materials and methods

1 Study plots

The study area is located in the municipality of São Francisco Xavier (SFX), in the state of São Paulo, Brazil, a region historically encompassed within the Atlantic Forest biome. Following the extensive deforestation that characterized this region, land use shifted predominantly toward agriculture and timber extraction.

Two plots are investigated in this study (Figure 6). The first, designated R40, is the focal plot of this work. It is currently under secondary forest cover, having been reforested approximately 40 years ago after an undetermined period of deforestation. The second plot is situated in close proximity to R40 and currently consists of grassland, it similarly underwent deforestation and is included as a comparative reference representing that land-use state. This plot is referred to as Pasture.

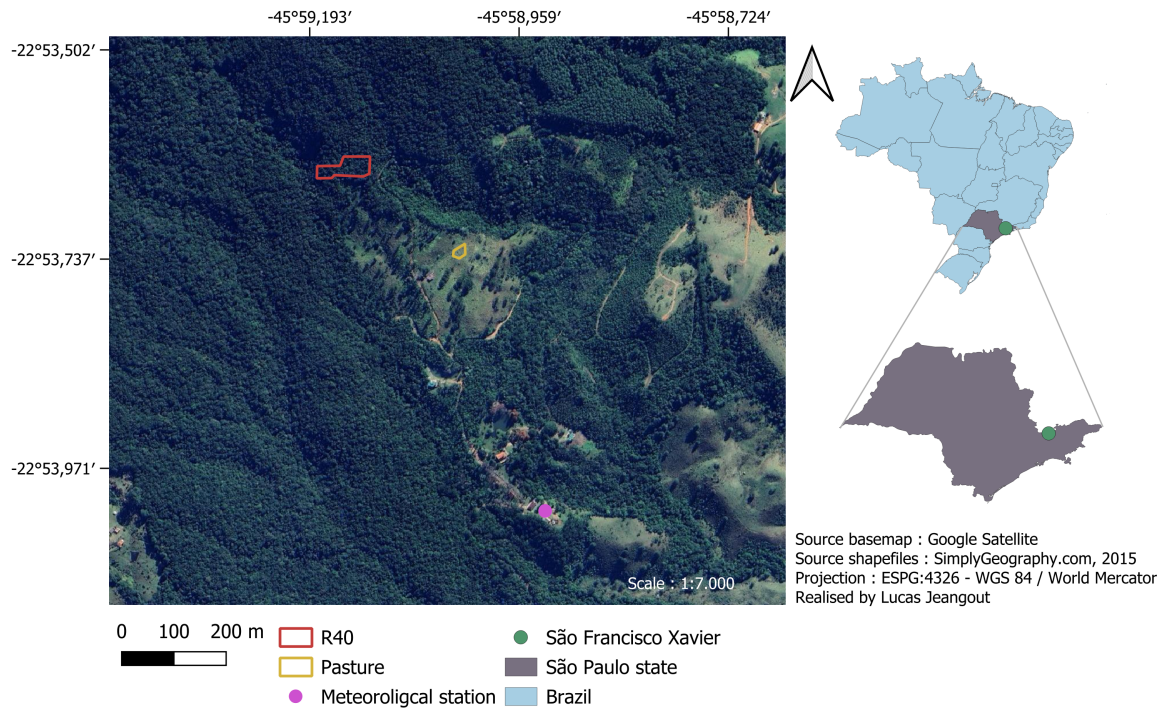


Figure 6: Map of the two study areas, R40 and Pasture, located in the São Francisco Xavier region, with the meteorological station indicated.

R40 has an area of approximately 370 m². It hosts several types of vegetation, covering the herbaceous, shrub, and tree layers. Its strongly established vegetation produces a dense canopy cover (Figure 7). The slope of the plot can reach up to 10 %. Pasture, not far away from R40, covers an area of around 80 m². It has entirely herbaceous vegetation and close to no slope.



Figure 7: Picture of the R40 plot, focused on the canopy and its high density. Taken on 05/02/2026.

2 Meteorological data

2.1 Meteorological Station

The meteorological station used in this study is shown in Figure 8. It belongs to the INPE. This station is situated in an open grass field with trees on the outskirts of the field, which could have an effect on the measured variables via wind attenuation or occasional shading for example. This station has been active since 10 January 2023. Another station was present from the 23 January 2019 until September 15th 2021 in the same place. These stations measure the same metrics, the only difference is the time step, hourly for the previous and 15 minutes for the present.



Figure 8: Picture of the São Francisco Xavier active meteorological station, taken on the 05/02/2026.

The six different metrics these stations provide are: wind speed [m/s], relative humidity [-], temperature [°C], solar radiation [W/m²], rain [mm] and pressure [kPa]. Those data are available from the 23rd January 2019 until February 2026, as of the time of writing.

During the evaluated period, several gaps of data are present. In total, 20.4 % of the data is missing. The repartition of those gaps is asymmetrical. Two periods mainly encompass those gaps. The first one, the period between the two stations, is missing 481 straight days and the second 39 days, representing 99.6 % of the missing data. Two further gaps last several hours, and the remaining eleven each miss a single time step. Due to the scale of the biggest ones, simple interpolations would not work, hence they would not take into account variations during these periods.

To fill those gaps and obtain a full time series, another meteorological station was used. The closest one is situated in Monte Verde, which is a neighboring city from SFX. The stations themselves are 6.5 km away from each other. This station belongs to the Brazilian National Institute of Meteorology, the INMET. It possesses a one hourly time step. Its data is available online on the official website of the INMET. This station overlaps the gaps of the INPE stations for 96.3 % in average across variables, except for solar radiation, which overlaps those periods for only 59.1 %.

The time steps did not match. To have homogeneous characteristics, the data from the current INPE station was aggregated from the 15 minutes time step. The following variables were aggregated by arithmetic mean: solar radiation, temperature, RH, wind speed, pressure. While the precipitations were summed.

A direct injection of the Monte Verde data would be possible, but would introduce a systematic bias arising from the station's site characteristics and measurement instruments. In order to avoid this bias, quantile mapping (QM) was used on the four biggest gaps. For the other eleven, linear interpolation was used.

2.1.1. Quantile mapping

QM is a non-parametric statistical bias correction method that transfers the full empirical distribution of a source dataset onto that of a target dataset (Gudmundsson et al., 2012). Although originally developed to bias-correct climate-model output, it has been successfully applied to fill gaps in station time series using neighboring stations (Grillakis et al., 2020), which is the use adopted here. Its mathematical formulation is: for any observation x drawn from the source distribution, the bias-corrected value x' is obtained as:

$$x' = F_t^{-1}(F_s(x)) \quad (7)$$

where F_s is the empirical cumulative distribution function (CDF) of the source and F_t^{-1} is the quantile function (inverse CDF) of the target, both estimated over a shared calibration period during which both stations recorded simultaneously complete observations.

This method was applied to five of the six variables, precipitation requiring a different approach due to the discontinuous nature of its distribution. For the QM of solar radiation, only daytime was used to prevent the nocturnal values changing the distribution.

Unlike continuous meteorological variables, precipitation data consists predominantly of zero values with sparse positive values. This structure makes empirical quantile estimation unviable, particularly where the proportion of zeros differs between stations and where intense convective events are too rare to be reliably sampled.

A multiplicative ratio correction is applied instead. Let Σ_t and Σ_s be the cumulative precipitation totals recorded at the target station and the source station respectively over the shared calibration period. The correction factor α is defined as:

$$\alpha = \frac{\Sigma_t}{\Sigma_s} \quad (8)$$

and the bias-corrected precipitation at any time step i within the gap is:

$$P'_i = \alpha \cdot P_{s,i} \quad (9)$$

where $P_{s,i}$ is the raw INMET precipitation value. This approach corrects the mean volumetric bias between the two stations while preserving the temporal structure of precipitation events, their timing, duration, and relative intensity, as recorded by the station. It does not make assumption about the shape of either distribution and requires only the estimation of a single scalar parameter, which is estimated over the calibration periods. The implicit assumption is that the ratio of mean precipitation between the two stations is stationary over time, which is physically reasonable given the proximity of the sites.

2.1.2. Residual gaps

Following the QM step, residual missing values arise from periods within the four large gaps during which the INMET reference station itself lacked observations. For temperature, wind speed, and atmospheric pressure, the residual gaps are short, two hours maximum, so linear interpolation was used. For precipitation, some gaps were longer, up to 10 hours, and linear interpolation was inappropriate because rainfall is an intermittent process. Residual missing values were instead set to zero.

Relative humidity required a distinct treatment. The humidity sensor was dysfunctional for extended periods, totaling 1200 missing hours. A linear regression model was fitted to the observed main-station data to predict relative humidity from temperature and the calendar month encoded as a cyclic pair. The model reads

$$\widehat{\text{RH}} = \beta_0 + \beta_1 T + \beta_2 \sin\left(\frac{2\pi m}{12}\right) + \beta_3 \cos\left(\frac{2\pi m}{12}\right) \quad (10)$$

where m is the month of the year. The cyclic month term captures the pronounced contrast between the wet and dry seasons of the Atlantic Forest region, which accounts for the dominant fraction of the explained variance (R^2 rising from 0.33 with temperature alone to 0.62 with the seasonal term added). The model was calibrated on all data points where both variables were simultaneously observed, yielding $R^2 = 0.62$ and an $RMSE$ of 0.096.

Solar radiation also required a dedicated treatment. Before quantile mapping, the INMET solar series was quality-controlled. Daytime values below 5 % of the mean for the same hour and date of the evaluated period were flagged as a sensor malfunction and excluded. This targeted a sensor failure from June to July 2022. Excluded values were replaced by the mean for the corresponding week and hour, in order to avoid day to day exceptional values. Residual gaps outside this period were also filled with their corresponding mean. Because only the magnitude is rescaled, the temporal structure of the filled segments is inherited from Monte Verde rather than reconstructed for São Francisco Xavier.

2.2 Reference and potential evapotranspiration

Reference evapotranspiration (ET_0) can be derived from the meteorological variables. This flux corresponds to the sum of evaporation and transpiration of a standard field, with a grass height of 12 cm, a surface resistance of 70 s/m and an albedo of 0.23, without being under water stress.

This flux can be computed using the Penman-Monteith equation, coupled with aerodynamic and surface resistance formulations. The FAO proposed a unified formula integrating these three equations (Allen et al., 1998). The hourly form of this expression is given below:

$$ET_0 = \frac{0.408 \Delta (R_n - G) + \gamma \frac{37}{T + 273} u_2 (e^\circ(T) - e_a)}{\Delta + \gamma(1 + 0.34 u_2)} \quad (11)$$

where ET_0 is the reference evapotranspiration [mm h^{-1}], R_n the net radiation at the grass surface [$\text{MJ m}^{-2} \text{h}^{-1}$], G the soil heat flux density [$\text{MJ m}^{-2} \text{h}^{-1}$], Δ the slope of the saturation vapour pressure curve [$\text{kPa } ^\circ\text{C}^{-1}$], γ the psychrometric constant [$\text{kPa } ^\circ\text{C}^{-1}$], T the mean hourly air temperature [$^\circ\text{C}$], $e^\circ(T)$ the saturation vapour pressure at temperature T [kPa], e_a the actual vapour pressure [kPa], and u_2 the mean hourly wind speed at 2 m height [m s^{-1}]. The details of this equation and the different variables are in Appendix A.

Once the site specific vegetation is taken into account, the ET_0 becomes the potential evapotranspiration (ET_p). It writes:

$$ET_p = K_c \cdot ET_0 \quad (12)$$

Where K_c [-] is the crop coefficient. Its value was derived by Oufini (2025) from the actual evapotranspiration (ET_a) values produced by the Sen-ET method of Parmentier (2024). Oufini (2025) exploited the contrast between the ET_a of R40 and that of a neighboring pasture,

assuming the pasture to be close to the FAO reference surface ($K_c \approx 1$). A linear regression of R40 ET_a against pasture ET_a , forced through the origin, yielded $K_c = 1.16$ ($R^2 = 0.99$). This value was therefore used for R40. For Pasture, under the assumption that its characteristics match to the reference one, its potential evapotranspiration equals the reference ET_0 .

3 Soil properties

3.1 Field monitoring

Volumetric water content (VWC) and MP are continuously measured in the field. The sensors used and their availability are displayed in Table 1. In each soil profile, the two metrics were evaluated at seven different depths: 10, 20, 30, 50, 100, 150 and 200 cm.

Table 1: Sensor deployment and specifications at the R40 and Pasture sites (METER group (n.d.-c), METER group (n.d.-d)).

	VWC		MP	
	R40	Pasture	R40	Pasture
Sensor	GS1	Teros 10	MPS-6	Teros 21
Start date	20/02/2019	16/07/2021	20/02/2019	16/07/2021
End date	26/11/2025	30/10/2024	09/04/2025	06/02/2025
Range	0.00 to 0.64 m ³ /m ³		−8 to −100 000 kPa (1.90 to 6.00 pF)	
Resolution	0.001 m ³ /m ³		0.1 kPa	
Precision	±0.03 m ³ /m ³		±(10% of reading + 2 kPa)	

All of those sensors have been developed by the METER group. The GS1 and MPS-6 got renamed in 2015 to their Teros counterpart. On the 16th February 2023, the R40 MPS-6 sensor for the 200 cm depth got replaced by a Teros 21 due to a malfunction. At a certain time, the logging interval was reduced from 15 min to 1 h. An arithmetic mean was applied to resample all datasets to an hourly time step.

To sort out anomalies, the data was filtered. Parmentier (2024), Dermauw (2024) and Oufini (2025) served as foundations for this process. A few adjustments were made so that anomalies could still be excluded without sacrificing as much valuable information. The following filters were applied:

- Duplicates were discarded to only keep the first recording.
- Persistence filter: if a sensor reports the exact same value for five consecutive time steps, all readings are discarded.
- VWC physical bounds: from 0 m³ m^{−3} to 0.64 m³ m^{−3}, according to the limitations of the sensors.

- MP physical bounds: from 2.1 to 6 pF, according to the limitations of the sensors and slightly raised from 2.0 pF because readings between 2.0 and 2.1 pF clustered unreliably at the sensor's wet limit.
- Hampel identifier: uses a rolling window of ± 24 hours to compute the local median and median absolute deviation (MAD). A point is flagged as an outlier if:

$$|value - median| > k \cdot 1.4826 \cdot MAD \quad (13)$$

where k is a scalar set to five as a threshold. It aims to target sudden spikes in the sensors. The value 1.4826 represents here the standard deviation under a Gaussian distribution.

- Seasonal backstop: computes the median and MAD over the whole hydrological season. A point is flagged as an outlier using Equation 13 with $k=10$. The role of this filter is to catch slow drifts or prolonged sensor failures that the Hampel identifier misses because they look locally smooth.
- Specific corrections: some anomalies identified visually still persisted, so local thresholds were applied to discard them.

3.2 Bulk density

The bulk density was determined by two distinct laboratories, IAC and LabEcoh. Both used the same methodology. The samples were collected using Kopecky rings with a standard volume of 250 mL. They were heated to 105 °C for 24 hours to obtain dry samples. They were then weighed on a scale with a precision of 0.1 g and a resolution of 0.01 g. The samples and results came from the 2024 campaign (Dermauw, 2024).

3.3 Retention curves

The drying retention curve $\theta(h)$ and the justification for choosing the single-modal VGM model are presented in Section III.2.1. The equation parameters (θ_s , θ_r , α , n , l) were determined for every depth from Hyprop (HYdraulic PROPERTIES analyser, METER group) and WP4C potentiometer (METER group) measurements.

The Hyprop device consists of two tensiometers positioned at two different depths within a soil sample placed in a Kopecky ring. The plane located midway between the two tensiometers corresponds to the horizontal plane of symmetry of the soil column. The sample is sealed at the base by contact with its balance. The soil surface is exposed to the ambient atmosphere, allowing soil water to evaporate. The sample's matric pressure value is calculated based on the average of the two measured tensions (Eq.14). The sample mass is weighed continuously during drying. The water content is calculated from the difference between the fresh and dry masses (Eq.15). The combination of the two equations provides the data points for the curve.

$$h = \frac{h_{t1} + h_{t2}}{2} \quad (14)$$

$$\theta(h) = \frac{M_{tf} - M_{ts}}{M_{ts}} \frac{\rho_a}{\rho_w} \quad (15)$$

Where M_{tf} is the total fresh mass of the sample [g], M_{ts} is the total dry mass of the sample after being oven-dried for 48h at 105°C [g], ρ_a is the bulk density [kg/m³] and ρ_w is the volumetric mass density of water [kg/m³]. The Hyprop tensiometers measuring range extends from +0.3 to -100 kPa. They have an accuracy of 0.1 kPa and a resolution of 0.001 kPa (METER group, n.d.-a). By using the cavitation point of the tensiometer and their air-entry value, the device is able to obtain one more data point above pF 3.

As Hyprop is not able to measure data points on higher suctions, WP4C can be used to better determine the dry range of the retention curves. A capsule is filled with a small dry unstructured sample. The capsule is placed inside the WP4C device. The sample and the air inside the sealed chamber reach equilibrium. The relative humidity of the air in the chamber containing the sample is then measured by the device to estimate the potential. This is done by cooling a mirror until condensation is observed on it. The device records the dew point temperature when it detects condensation, which enables the determination of the air's relative humidity and, consequently, the potential. The VWC is calculated using equation 15, by weighing the sample before and after it has been placed in the oven for 48 hours at 105°C. WP4C range extends from 0 to -300 MPa. Its accuracy is ± 0.05 MPa from 0 to -5 MPa and 1 % from -5 to -300 MPa. The resolution is not available (METER group, n.d.-e).

Curves are then fitted to the data points via the Labros software. As several Hyprop and WP4C pairs were available (Table 2), all data points were grouped by depth and then fitted. The number varies due to sample issues, making them unusable. In total, three sets of measurements were conducted per plot.

Table 2: Number of Hyprop and WP4C measurement pairs per depth at the R40 and Pasture sites.

Depth [cm]	Number of Hyprop + WP4C pairs	
	R40	Pasture
10	3	2
20	3	2
30	3	3
50	2	2
100	3	3
150	3	3
200	3	2

Based on the results of the field and laboratories measurements, some depths were grouped

together due to their similarities in hydraulic properties. This allowed to reduce the complexity of the model, particularly for the inverse modeling (Section V.5.4), by reducing the number of parameters. The new groups can be seen in Table 3.

Table 3: Depth groups for the R40 and Pasture sites, associated by similarities in the field and laboratories measurements.

Group order	Depth [cm]	
	R40	Pasture
1	10	10
2	20+30	20+30+50
3	50+100	100+150
4	150+200	200

3.4 Conductivity curves

The unsaturated conductivity function $K(S_e)$ and its VGM formulation are described in Section III.2.2. Conductivity data points were obtained with Hyprop, complemented by the Ksat device (METER group) for the saturated value of the hydraulic conductivity.

Hyprop uses the Buckingham-Darcy equation (Eq.4) to obtain the data points to construct the conductivity curve. It measures the evaporation flux continuously based on measured mass changes. The potential gradient is computed using the two tensiometers. By isolating the hydraulic conductivity from the Buckingham-Darcy equation, the conductivity value $[\text{cm d}^{-1}]$ is calculated for the potential h [cm], corresponding to the mean of the measured potentials as:

$$h = \frac{h_{\text{tensio1}}(t_i) + h_{\text{tensio2}}(t_i) + h_{\text{tensio1}}(t_{i+1}) + h_{\text{tensio2}}(t_{i+1})}{4} \quad (16)$$

and ΔH [cm] the potential gradient calculated as:

$$\Delta H = \Delta h + \Delta z = \frac{\Delta h(t_i) + \Delta h(t_{i+1})}{2} + L \quad (17)$$

where Δz [cm] is the corresponding difference in gravitational head. Since the gravitational head difference equals the vertical distance separating the two tensiometer cups, Δz is equal to L [cm], the fixed vertical spacing between the tensiometers in the soil sample.

The hydraulic conductivity at saturation is obtained apart from this process with the Ksat device. It measures saturated hydraulic conductivity using the constant head method. A soil core is placed in the device, water is supplied at a fixed hydraulic head above the sample, and the instrument automatically records the flow rate through the soil over time. K_s is then derived from Darcy's law, relating flow rate to the hydraulic gradient and sample geometry. The device automates head maintenance and flow measurement, and works on undisturbed

cores to preserve natural soil structure. Three distinct measurements are done by sample. It has a range from 0.01 to 5000 cm d⁻¹ (METER group, n.d.-b). It also uses Kopecky rings. The samples were measured on the Ksat device before going through the HYPROP process.

Conductivity curves are aggregated the same way as retention curves. K_s values are obtained first by geometric means for the same sample and then by geometric means between the samples for each group (Table 3). Only the samples that also yielded successful Hyprop measurements were used.

3.5 Hysteresis

The physical basis of hysteresis, the Lenhard et al. (1991) formulation adopted here and its suppression of the pumping effect are presented in Section III.2.4. In practice, the retention function (Eq. 2) is applied separately to the main drying and main wetting curves, each with its own α (α and α_w), while n , θ_r , l and K_s are shared between branches. The ratio α_w/α was set to two, a common simplification for deriving the wetting branch from the more easily measured drying curve. This value is supported empirically by Likos et al. (2014), who reported a mean α_w/α_D of 2.24 ± 1.25 across a wide range of soil types. Scanning curves are interpolated from the two main branches and constrained to the previously visited reversal points, which is what removes the pumping artifact.

The HYDRUS-1D parametrization exposes two further water contents near saturation: θ_m , a fictitious upper anchor of the retention function ($\theta_m \geq \theta_s$),¹ and θ_{sw} , the saturated content actually reached on the main wetting curve, which falls below the drying-branch value when air is trapped on re-wetting. Both were fixed to the measured θ_s , so that $\theta_m = \theta_{sw} = \theta_s$. This recovers the unmodified van Genuchten retention and conductivity functions, switches off the air-entrapment reduction, and leaves α as the only parameter distinguishing the wetting and drying branches. These values were fixed rather than optimized because no direct measurement constrained them, and holding them at θ_s avoids enlarging the inverse problem (Section V.5.4).

4 Vegetation data

4.1 Root density

The model relies on root density as an input parameter. To derive this quantity, three soil samples for each depth for both sites, collected during the 2023 field campaign (Delvoie, 2023), were oven-dried and subsequently weighed using a scale with a resolution of 0.01 g and an accuracy of 0.1 g. Roots were then mechanically separated from coarse soil particles by passing the samples through two consecutive sieves of decreasing mesh sizes (1 mm and 0.5

¹ θ_m originates in the modified van Genuchten formulation of Vogel and Cislérova (1988), which inserts a small air-entry break into the retention curve and rescales the unsaturated branch to a fictitious water content $\theta_m \geq \theta_s$ to improve the behavior of the conductivity function near saturation. It is a mathematical anchor rather than a physically attained content, setting $\theta_m = \theta_s$ recovers the unmodified van Genuchten model.

mm), enabling the isolation of fine roots (diameter < 2 mm) from the rest. Fine root density [g kg^{-1}] was computed for each sample as the ratio of the dry fine root mass [g] to the dry soil mass [kg]. Finally, the average root density of each triplicate was adjusted according to the bulk density of the soil [g cm^{-3}], in order to obtain a fine root density expressed in terms of soil volume in [g dm^{-3}].

The analysis focused specifically on fine roots, as it is established in the literature that they account for the large majority of water uptake by plants (Vercambre et al., 2002). The depth-wise fine root densities were normalized over the 0 cm to 200 cm profile to obtain the root distribution function $b(z)$ required by HYDRUS-1D (Eq. 20), with values linearly interpolated between sampling depths.

4.2 Leaf area index

The leaf area index (LAI) is defined as the total one-sided green leaf area per unit of ground surface area [$\text{m}^2 \text{ m}^{-2}$]. Its value characterizes the density of the vegetation canopy. It captures the seasonal dynamics of the tree canopy, which undergoes meaningful variation even in evergreen forests due to the alternation of dry and wet seasons. This variable is useful for partitioning evaporation from transpiration when estimating evapotranspiration. The LAI was also used to obtain the canopy interception.

LAI was not measured directly in the field. Instead, it was retrieved through radiative transfer modeling (RTM). A RTM is a physically based mathematical framework that describes the propagation of electromagnetic radiation through a medium by accounting for the processes of absorption, scattering, and emission. In this case, it simulates the spectral reflectance that a satellite sensor would measure given a set of biological characteristics, including LAI, and physical characteristics such as illumination and observation angle. The model can generate large synthetic datasets of reflectance spectra paired with known biophysical parameters, which can then be used to train machine learning models for the inverse problem of retrieving those parameters from actual satellite observations. The LAI was retrieved using the Sentinel-2 Biophysical Processor, available within European Space Agency's (ESA) Sentinel Application Platform (SNAP). The resulting product has a spatial resolution of 20 m, using two of the 10 m resolution bands and six of the 20 m resolution bands.

This approach was applied to retrieve a time series of LAI values over R40 and Pasture across the study period. Sentinel-2 images were filtered to retain only acquisitions with less than 20 % cloud cover. Despite the five-day revisit time of the Sentinel-2 constellation (comprising satellites 2A and 2B), the number of usable observations is limited by the frequently cloudy conditions at São Francisco Xavier, particularly during the wet season. The resulting irregular LAI series was linearly interpolated to a continuous hourly series for model input. This assumes a smooth trajectory between valid retrievals.

Regarding spatial representativeness, the central coordinates of each plot were used to extract LAI from the corresponding 20 m pixel. Given this pixel size, adjacent areas outside the plot boundary should not contribute to the signal. However, if it does, the vegetation in those

immediately surrounding areas of both plots is similar in structure to that within the plots, so the effect on the retrieved LAI is expected to be minor.

4.3 Actual evapotranspiration

In HYDRUS-1D, the conversion of ET_p into actual evapotranspiration ET_a is split in two distinct steps, partitioning and stress-based reduction.

ET_p is first split into potential soil evaporation (E_p) and potential transpiration (T_p) by partitioning the solar radiation component of the energy budget via interception by the canopy using Beer's law (Ritchie, 1972). It describes as:

$$T_p = ET_p (1 - e^{-k \cdot LAI}) = ET_p SFC \quad (18)$$

$$E_p = ET_p e^{-k \cdot LAI} = ET_p (1 - SFC) \quad (19)$$

where SFC is the soil fraction that is covered [-] and k is the radiation extinction coefficient of the canopy [-]. Since no plot-level measurements of photosynthetically active radiation above and below the canopy were conducted, from which the extinction coefficient k is typically derived, a literature-based value was used instead. Following Zhang et al. (2014), the radiation extinction coefficient was set to 0.59 for R40 and to 0.50 for Pasture, corresponding respectively to the mean values reported for broadleaved forests and grasslands.

Actual soil evaporation is computed via the surface boundary condition of the model. When the soil is sufficiently moist, evaporation is atmosphere-controlled and $E_a = E_p$. When the hydraulic conductivity of the soil near the surface is insufficient to sustain E_p , HYDRUS-1D switches to a minimum pressure head h_s at the surface, corresponding to oven-dry conditions. E_a is then determined by the soil's ability to conduct water upward, it becomes soil-controlled.

Actual root water uptake is computed by applying a water stress response function $\alpha(h)$ [-] to the potential uptake, following:

$$S(z, t) = \alpha(h) b(z) T_p \quad (20)$$

where $b(z)$ [L^{-1}] is the normalized root density distribution function and $S(z, t)$ [T^{-1}] is the volumetric sink term. Four pressure heads define the vegetation's response to soil water status (Figure 9), described here from the wettest to the driest condition. Near saturation, for suctions below $|h_0|$, water logging induces anoxic conditions that suppress root water uptake entirely. As the soil drains and suction increases from $|h_0|$ toward $|h_1|$, progressive aeration relieves this oxygen limitation and uptake rises from zero to its maximum. Between $|h_1|$ and $|h_2|$ the plant transpires at its full potential rate ($T_a = T_p$), water being limiting neither by excess nor by deficit. Beyond $|h_2|$, soil water availability can no longer meet atmospheric demand and the plant restricts stomatal aperture, so actual transpiration falls below potential ($T_a < T_p$). The threshold $|h_2|$ is not fixed but depends on evaporative demand. Under high demand (r_{2h})

stomatal regulation begins at a lower suction ($|h_{2H}|$), whereas under low demand (r_{2L}) full transpiration is maintained to a higher suction ($|h_{2L}|$). At $|h_3|$, the permanent wilting point, uptake ceases completely (Homaei et al., 2002).

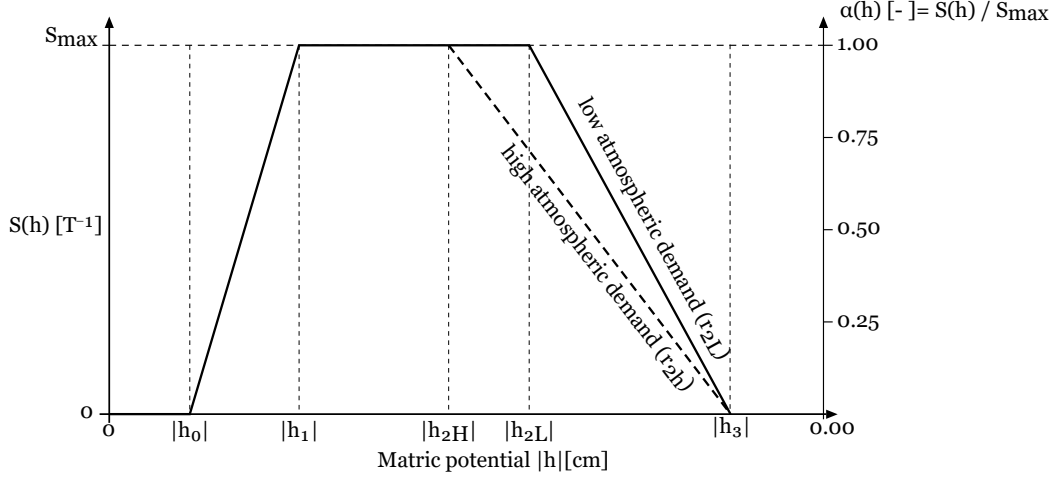


Figure 9: General shape of the water stress response function, as described by Feddes et al. (1978).

Owing to a lack of data on those properties for forested sites, the reference parameters for a standard field were used (Wesseling, 1991), with a change to h_3 to match the wilting point (from pF 3.9 to pF 4). As shown in a sensitivity analysis by Delvoie (2023), the model is relatively insensitive to the h_{2L} and h_3 parameters.

Actual transpiration is then obtained by integrating the sink term (Eq. 20) over the root zone, which corresponds to the full 200 cm profile for both Pasture and R40. Finally, ET_a is computed as the sum of E_a and T_a .

4.4 Canopy interception

Throughfall was measured on the R40 site using 19 manual collectors distributed across the area, three of which were external gauges recording gross rainfall. Gauges were read at irregular intervals of approximately two weeks. The measurement campaign started on the 15th December 2020 and ended the 1st January 2022, yielding 18 usable measurement periods. Each recorded value represents the volume accumulated since the preceding measurement. For each period, gross rainfall P was taken as the mean of the three external gauges, and throughfall TF as the mean of the 16 internal gauges. Observed interception was then computed as $I_{obs} = P - TF$ per period. Stem flow was not explicitly measured and was not added into the model, representing a limitation.

HYDRUS-1D computes the daily interception with the Von Hoyningen-Hüne and Braden formula (Von Hoyningen-Hüne (1983), Braden (1985) and van Dam et al. (1997)), which describe as:

$$I = a \cdot LAI \left(1 - \frac{1}{1 + \frac{b \cdot P}{a \cdot LAI}} \right) \quad (21)$$

with

$$b \approx SFC = 1 - e^{-k \cdot LAI} \quad (22)$$

where P is daily gross rainfall [mm], b is an empirical parameter that approximates to SFC [-], and a [mm d⁻¹] is the canopy storage capacity per unit LAI. The parameter a was calibrated against the field measurements.

A systematic bias was observed at the on site external pluviometer gauges relative to the meteorological station. Per-period ratios ranged from 1.1 to 2.9, with a mean of 1.5. This represents a limitation for those measurements. To avoid this bias, calibration was performed in terms of interception fraction rather than absolute interception quantity. This discrepancy can be caused by various reasons as the external gauges may still be affected by the trees via wind or dripping points for example. For each measurement period, the predicted interception fraction was computed as the ratio of the summed daily model predictions to the summed daily meteorological station rainfall:

$$f_{\text{pred}} = \frac{\sum I_{\text{pred},t}}{\sum P_{\text{meteo},t}} \quad (23)$$

The corresponding observed fraction was computed from the pluviometer readings:

$$f_{\text{obs}} = \frac{I_{\text{obs}}}{P_{\text{pluvio}}} \quad (24)$$

In both cases, the division is applied to period totals, after aggregating all daily values, such that the non-linearity of the Equation 21 is preserved. The parameter a was identified by minimizing the sum of squared errors between predicted and observed interception fractions across all 18 periods:

$$\mathcal{L}(a) = \sum_{\text{periods}} (f_{\text{pred}} - f_{\text{obs}})^2 \quad (25)$$

Each period is weighted equally regardless of its total rainfall, which avoids the disproportionate influence of large wet periods on the calibration.

The canopy-storage parameter a was calibrated against the throughfall measurements at R40 only. For Pasture, no throughfall campaign was conducted. The canopy interception by the herbaceous cover was therefore set to 0.25 mm/d, the standard value recommended for ordinary agricultural crops and grasses in the HYDRUS-1D modeling framework.

5 Water flow modeling

5.1 HYDRUS-1D presentation

Soil water dynamics were simulated with HYDRUS-1D (version 4.17.0140, Šimůnek et al. (2024)), a finite-element model that solves the one-dimensional Richards equation (Eq. 6) to compute the vertical distribution of volumetric water content and matric potential over time. The governing physics, the VGM constitutive relations, and the Lenhard hysteresis formulation adopted here are presented in Section III.2, the root-water-uptake and evapotranspiration schemes are described in Section V.4.3, and the canopy interception model is detailed in Section V.4.4. The present section focuses on the numerical setup, boundary conditions, and initial state.

5.1.1. Spatial and temporal discretization

The soil profile was discretized into 201 nodes spaced 1 cm apart, covering the depth interval 0 to 200 cm below the soil surface. The same uniform mesh was used for both the R40 and Pasture profiles. The 201 nodes were partitioned into four material layers following the depth groupings established from laboratory analysis (Table 3). An adaptive time-stepping scheme was adopted. The minimum time step was set to 10^{-5} days (approximately 0.86 s) for R40 to resolve sharp wetting fronts on extreme rainfall events, and to a lower 10^{-6} days for Pasture, as it was running into convergence issues on those same exact events. The maximum time step to $\frac{1}{48}$ days (30 minutes) to speed up the simulation during calmer periods. HYDRUS-1D is able to modify its time step based on the number of iterations required to reach convergence of the Richards equation at each step. If the iterative solution converges in few iterations, the time step is increased, whereas if convergence is slow or fails, the time step is reduced, thereby balancing numerical stability and computational efficiency.

5.1.2. Boundary conditions

The upper boundary was set to an atmospheric condition, with hourly net precipitation and potential evapotranspiration supplied from the meteorological time series described in Section V.2. Canopy interception was computed internally by HYDRUS-1D using the Von Hoyningen-Hüne Braden model (Section V.4.4), with the calibrated canopy-storage parameter a and extinction coefficient k derived as described in that section. Surface ponding was permitted up to a maximum pressure head of 100 cm. The lower boundary condition was set to the time variable prescribed pressure head at the 200 cm depth, derived from the in situ measurements time series. This condition allows the model to represent the influence of the fluctuating shallow water table on deep soil moisture and the potential for capillary rise toward the root zone during the dry season.

5.1.3. Initial conditions and simulation periods

Both plots are initialized from the soil water content and pressure head profiles at the first observation date. A one-year spin-up period is imposed before the evaluation window to allow the simulated state to depart from the prescribed initial conditions and reach a regime driven by the meteorological forcing rather than by initialization artifacts. For R40, the full simulation covers 24 January 2019 to 31 December 2025 (2534 days), for Pasture it covers 16 July 2021 to 2 November 2024 (1206 days), the cut-off being set by the end of the VWC record.

5.2 Global sensitivity analysis

Prior to inverse modeling, a global sensitivity analysis (GSA) was conducted on each site independently to identify the VGM parameters that most strongly influence the model outputs. This step serves two complementary purposes: it provides an objective, quantitative basis for the selection of parameters to optimize in the inverse problem (Section V.5.4), and it directly quantifies which parameters control the water balance outputs of interest, independently of the fitting procedure. The Morris elementary effects method (Morris, 1991) is chosen as the GSA approach. This screening technique evaluates the effect of individually perturbing each parameter across its feasible range and yields two sensitivity indices: μ^* , the mean absolute elementary effect (overall importance), and σ , the standard deviation of elementary effects (non-linearity and interactions). The analysis is implemented using the SALib Python library (Herman & Usher, 2017). The sensitivity problem encompasses 40 parameters: the six drying-branch VGM parameters (θ_r , θ_s , α , n , K_s , l) together with four wetting-branch parameters expressed as dimensionless ratios to their drying-branch counterparts (Equation 26), for each of the four material layers:

$$r_{\theta_{sW}} = \frac{\theta_{sW}}{\theta_s}, \quad r_{\alpha_W} = \frac{\alpha_W}{\alpha_D}, \quad r_{\theta_m} = \frac{\theta_m}{\theta_s}, \quad r_{K_{sW}} = \frac{K_{sW}}{K_s} \quad (26)$$

Expressing the wetting-branch parameters as ratios to their drying-branch counterparts serves a dual purpose. Physically, it encodes the constraints that must hold regardless of parameter values ($\alpha_W \geq \alpha_D$, $\theta_{sW} \leq \theta_s$, $\theta_m \geq \theta_s$). Even though the hysteresis takes effect only on the retention curve with the selected model, HYDRUS-1D still imposes a K_{sW} parameter, using it in the GSA will inform if it takes any part in the model output. The parameters α and K_s are sampled on a $\log_{10}$ scale to provide equal Morris resolution across each decade of their ranges. The sampling bounds for all ten parameter slots are given in Table 4. Bounds for the drying-branch parameters are centered on the laboratory base values, while bounds for the wetting-branch ratios are fixed and site-independent.

Table 4: Morris sampling bounds for each of the ten parameter slots, applied identically to all four material layers. Drying-branch bounds are expressed relative to laboratory base value, wetting-branch bounds are fixed. Parameters marked *log* are sampled uniformly in $\log_{10}$ space.

Slot	Scale	Lower bound	Upper bound
θ_r	linear	$0.5 \times \text{base}$	$\min(0.40, 0.60 \times \theta_{s,\text{base}})$
θ_s	linear	$0.85 \times \text{base}$	$\min(0.90, 1.10 \times \text{base})$
α	log	$0.5 \times \text{base}$	$\text{base} \times 2$
n	linear	$\max(1.05, 0.9 \times \text{base})$	$1.15 \times \text{base}$
K_s	log	$\frac{1}{3} \times \text{base}$	$\text{base} \times 30$
l	linear	$\text{base} - 1.5$	$\text{base} + 1.5$
$r_{\theta_{sw}}$	linear	0.70	1.00
r_{α_w}	linear	1.00	3.00
r_{θ_m}	linear	1.00	1.30
$r_{K_{sw}}$	linear	0.30	2.00

This Morris design uses 50 trajectories with 4 levels per parameter axis (Saltelli et al., 2008), generating $50 \times (40 + 1) = 2050$ HYDRUS-1D runs per site. Simulations are performed over the calibration period, after a one year spin-up to avoid initial conditions bias. Model performance is assessed against four scalar outputs:

1. RMSE_θ : the root-mean-square of normalised residuals on volumetric water content, pooled across all seven observation depths, where residuals are normalised by the pooled standard deviation of the observed VWC time series
2. RMSE_{pF} : the root-mean-square of normalized residuals on pF, similarly pooled and normalized
3. D_{cum} : cumulative bottom-boundary flux over the evaluation window
4. ET_{cum} : cumulative actual evapotranspiration over the evaluation window.

The first two outputs represent calibration fit and identify parameters that determine how well the model reproduces the observed soil water state. The last two represent water balance quantities, independently of model–data agreement, and ensure that the sensitivity ranking reflects the parameters that govern the fluxes used to answer the study question. Parameters are ranked by their mean μ^* across the four outputs.

5.3 Scenarios analysis

Prior to the full Bayesian calibration, three pre-calibration scenario analyses are conducted to determine the optimal configuration for structural model choices whose effect on simulation

quality is not known a priori. All analyses are performed with soil hydraulic parameters fixed at their laboratory base values, so that the results reflect model structural choices rather than parameter uncertainty. In each case the evaluation criterion is the coefficient of determination R^2 computed separately on the VWC and MP time series, pooled across all seven observation depths and the full evaluation window.

5.3.1. Root water uptake compensation

In HYDRUS-1D, a root compensation mechanism allows roots to increase their uptake in wetter portions of the root zone when uptake is locally inhibited by water stress elsewhere, partially maintaining total transpiration at its potential level. The degree of compensation is governed by the dimensionless critical water-stress index ω . When $\omega = 1$ no compensation is enabled and $\omega = 0$ means full compensation. To assess whether reducing this value improves the simulation of observed soil moisture dynamics, six values in the set $\{1.0, 0.9, 0.8, 0.7, 0.6, 0.5\}$ are tested. Based on the results, a single value ω is adopted for all subsequent simulations of both sites.

5.3.2. Near-saturation retention behavior

The standard van Genuchten model (Equation 2) has a vertical tangent as $h \rightarrow 0^-$, which can cause numerical difficulties when the soil approaches full saturation during intense rainfall events. An alternative formulation introduces a small air-entry pressure head $h_{\text{air}} = -2$ cm below which the van Genuchten curve is followed and above which the soil is held at θ_s , eliminating the infinite slope and regularizing the conductivity function near saturation. Two configurations are compared, the standard model ($h_{\text{air}} = 0$) and the modified formulation with $h_{\text{air}} = -2$ cm. Based on the result, one option is adopted for all subsequent simulations of both sites.

5.3.3. Canopy interception

As seen in Section V.4.4, the parameter a controls the magnitude of the intercepted flux. The value calibrated against the R40 throughfall campaign was comparatively high (Section 1.3.3.), so a sensitivity analysis was conducted to assess how strongly this parameter affects the simulated dynamics. Six values spanning the plausible range, $\{0.25, 0.50, 1, 2, 3, 3.59\} \text{mm d}^{-1}$, were tested against the in-situ observations. As in the previous two scenarios, the evaluation criterion is the R^2 on the VWC and pF time series, and based on the results a single value is adopted for all subsequent simulations.

5.4 Inverse modeling

The VGM parameters are initially derived from HYPROP retention measurements and Ksat device data (Sections III.2.1 and III.2.2). These laboratory estimates characterize small, undis-

turbed cores under controlled evaporation and may not represent the effective hydraulic behavior at the profile scale under field dynamics. Saturated conductivity in particular is known to be scale-dependent owing to preferential flow paths and lateral heterogeneity not captured by core-scale measurements. Inverse estimation is therefore used to refine those parameters. The inverse problem is cast in a Bayesian framework. Rather than seeking a single optimal parameter set, the objective is to infer the full posterior distribution of the hydraulic parameters given the in-situ observations, so that parameter uncertainty is explicitly propagated to the simulated water balance fluxes and storage terms. Following Bayes' theorem:

$$p(\theta | y_{\text{obs}}) = \frac{p(y_{\text{obs}} | \theta) \cdot p(\theta)}{p(y_{\text{obs}})} \quad (27)$$

where θ is the parameter vector, y_{obs} the observation vector, $p(y_{\text{obs}} | \theta)$ the likelihood of the observed data treated as a function of θ for fixed y_{obs} , $p(\theta)$ the prior distribution, and $p(y_{\text{obs}}) = \int p(y_{\text{obs}} | \theta) p(\theta) d\theta$ the marginal likelihood. The marginal likelihood is an intractable high-dimensional integral that serves as a normalizing constant with respect to θ . The Bayesian inverse modeling framework applied here follows the approach of Scharnagl et al. (2011), who demonstrated its use for calibrating VGM parameters against in-situ soil moisture observations using informative laboratory-derived priors and a jointly estimated measurement error structure.

5.4.1. Parameters optimized

The global sensitivity analysis identified n as the parameter with the highest mean absolute elementary effect on the model outputs (Section VI.2). Nevertheless, n is not included among the optimized parameters as the Labros fit procedure yields tight per-material bounds (all parameter bounds can be seen in Appendix B), reflecting the fact that the shape exponent is well constrained by the full drying curve measured in the laboratory. Optimizing a parameter that is already tightly constrained by direct measurement would add complexity to the inverse problem without improving the physical realism of the estimate. By contrast, the saturated hydraulic conductivity K_s and the parameter α are subject to known scale dependence between the core and the field. K_s in particular spans orders of magnitude depending on the spatial extent of macropores and preferential flow paths that may not be captured by 250 mL cores, while α , which controls the air-entry pressure and the shape of the retention curve near saturation, can differ between structured field soils and the same soil after disturbance. Optimizing α carries an additional advantage in the presence of hysteresis, since the wetting-branch air-entry pressure is defined as $\alpha_w = 2\alpha$, any update to α during calibration automatically propagates to α_w , allowing the hysteresis model to adapt to field conditions without introducing an extra free parameter. At the same time, the Labros fit procedure provides an informative prior on α , anchoring the optimization within a physically credible range. These two parameters are therefore selected for optimization across all four material layers, yielding eight physical parameters in total. Both are explored on a $\log_{10}$ scale given their strictly positive nature and their order-of-magnitude variability. All remaining drying-branch parameters

are held fixed at their laboratory-derived base values. For the wetting branch, α_w/α_d is fixed at 2.0 and all other hysteresis parameters are set equal to their drying-branch counterparts ($\theta_{sw} = \theta_s$, $\theta_m = \theta_s$, $K_{sw} = K_s$). The full parameter vector θ therefore has ten dimensions: the eight log-transformed physical parameters plus two measurement error hyperparameters described below.

5.4.2. Likelihood function

The model–data misfit is quantified through a diagonal Gaussian likelihood treating the residuals on volumetric water content and pF independently (Scharnagl et al., 2011):

$$\ln p(\mathbf{y} | \theta) = -N_\theta \ln \sigma_{\text{VWC}} - \frac{1}{2\sigma_{\text{VWC}}^2} \sum r_\theta^2 - N_{\text{pF}} \ln \sigma_{\text{pF}} - \frac{1}{2\sigma_{\text{pF}}^2} \sum r_{\text{pF}}^2 + \text{const.} \quad (28)$$

where r_θ and r_{pF} are the pointwise residuals between simulated and observed volumetric water content [$\text{m}^3 \text{m}^{-3}$] and pF [–] at all seven observation depths², N_θ and N_{pF} are the total numbers of valid VWC and pF observations respectively, and σ_{VWC} and σ_{pF} are measurement error standard deviations treated as unknown hyperparameters jointly estimated with the physical parameters. Treating σ_{VWC} and σ_{pF} as unknowns rather than fixing them avoids the need to specify sensor precision a priori. A misspecified measurement error would inflate or deflate posterior confidence and bias the inferred parameter uncertainties. The measurement error hyperparameters are bounded to $\sigma_{\text{VWC}} \in [10^{-3}, 0.10][\text{m}^3 \text{m}^{-3}]$ and $\sigma_{\text{pF}} \in [10^{-2}, 1.00] \text{ pF}$, with log-uniform priors within these bounds (Scharnagl et al., 2011).

5.4.3. Prior distributions

Informative priors are constructed from the available laboratory measurements, following Scharnagl et al. (2011). All eight physical parameters use lognormal priors in $\log_{10}$ space, but the source of the prior centre differs by parameter type. For α , the centre is the geometric mean of the Hyprop replicates. For K_s , the centre is the geometric mean of the two to six direct Ksat measurements available per material layer (Section III.2.2). These point measurements characterize saturated conductivity at the core scale. The prior standard deviation is derived as $\sigma_{\log} = \text{std}(\log_{10} K_s)$ from the individual measurements per material layer wherever two or more replicates are available, yielding material-specific values that reflect the observed core-scale variability. For α at both sites, per material HYPROP [$p_{2.5}, p_{97.5}$] bounds are available from the fitting procedure. The prior standard deviation is derived as $\sigma_{\log} = (\log_{10} p_{97.5} - \log_{10} p_{2.5})/(2 \times 1.96)$, placing 95 % of the prior mass within the observed laboratory range. The resulting base values, prior standard deviations, and physical bounds for both sites are summarized in Table 5.

²Because the pressure head at 200 cm is prescribed as the lower boundary condition (Section 5.1.2.), the pF residual at that depth is structurally zero and is excluded from the likelihood. Only the six shallower depths contribute to the pF term, while volumetric water content is scored at all seven.

Table 5: Prior specifications for the eight optimized parameters. All priors are lognormal in $\log_{10}$ space. The K_s priors are centered on the geometric mean of the 2–6 direct Ksat measurements per material layer, the α priors are centered on the geometric mean of the Hyprop replicates. The $K_s \sigma_{\log}$ is derived as $\text{std}(\log_{10} K_s)$ from the individual Ksat measurements per material layer. The $\sigma_{\log}$ for α is derived from the Hyprop range as $(\log_{10} p_{97.5} - \log_{10} p_{2.5}) / (2 \times 1.96)$ for both sites. Physical bounds apply to both sites.

Parameter	Material	Base value		$\sigma_{\log}$		Physical bounds
		R40	Pasture	R40	Pasture	
K_s [cm d^{-1}]	1	1962	65	0.59	0.11	[0.01, 10^4]
	2	456	424	0.25	0.43	
	3	336	13	0.27	0.19	
	4	99	25	0.60	0.50	
α [cm^{-1}]	1	0.0721	0.0459	0.060	0.043	[10^{-5} , 0.5]
	2	0.0903	0.0587	0.041	0.048	
	3	0.0869	0.0349	0.058	0.060	
	4	0.0699	0.0417	0.041	0.171	

5.4.4. Posterior sampling and propagation

Because the marginal likelihood $p(y_{\text{obs}})$ is intractable, the posterior in Eq. 27 cannot be evaluated in closed form. Markov Chain Monte Carlo (MCMC) methods circumvent this by constructing a Markov chain whose stationary distribution is the target posterior. Each state depends only on the previous one, and after a sufficient number of steps the visited states constitute samples drawn from $p(\theta | y_{\text{obs}})$. Crucially, the transition rule requires only the unnormalized posterior $p(y_{\text{obs}} | \theta) p(\theta)$, so the unknown normalizing constant cancels and never has to be computed. DREAM(ZS) (Differential Evolution Adaptive Metropolis with Snooker updater) is an MCMC sampler designed to remain efficient in high-dimensional problems with strongly correlated parameters (ter Braak & Vrugt, 2008).

In the present work, the algorithm maintains $N_c = 5$ parallel Markov chains whose proposals are generated by differential evolution jumps formed from pairs of states drawn at random from a shared history archive of accepted samples. This mechanism automatically scales proposals to the covariance structure of the posterior, making the sampler efficient for high dimensional problems with correlated parameters. The Snooker updater complements the standard differential evolution jump by projecting the current state onto the subspace defined by two archive points, which improves mixing in the directions of strongest correlation and reduces the risk of premature convergence to a local mode.

For R40, the calibration is performed on the truncated 1072-day simulation (24 January 2019 to 31 December 2021), with the scoring window restricted to 1 January 2020 to 31 December 2021 (days 343 to 1073). Each proposal requires one complete HYDRUS-1D run on this truncated period. For Pasture, calibration and posterior propagation cover 16 July 2021 to

16 April 2024 (1006 days), shortened from the full record due to convergence issues. For the inter-plot comparison (Sections VI.6 and VI.7), a single run using the posterior-median parameters is extended over the complete record, 16 July 2021 to 2 November 2024 (1206 days).

Chains are initialized by drawing directly from the joint prior. The history archive is pre-seeded with $10 \times \text{parameter dimension} = 100$ draws from the prior, ensuring that the differential-evolution proposals are well-conditioned from the first step. The sampling is run for 1120 steps per chain. The first quarter of all steps is discarded as burn-in. Convergence of the retained sample is assessed using the Gelman–Rubin statistic: a chain is considered converged when $\hat{R} < 1.1$ for all parameters (Gelman & Rubin, 1992). The effective sample size per parameter is also computed using the initial positive sequence estimator to verify that the retained draws are sufficiently independent. The prior-to-posterior contraction ratio ($\sigma_{\text{post}}/\sigma_{\text{prior}}$) is reported per parameter to verify that the field observations informed each parameter beyond the prior. Once convergence is confirmed, 200 parameter draws are sampled at random from the joint posterior and each is propagated over the full simulation period. For each draw, HYDRUS-1D produces the complete simulated VWC and pressure head time series at all seven observation depths, from which the 2.5th, 50th, and 97.5th percentile predictive bands are assembled across the ensemble. Cumulative drainage and evapotranspiration over the scoring window are also extracted per draw, yielding full posterior predictive distributions for these water balance components.

5.5 Results expectation

Based on the processes reviewed in Sections III.3 and III.4, two opposing hydrological effects of reforestation are expected to determine the sign of ΔS for R40 relative to Pasture. On the one hand, the dense forest canopy intercepts a larger fraction of gross rainfall and the deep root system sustains higher transpiration rates throughout the year, including during the dry season. Both processes constitute larger water losses from the system. On the other hand, 40 years of biological activity are expected to have produced higher near surface macroporosity and saturated hydraulic conductivity at R40 relative to Pasture soil, favoring infiltration, so that a larger fraction of gross precipitation enters the profile. In addition, the greater draw down of the root zone during the dry season may intensify the upward hydraulic gradient at the base of the profile, potentially driving a larger upward capillary flux through the lower boundary condition. Whether R40 gains or loses water stock on an annual basis relative to Pasture therefore depends on the net outcome of these competing processes under the specific soil and climate of São Francisco Xavier. The model-based quantification presented in Section VI. is intended to determine which term dominates and to attribute the difference in ΔS between the two land-use states to the individual fluxes of Equation 1.

VI. Results and discussion

1 Collected data

1.1 Meteorological data

The completed meteorological series are shown in Appendix C. Imputation quality is summarized in Figure 10, which compares quantile-mapped predictions against observed INPE values over the overlap period. Relative humidity, pressure and temperature all reach coefficient of determination $R^2 > 0.97$, indicating reliable imputation. Solar radiation and wind speed score lower, consistent with site-specific variability only partially captured by quantile mapping. Precipitation is the least well reproduced variable, as expected given the strong spatial variability of rainfall intensity and duration. However, the comparatively low mean absolute error suggests that imputed rainfall *totals* remain close to the true amounts, which is the quantity most relevant to the water balance.

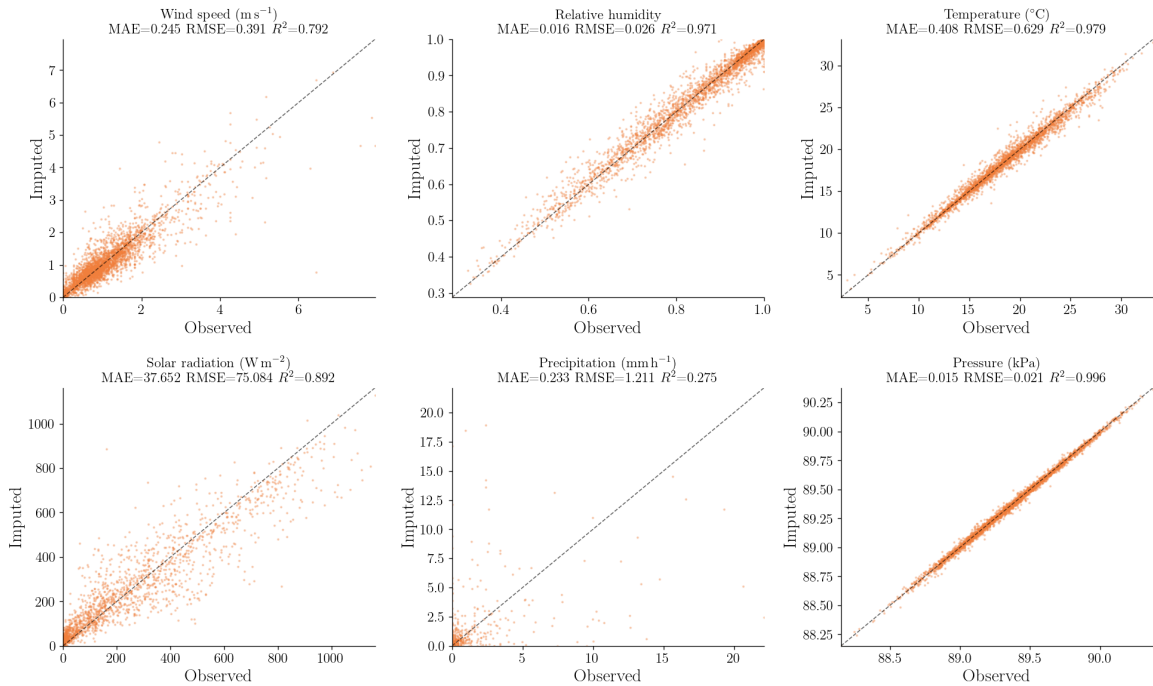


Figure 10: Graph comparing the predicted and observed values for the six meteorological variables. For each variable, the mean absolute error (MAE), the root mean squared error (RMSE) and the determination coefficient (R^2) are evaluated. A 1:1 line is also displayed on each sub-graph.

From those filled datasets, ET_p was derived using the FAO-56 Penman-Monteith method and is displayed in Figure 11. The mean cumulated potential evapotranspiration³ over the entire period is 999.5 mm/year, ranging from 888.0 mm in 2022 to 1084.7 mm in 2020. This figure clearly illustrates the hydrological seasonal pattern of the site: the mean ET_p during the dry season is 72.1 mm/month, compared to 93.2 mm/month during the wet season, reflecting the

³This is the potential evapotranspiration for the R40 site, computed as in Equation 12, this flux for Pasture corresponds to the reference evapotranspiration and is therefore always smaller.

higher atmospheric water demand driven by increasing solar radiation in the austral spring and summer. The minimum monthly ET_p recorded over the full period is 44.5 mm (June 2022), while the maximum reaches 130.7 mm (December 2023). On a climatological basis, September presents the highest average ET_p at 107.5 mm/month. This is a consequence of clear sky conditions at the end of the dry season combined with rapidly increasing spring solar radiation. June on the other hand exhibits the lowest at 52.3 mm/month. Regarding precipitation, the mean annual rainfall is 1983 mm, with a strong seasonal contrast: the wet season averages 250.1 mm/month against only 52.8 mm/month during the dry season. January is the rainiest month on average at 360.7 mm, and July the driest at 35.9 mm. Some interannual variability is visible in both variables, but the general pattern consistently follows the expected wet and dry seasonal cycle characteristic of the humid subtropical highland climate of the Serra da Mantiqueira. The site total (1983 mm yr^{-1}) exceeds the Campos do Jordão normal (1565 mm yr^{-1}), consistent with the latter being a regional reference rather than a local proxy (Section II.2).

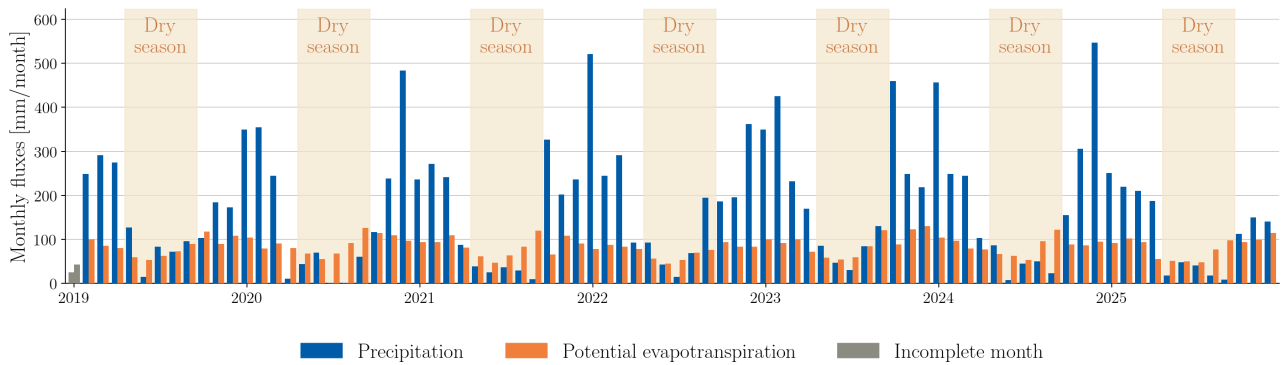


Figure 11: Graph of the precipitation and the potential evapotranspiration fluxes evaluated for the whole study period. The wet and dry theoretical seasons are also displayed on the graph. The dry season ranges from May to September while the wet season occupies the rest of the months. The ET_p is derived from the meteorological variables using the FAO-56 Penman-Monteith method (Eq. 11).

1.2 Soil data

1.2.1. Volumetric water content

The filtered VWC time series for both plots are shown in Figure 12. Both display strong seasonality, with the largest amplitude in the near-surface layers and progressively damped, more stable dynamics at depth. R40 reaches both drier and wetter states than Pasture, a wider overall range, consistent with its higher evaporative demand in the dry season and higher infiltration capacity during rainfall.

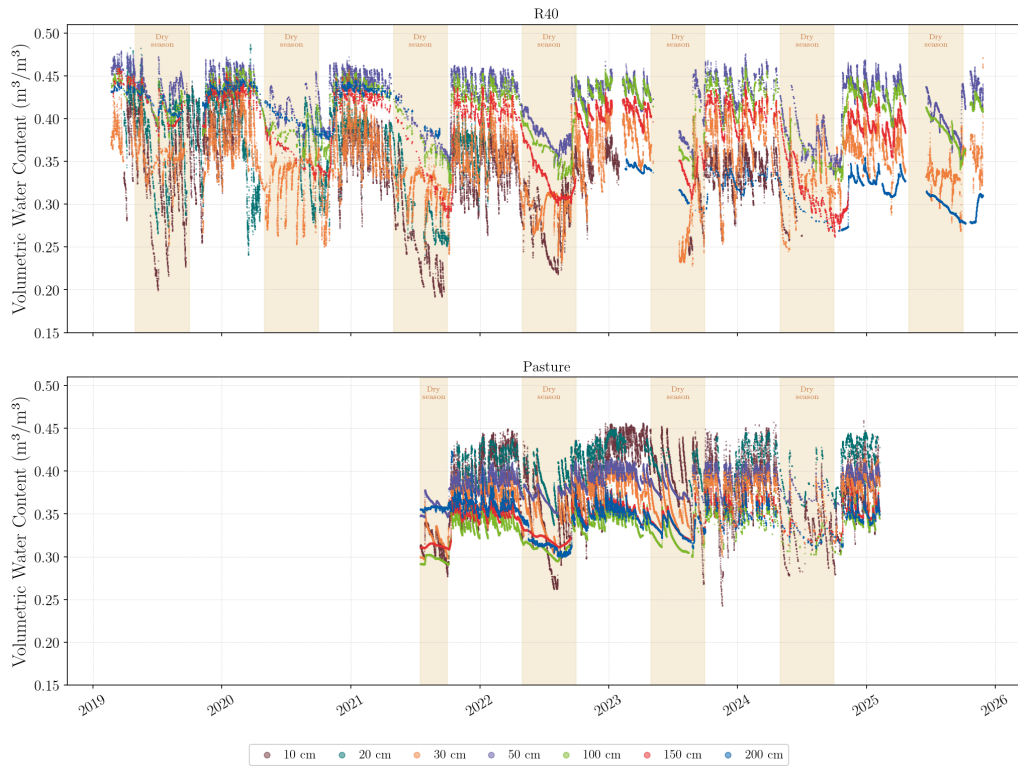


Figure 12: Volumetric water content measured at both sites, R40 (above) and Pasture (below). The different colors represent the different depths. The wet and dry theoretical seasons are also displayed on the graph. The dry season ranges from May to September while the wet season occupies the rest of the months.

1.2.2. Matric Potential

The time series for the MP measured in situ and after being filtered are displayed in Figure 13 for both plots. Similar to VWC, the seasonality is visible but it is going through up to two order of magnitudes for the MP, as pF is a logarithmic scale. The data gaps are also more pronounced. The sensors limit of pF 2 will affect the simulation as the true values likely dropped below pF 2 during the wet season, which the sensors could not capture. The Pasture never reaches states as dry as R40. This could indicate a higher water table at this site but it more strongly reflects the lower evapotranspiration demand of that plot.

Because the 200 cm pressure head series is prescribed as the lower boundary condition (Section 5.1.2.), this wet end saturation directly limits the simulated capillary rise flux. The imposed boundary cannot become wetter than pF 2 even when the field likely did. The consequences for the inferred capillary contribution are addressed in Section VIII..

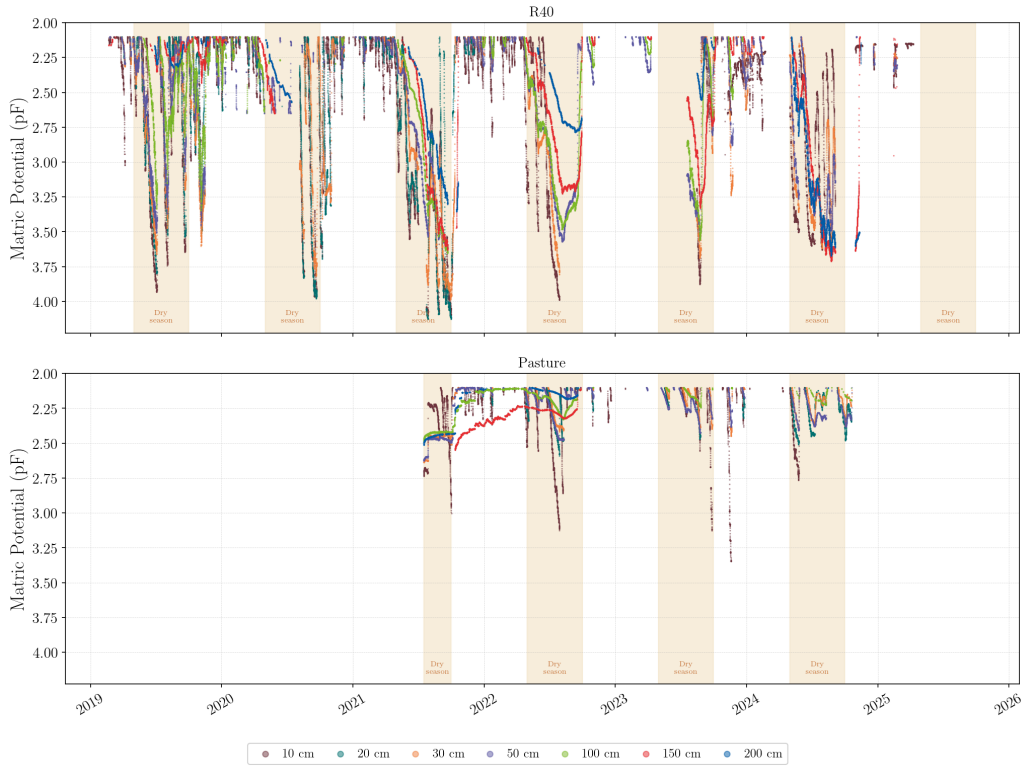


Figure 13: Matric potential measured at both sites, R40 (above) and Pasture (below). The different colors represents the different depths. The wet and dry theoretical seasons are also displayed on the graph. The dry season ranges from May to September while the wet season occupies the rest of the months.

1.2.3. Soil hydraulic properties

Depth groups were defined from the field and laboratory data (Section V.3.3). The grouped van Genuchten–Mualem parameters with their tied hysteresis values are given in Table 6, and the hysteretic retention curves are shown in Figures 14 and 15. The near-surface saturated conductivity K_s is consistently higher at R40 than at the Pasture, consistent with the structural recovery of macroporosity under 40 years of forest cover. This contrast is what motivates treating K_s as a calibrated parameter: the 250 mL rings used for the K_{SAT} measurements cannot capture the connected macropore network that governs field-scale infiltration, so the laboratory K_s is expected to be a downward-biased estimate of the effective profile conductivity (Section V.5.4).

The fitted pore-connectivity parameter l reaches the fitting lower bound of -5 in several groups. These are HYPROP-fit artifacts rather than physical estimates, the conventional Mualem value being $l \approx 0.5$. Because l is held fixed in the inverse problem, its influence on $K(h)$ was screened in the global sensitivity analysis (Section VI.2). It ranks last (*ex aequo*) of the ten parameter slots, well below K_s and α . Its low sensitivity justifies retaining the fitted values.

Table 6: Soil hydraulic parameters by depth group and by sites, obtained through laboratory measurements. The hysteretic parameters are tied to their drying counterparts ($\theta_m = \theta_{sW} = \theta_s$, $K_{sW} = K_s$), while $\alpha_W = 2\alpha$.

	Depth group [cm]	θ_r [cm ³ cm ⁻³]	θ_s [cm ³ cm ⁻³]	α [cm ⁻¹]	n [-]	K_s [cm d ⁻¹]	l [-]	θ_m [cm ³ cm ⁻³]	θ_{sW} [cm ³ cm ⁻³]	α_W [cm ⁻¹]	K_{sW} [cm d ⁻¹]
R40	10	0.171	0.646	0.0721	1.387	1962	1.39	0.646	0.646	0.1442	1962.45
	20+30	0.107	0.625	0.0903	1.230	456	-0.45	0.625	0.625	0.1806	456.38
	50+100	0.160	0.591	0.0869	1.153	336	-2.117	0.591	0.591	0.1738	336.39
	150+200	0.007	0.525	0.0699	1.131	99	-5.000	0.525	0.525	0.1398	99.23
Pasture	10	0.050	0.427	0.0459	1.103	65	2.359	0.427	0.427	0.0918	65.33
	20+30+50	0.050	0.547	0.0587	1.123	423	0.766	0.547	0.547	0.1174	423.99
	100+150	0.123	0.500	0.0349	1.180	13	-5.000	0.500	0.500	0.0698	13.51
	200	0.093	0.501	0.0417	1.167	25	-5.000	0.501	0.501	0.0834	25.50

Figures 14 and 15 overlay the fitted curves on the filtered in-situ data. Points from the different depths within a group overlap consistently, supporting the grouping. The laboratory curves capture part but not all of the field behavior. Agreement is best for the shallowest group and degrades with depth. Optimizing α should shift the curve toward the data cloud (Section V.5.4). The Pasture shows the same pattern with fewer usable points.

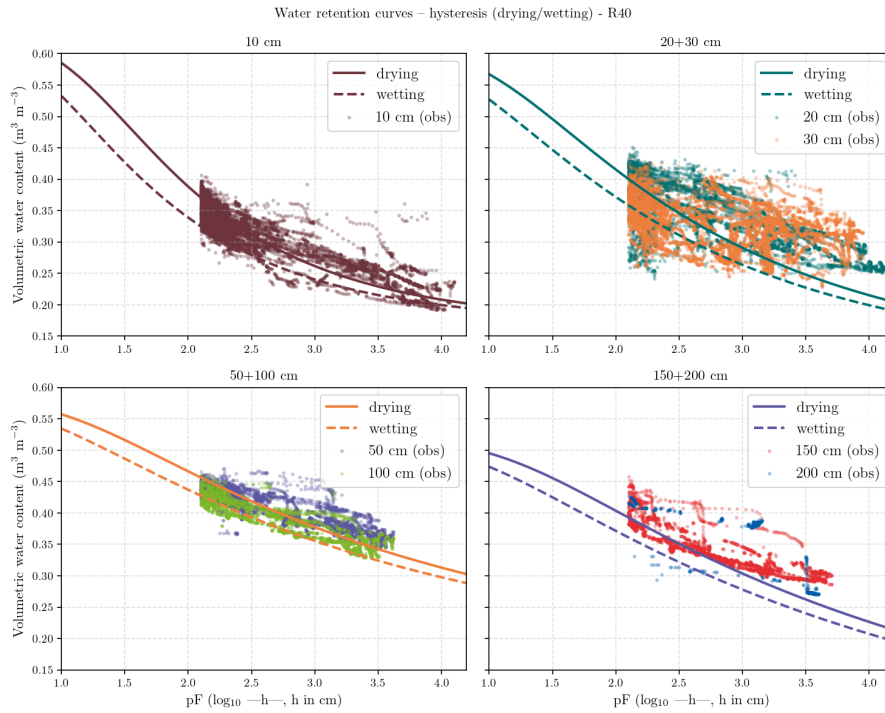


Figure 14: Laboratory retention curves by depth groups for the R40 site. The hysteretic parameters are tied to their drying counterparts : θ_m , θ_{sW} and K_{sW} have a tie of one while α_W equals to two times α . The in situ measurements are also represented, their color depends on measurements depth.

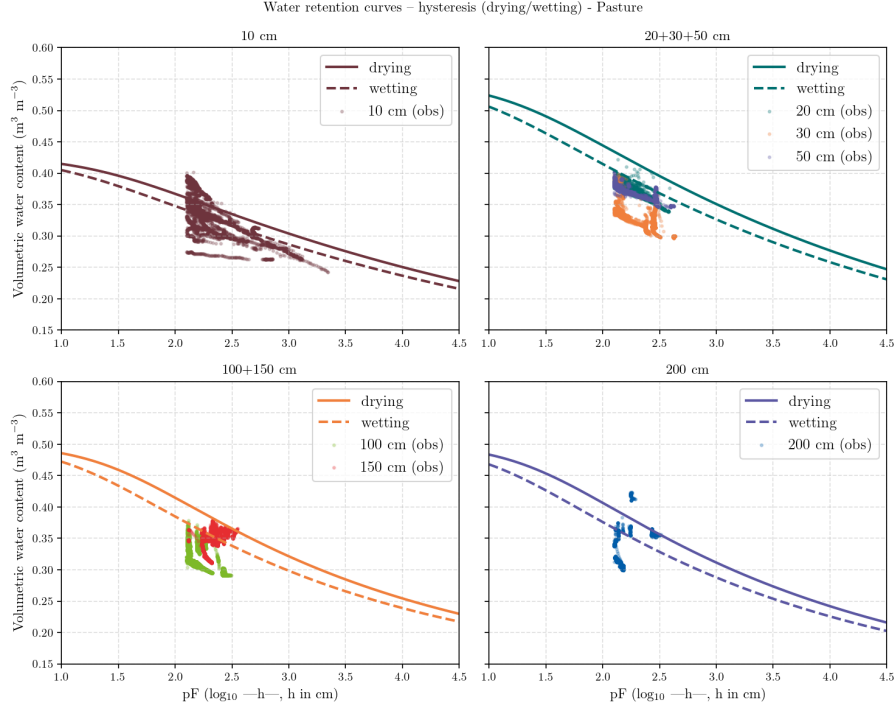


Figure 15: Laboratory retention curves by depth groups for the Pasture site. The hysteretic parameters are tied to their drying counter parts : θ_m , θ_{sW} and K_{sW} have a tie of one while α_W equals to two times α . The in situ measurements are also represented, their color depends on measurements depth.

1.3 Vegetation parameters

1.3.1. Root density

Fine-root density (Figure 16, after Dermauw (2024)) decreases with depth at both plots, with most fine roots in the 0 cm to 10 cm layer (50 % at R40, 46 % at the Pasture). The Pasture shows a higher fine-root density throughout the profile. This is not anomalous as grasses concentrate biomass in a dense, shallow fine-root mat, whereas trees invest proportionally more in coarse woody roots distributed deeper (Jackson et al., 1996). In the topsoil a grassland can therefore match or exceed a forest in fine root density even though the forest holds far greater total root biomass. In the model the absolute densities are not determinant as only the *shape* of the normalized distribution $b(z)$ enters Equation 20. Both profiles decreases at roughly the same rate with depth.

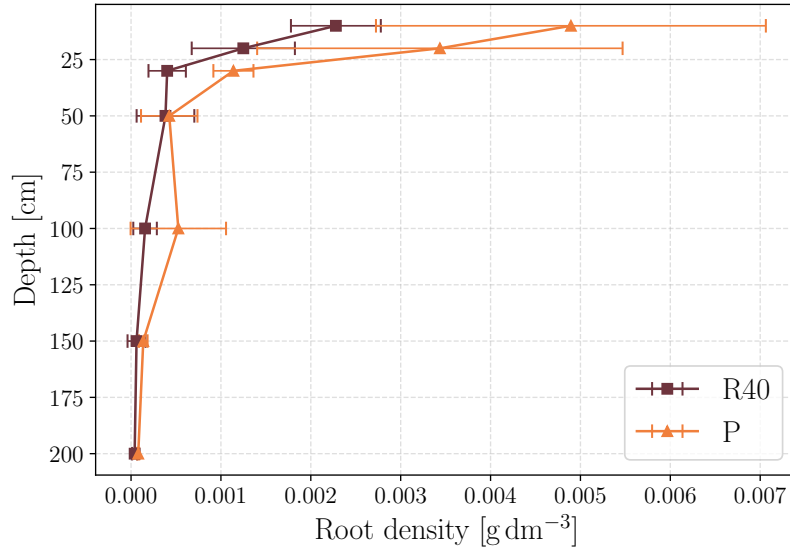


Figure 16: Vertical soil profile of the fine root density [g dm^{-3}] for R40 (in brown) and Pasture (in orange). As the center value is the mean of triplicates for each depths, the standard deviations are shown for each data point.

1.3.2. Leaf area index

The LAI time series for both plots are shown in Figure 17, monthly box plots are available in Appendix D. Over the study period R40 yielded 149 valid retrievals and the Pasture 152. R40 averages $2.47 \text{ m}^2 \text{ m}^{-2}$ annually, with a seasonal contrast ($2.28 \text{ m}^2 \text{ m}^{-2}$ in the dry season, $2.74 \text{ m}^2 \text{ m}^{-2}$ in the wet with a range from $1.08 \text{ m}^2 \text{ m}^{-2}$ to $3.55 \text{ m}^2 \text{ m}^{-2}$). The Pasture averages $1.20 \text{ m}^2 \text{ m}^{-2}$ with a much stronger seasonal swing ($0.85 \text{ m}^2 \text{ m}^{-2}$ dry season, $1.65 \text{ m}^2 \text{ m}^{-2}$ wet season with a range from $0.32 \text{ m}^2 \text{ m}^{-2}$ to $2.42 \text{ m}^2 \text{ m}^{-2}$).

The forest's modest seasonal drop, $\sim 20\%$ between seasons, is consistent with an evergreen canopy that retains much of its foliage through the dry season, whereas the Pasture's larger drop, $\sim 50\%$, reflects herbaceous senescence. The forest's dry season minimum stays well above the Pasture's wet season maximum, confirming their distinct structure. The higher R40 LAI increases interception and shifts the evaporation–transpiration partition toward transpiration (Eq. 19).

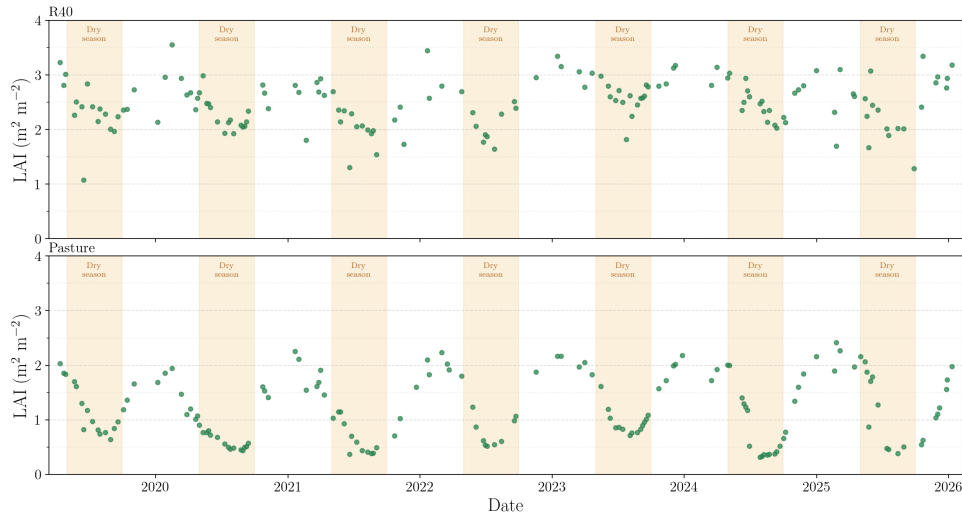


Figure 17: Time series of leaf area index [$m^2 m^{-2}$] values for the R40 (above) and the Pasture (below) plots. The wet and dry theoretical seasons are also displayed on the graph. The dry season ranges from May to September while the wet season occupies the rest of the months.

1.3.3. Canopy interception

R40

The Von Hoyningen-Hüne and Braden interception model (Equation 21) was calibrated against the 18 pluviometer measurement periods collected on the R40 site between December 2020 and January 2022. As described in Section V.4.4, the objective function minimized the sum of squared errors between the predicted and observed period-level interception fractions (I/P), which avoids any dependence on the absolute rainfall reference and renders the calibration consistent with the meteorological forcing used in the simulation.

The calibrated storage capacity parameter is $a = 3.59 \text{ mm d}^{-1}$. The goodness of fit over the 18 calibration periods is shown in Figure 18. A Nash–Sutcliffe efficiency of $NSE = 0.53$ indicates that the model captures the majority of the inter-period variability in interception, although a non-negligible scatter remains, reflecting both model structural limitations and the inherent spatial variability of throughfall measured by a fixed gauge network.

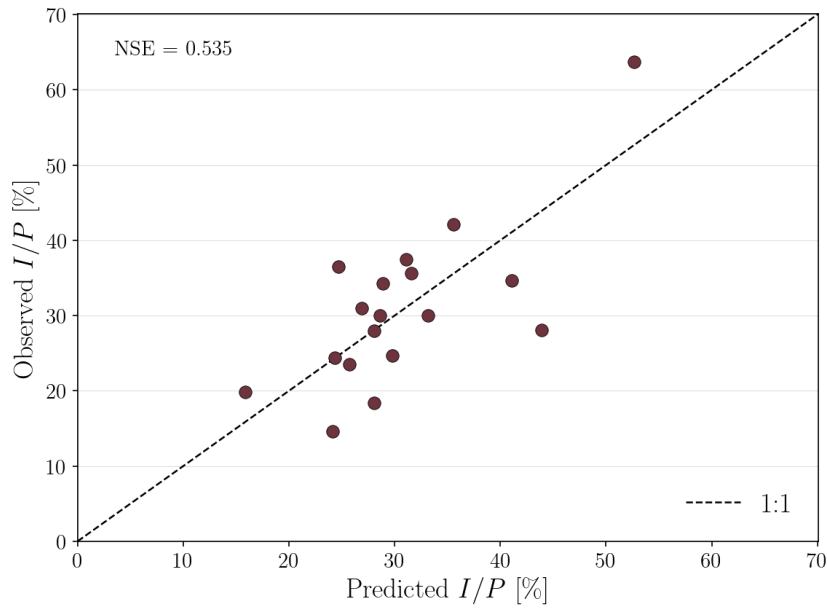


Figure 18: Goodness of fit of the calibrated Von Hoyningen-Hüne Braden interception model at R40 over the 18 pluviometer periods. Observed period interception is plotted against the model prediction. The dashed line is the 1:1 reference.

The calibrated value of a is higher than the usual values for crop field. But because HYDRUS-1D usage on forests is unusual, no values for a on forested sites were available in the literature. Because of that, comparison on the a parameter itself cannot be done. Several factors could have biased the calibrated value nevertheless. First, calibration was performed at the period level rather than at the event level, period aggregation is known to inflate the apparent canopy storage because the model cannot distinguish individual storm filling and draining cycles within a period. Second, despite the ratio-based objective function adopted here, the systematic under catch of the meteorological station relative to the pluviometer gauges, with a mean scale factor ≈ 1.5 , implies that the model is driven by lower daily rainfall depths than those that actually reached the canopy, which pushes the fitted a upward to compensate. The calibrated value should therefore be interpreted as an effective parameter tuned to reproduce observed interception fractions under the available meteorological forcing, rather than a direct physical estimate of canopy storage capacity.

The calibrated model was applied to the full meteorological record from January 2019 to December 2025 using daily rainfall and LAI. The resulting annual totals are presented in Figure 19 and summarized below. The daily interception–rainfall relationship is provided in Appendix E.

Over the almost seven-year period, total gross rainfall amounted to 13 879 mm, of which 4 160 mm (30.0%) was attributed to canopy interception. Mean annual rainfall and interception were $1\,983\text{ mm yr}^{-1}$ and 594 mm yr^{-1} , respectively. Annual interception fractions ranged from 27.8% in 2021 to 34.4% in 2025. Interannual variability is therefore moderate, reflecting the relatively stable LAI and the non-linear but smooth response of the model to changing rainfall amounts (Figure 19).

A consistent seasonal signal is present. Interception fractions are slightly higher during the

dry season (months May–September; mean $I/P = 32.9\%$) than during the wet season (mean $I/P = 29.5\%$). This counterintuitive result arises from the non-linearity of the model. Dry season events tend to be smaller and more numerous, keeping the model in its near-linear regime where each millimeter of rainfall generates a higher relative interception, whereas large wet-season events approach canopy saturation and reduce the fractional loss.

The simulated interception fraction of approximately 30% is notably higher than the 10–20% range commonly reported for comparable density canopies in the literature (Section III.1). This discrepancy is a direct consequence of the elevated calibrated a value discussed above. Given the measurement constraints, achieving a more physically consistent parameter estimate was not possible with the available data. The approach adopted here minimizes the calibration bias as much as possible by calibrating directly on dimensionless interception fractions, ensuring internal consistency between the simulated fluxes and their meteorological drivers. The resulting interception estimates should nonetheless be interpreted with caution, and the absolute simulated values are likely to overestimate the true canopy loss. This could have a large impact on the simulated hydrological behavior of the site. The impact of this potential overestimate is evaluated in the limitations in Section VIII.

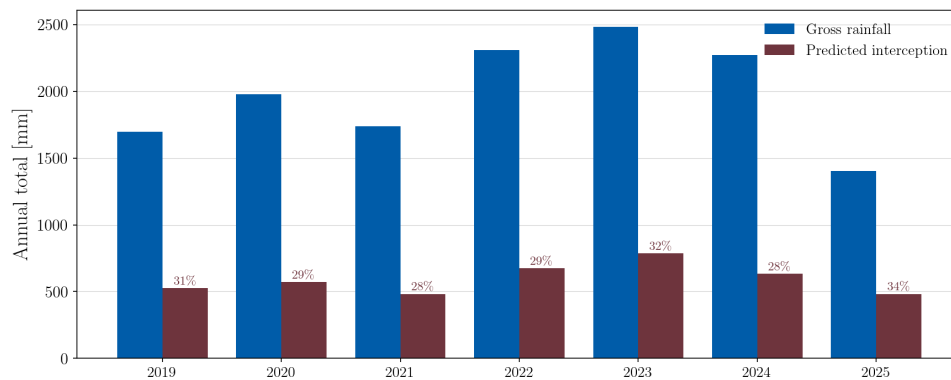


Figure 19: Gross rainfall and predicted canopy interception at R40, 2019–2025. Labels give the annual interception fraction I/P .

Pasture

As no field interception measurements were available for the Pasture site, the storage capacity parameter was therefore set to the recommended value for ordinary agricultural crops and grasses of $a = 0.25 \text{ mm d}^{-1}$. All diagnostic figures for this site are provided in Appendix E.

Over the seven-year period, total gross rainfall amounted to 13 879 mm, of which 404 mm (3%) was attributed to canopy interception. Mean annual interception was 58 mm yr^{-1} . Annual interception fractions ranged from 2.5% in 2020 to 3.9% in 2025, with a seven-year mean of 3.0%. The seasonal contrast is negligible: the dry-season mean I/P of 3.0% is virtually identical to the wet-season value of 2.9%, indicating that the seasonal rainfall regime has little effect on the interception fraction at this low storage capacity.

The simulated interception fraction of approximately 3% is an order of magnitude below the value for R40. This low value is a direct consequence of the small effective storage capacity

$a \cdot \overline{\text{LAI}} \approx 0.25 \text{ mm}$: for any daily rainfall event exceeding approximately 1 mm, the canopy saturates almost immediately and the residual rainfall passes through as throughfall. The interception loss is therefore nearly independent of event size for the majority of rainy days, and the annual total is determined almost entirely by the number of rainy days rather than their intensity. The main source of uncertainty is the LAI time series, whose magnitude sets the effective storage capacity and the canopy cover fraction b .

2 Sensitivity analysis

The Morris screening ranks the forty perturbed parameters by their mean absolute elementary effect μ^* , averaged over the four outputs (RMSE_θ , RMSE_{pF} , D_{cum} and ET_{cum}) and over the four material layers. The resulting per-type ranking is given in Table 7. The results for every parameters across all outputs for both sites, represented as a heat map of μ^* , is available in Appendix F.

Before these ranks are interpreted, a numerical caveat must be raised. The hysteretic parametrization proved to be a source of instability. Of the 2050 HYDRUS-1D runs generated by the Morris design, only 840 ran to completion at R40 (41%) and 204 at the Pasture (10%), the remainder failing to converge. Because only the convergent trajectories enter the index, the surviving subset may not sample parameter space uniformly. The Pasture ranking should therefore be read as indicative rather than quantitative, and is reported chiefly for its agreement with the more completely sampled R40 screening. The sensitivity indices are therefore computed over the completed subset of trajectories. The markedly lower completion rate at the Pasture foreshadows the convergence difficulties that the hysteretic tied-parameter set raised more generally, and that were only resolved during the Bayesian calibration. This failure rate is itself a result: it confirms that the wetting-branch parameters affect the numerical behavior of the model and not merely the shape of the retention curve.

Despite the reduced sample, the ranking is remarkably consistent between the two land uses, which lends confidence to the screening. At both sites the shape exponent n is the single most influential parameter, followed by the saturated hydraulic conductivity. These two parameters dominate both the calibration fit outputs and the cumulative water-balance fluxes, confirming that the partitioning between drainage and evapotranspiration is governed primarily by the steepness of the retention curve and by the conductivity at the wet end. A middle tier groups the residual and saturated water contents (θ_r , θ_s), the air-entry parameter α and its wetting-branch ratio r_{α_w} , all with comparable μ^* and no stable ordering between sites. At the bottom, the wetting saturated conductivity ratio $r_{K_{sw}}$ and the Mualem pore-connectivity exponent l are jointly the least influential parameters at both sites, exerting no detectable effect on any of the four outputs.

These results directly inform the parameter selection of the inverse problem. Although n carries the largest elementary effect, it is excluded from optimization as it is tightly constrained by the full drying curves measured with Hyprop, so refining it against the in situ data would

add a strongly sensitive but already well identified dimension to the calibration without improving physical realism. K_s , by contrast, is both highly sensitive and subject to the core to field scale dependence discussed in Section 5.4.1., which makes it the prime candidate for inverse estimation. The air-entry parameter α occupies only an intermediate rank, yet is retained for optimization because it propagates to the wetting branch through the tied relation $\alpha_w = 2\alpha$ and because the Labros fits supply an informative prior, optimizing α therefore adapts the near-saturation retention to field conditions at no additional cost in free parameters. The eight optimized parameters (K_s and α across the four layers) thus combine the most sensitive identifiable parameter with the one most affected by scale.

The screening also settles the open question raised in the design of the analysis (Section V.5.2) concerning the wetting saturated conductivity K_{sw} . Although HYDRUS-1D requires this parameter even when hysteresis is restricted to the retention curve, its ratio $r_{K_{sw}}$ ranks last together with l , confirming that it has no bearing on the simulated outputs. Both are accordingly held fixed at their base values in the subsequent calibration. Among the wetting-branch ratios that do matter, the saturated water content ratio $r_{\theta_{sw}}$ is the most influential, reflecting its direct control on the amount of water stored near saturation along the wetting path, whereas the air-entry and θ_m ratios remain secondary.

Table 7: Overall ranks of the Morris screening of the parameters for both sites. The average rank is the mean by parameter type for all depth group. The individual parameter own rank are the mean rank over the four outputs evaluated: $RMSE_\theta$, $RMSE_{pF}$, D_{cum} and ET_{cum} . The rank of the parameters is based on μ^* values, the mean absolute elementary effect.

Rank	R40		Pasture	
	Variable	Average rank	Variable	Average rank
1	n	8.69	n	5.50
2	K_s	10.12	K_s	7.81
3	$r_{\theta_{sw}}$	10.56	θ_r	12
4	θ_r	18.12	$r_{\theta_{sw}}$	16.19
5	θ_s	18.50	α	21.19
6	α	21.19	θ_s	21.81
7	r_{α_w}	22	r_{θ_m}	23.56
8	r_{θ_m}	22.81	r_{α_w}	23.94
9	$r_{K_{sw}}$	33	$r_{K_{sw}}$	33
10	l	33	l	33

3 Structural model choices

The near-saturation behavior of the retention function was examined first, as it conditions the numerical stability on which the remaining analyses depend. Table 8 compares the standard van Genuchten formulation ($h_{air} = 0$) with the regularized form introducing a small air-entry

head $h_{\text{air}} = -2$ cm (Section 5.3.2.). The VGM parameters used here are the non-hysteretic laboratory values, retained for this screening because the hysteretic set was less numerically stable. The regularization improves the fit at R40 (R^2_{VWC} rising from 0.48 to 0.57 and R^2_{pF} from 0.49 to 0.54) and is decisive at the Pasture, where the standard formulation does not converge at all. By removing the vertical tangent of the retention curve as $h \rightarrow 0^-$, it also reduced the number of iterations needed per time step and shortened run times appreciably. The regularized formulation is therefore adopted for both sites.

Table 8: Coefficients of determination (R^2) for in situ VWC and pF, with and without the near-saturation air-entry regularization ($h_{\text{air}} = -2$ cm), for the R40 and Pasture sites.

	R40		Pasture	
	R^2 VWC	R^2 pF	R^2 VWC	R^2 pF
With $h_{\text{air}} = -2$ cm	0.5701	0.5410	0.1348	0.1427
Without ($h_{\text{air}} = 0$)	0.4843	0.4896	did not converge	

With the air-entry regularization in place, the degree of root-water-uptake compensation was screened by varying the critical stress index ω from 1.0 (no compensation) to 0.5 (Table 9). At R40, increasing the compensation (lowering ω) lets roots offset local water stress by drawing more from wetter portions of the profile, yet this degrades the fit monotonically: R^2_{VWC} falls from 0.570 at $\omega = 1.0$ to 0.565 at $\omega = 0.5$, and R^2_{pF} from 0.541 to 0.526. At the Pasture the fit is essentially insensitive, activating compensation shifts R^2 only at the fourth decimal. It saturates immediately, consistent with a shallow root system that holds little deeper water from which to compensate. In both cases the total spread remains small (below 0.01 for R^2_{VWC} and below 0.015 for R^2_{pF}), indicating that water stress is rarely active over the simulated period and that the uptake scheme exerts only a marginal control on the simulated states.

Table 9: Coefficients of determination (R^2) for in situ VWC and pF as a function of the critical stress index ω for root water uptake, for the R40 and Pasture sites.

Critical stress index ω	R40		Pasture	
	R^2 VWC	R^2 pF	R^2 VWC	R^2 pF
1.0	0.5701	0.5410	0.1348	0.1427
0.9	0.5692	0.5390	0.1350	0.1431
0.8	0.5678	0.5339	0.1350	0.1431
0.7	0.5665	0.5298	0.1350	0.1431
0.6	0.5655	0.5276	0.1350	0.1431
0.5	0.5647	0.5263	0.1350	0.1431

Independently of these structural choices, the Pasture fit ($R^2_{\text{VWC}} \approx 0.13$) is markedly weaker than that of R40 ($R^2_{\text{VWC}} \approx 0.57$). Neither scenario examined here closes this gap, whose likely

origin is discussed in Section VIII.. To preserve comparability between the plots, and because the non compensated scheme is the best supported configuration at R40 while being almost indifferent at the Pasture, all subsequent simulations use the air-entry regularization $h_{\text{air}} = -2$ cm and a critical stress index $\omega = 1.0$.

Finally, because the calibrated canopy storage capacity produced an interception fraction of about 30%, well above the values usually reported in the literature (Section III.1), the influence of this parameter on the simulation was tested by varying a across a range of values (Table 10). Although the mean annual interception fraction rose from 4.8% at $a = 0.25 \text{ mm d}^{-1}$ to 30% at the calibrated value of $a = 3.59 \text{ mm d}^{-1}$, the effect on fit quality was negligible: R^2_{VWC} stayed within 0.569–0.570 and R^2_{pF} within 0.538–0.541 over the entire range. The simulated water content and matric potential are therefore largely insensitive to the amount of intercepted rainfall, at least over the range tested and under the in-situ evaluation used here. The marginally best fit was in fact obtained at the calibrated value, which was consequently retained for all subsequent runs.

The insensitivity of the model to the a parameter, and thus the total interception, was very surprising. After a deeper analysis of the different fluxes, the explanation lies in the way HYDRUS-1D implements interception. The model computes the intercepted amount I_c from the canopy storage model and subtracts it from the potential transpiration demand T_p , leaving the potential soil evaporation and the water reaching the soil surface unchanged. Interception is thus treated as transpiration that is satisfied by water evaporating from the canopy rather than withdrawn from the root zone, so it does not reduce the throughfall that infiltrates. Since actual transpiration closely tracks potential transpiration when the soil is not water-stressed, raising the interception simply lowers actual transpiration by almost the same amount, with no compensating change in infiltration. The intercepted water therefore leaves the demand side of the balance rather than the supply side, which is why the simulated soil water state, and the drainage budget, remain nearly unaffected by the interception parameter (Šimůnek et al., 2024). This same mechanism explains the near-equal infiltration of the two plots despite their very different canopy interception, discussed in Section VI.6.

Table 10: Coefficients of determination (R^2) for in situ VWC and pF, and the resulting mean annual interception fraction, as a function of the canopy storage capacity parameter a , at the R40 site.

$a \text{ [mm d}^{-1}\text{]}$	R^2_{VWC}	R^2_{pF}	Avg. yearly intercepted [%]
0.25	0.5686	0.5385	4.8
0.50	0.5687	0.5384	8.2
1	0.5691	0.5384	13.6
2	0.5696	0.5388	21.4
3	0.5700	0.5401	27.2
3.59	0.5701	0.5410	30

Two points should be kept in mind for this scenario analysis as a whole. First, the R^2 values

are computed against the in-situ probe measurements, which are themselves subject to uncertainty. They should therefore be read not as an absolute assessment of the hydrological behavior at each site, but as the best available means of ranking the configurations relative to one another. Second, all runs in this section were performed with the non-hysteretic laboratory parameters. The hysteretic tied-parameter set was not used here because it caused convergence failures, a problem subsequently resolved during the optimization. Because these scenarios necessarily preceded the calibration, they could not rely on the optimized parameters, had those been available, the relative behavior of the configurations might have differed. The choices made here therefore reflect the best information available at this stage of the workflow.

4 Inverse modeling

The eight hydraulic parameters and the two measurement-error hyperparameters were estimated with DREAM(ZS) against the joint in-situ VWC and pF records (Section V.5.4). Each site was sampled with five chains of 1120 steps. The first quarter was discarded as burn-in, leaving 4200 posterior samples. The laboratory priors are informative, their bounds substantially tighter than the generic $\sigma_{\log} = 0.35$ prior, so the updating below reflects information added by the field data rather than prior tightening.

R40

The sampler converged but mixed slowly. The Gelman–Rubin statistic stays below the adopted $\hat{R} < 1.1$ threshold for every parameter, indicating convergence and reaching at most 1.014 for the physical parameters and 1.024 for the error hyperparameters, at a mean acceptance fraction of 0.65. Effective sample sizes are nonetheless modest, between 20 and 76 (Table 11), reflecting the strong pairwise correlations in the posterior. The well-identified parameters settle within the first hundred steps, whereas the surface K_s and α wander throughout. The traces and the full prior-to-posterior updating are in Appendix G (Figures 36 and 35). The credible intervals below should be read with the corresponding Monte-Carlo uncertainty, especially for the surface parameters.

Identifiability is strongly depth-dependent, and the contraction ratio $\sigma_{\text{post}}/\sigma_{\text{prior}}$ (Table 11) must be read with care wherever a posterior abuts a bound. The 10 cm surface layer is essentially unconstrained as its K_s and α contract to 2.34 and 1.13, no narrower than the prior and, for K_s , broader. This is the signature of the negative K_s – α correlation seen for that layer in the corner plot (Figure 20), the two trading off along a ridge the data cannot resolve. The deeper α are genuinely informed, contracting to 0.77 (α_2), 0.13 (α_3) and 0.01 (α_4), and $K_{s,2}$ likewise (0.37). The deep conductivities $K_{s,3}$ and $K_{s,4}$ only appear well constrained as their low contraction (0.02 and 0.11) is bound truncation, not identification.

Table 11: Posterior estimates of the eight van Genuchten–Mualem parameters and the two measurement-error hyperparameters for R40, obtained from 4200 DREAM(ZS) samples. The median (p50) and the 95 % credible interval (p2.5–p97.5) are reported, together with the Gelman–Rubin statistic $\hat{R}$, the effective sample size (ESS) and the prior-to-posterior contraction ratio $\sigma_{\text{post}}/\sigma_{\text{prior}}$. K_s and α were sampled in $\log_{10}$ space.

Parameter	Mat.	Median	95 % CI	$\hat{R}$	ESS	$\sigma_{\text{post}}/\sigma_{\text{prior}}$
K_s [cm d ⁻¹]	1	698	[152, 1306]	1.006	24	2.34
	2	4740	[1801, 7064]	1.005	33	0.37
	3	9920	[9662, 10000]	1.011	63	0.02
	4	8253	[6825, 10000]	1.003	20	0.11
α [cm ⁻¹]	1	0.0656	[0.0617, 0.0881]	1.009	50	1.13
	2	0.0834	[0.0757, 0.1007]	1.006	36	0.77
	3	0.1101	[0.1061, 0.1128]	1.014	76	0.13
	4	0.0581	[0.0580, 0.0589]	1.013	47	0.01
σ_{VWC} [m ³ m ⁻³]	—	0.0516	[0.0492, 0.0554]	1.007	24	0.02
σ_{pF} [pF]	—	0.645	[0.610, 0.689]	1.024	57	0.02

The conductivity estimates are not trustworthy at depth. The laboratory cores decrease monotonically with depth (Table 6), as expected of a forest topsoil whose structure and macroporosity concentrate near the surface (Section III.3). The calibration inverts this profile. The surface K_s is pulled below its core value, while the subsurface and deep layers rise by one to nearly two orders of magnitude and materials 3 and 4 press against the upper bound of 10⁴ cm d⁻¹. A deep conductivity of that order is implausibly high, and together with the inversion it marks a numerical artifact rather than a recovered property.

Two mechanisms explain it. First, the deep layers sit on the prescribed-head lower boundary at 200 cm, which stays wet throughout. Raising $K_{s,3}$ and $K_{s,4}$ couples the deep nodes more tightly to that imposed head, so the likelihood keeps rewarding higher conductivity with no interior optimum: the chains climb until the bound stops them, and the tight posterior is the bound truncating an open distribution rather than information gained. This is consistent with the Morris screening, where both deep conductivities rank among the influential parameters: they matter to the fit, but only by forcing the profile toward the boundary, not by converging on a physical value. Widening the bound would not change this. Second, the problem is high-dimensional: ten correlated parameters and effective sample sizes of 20–76 place the run at the lower edge of adequate sampling, where $\hat{R} < 1.1$ is necessary but not sufficient, so incomplete exploration along the surface ridge remains possible. More chains would sharpen the surface estimates but not resolve the deep conductivities, whose non-identifiability is structural. The deep K_s should therefore be read as bound-limited, not as recovered properties.

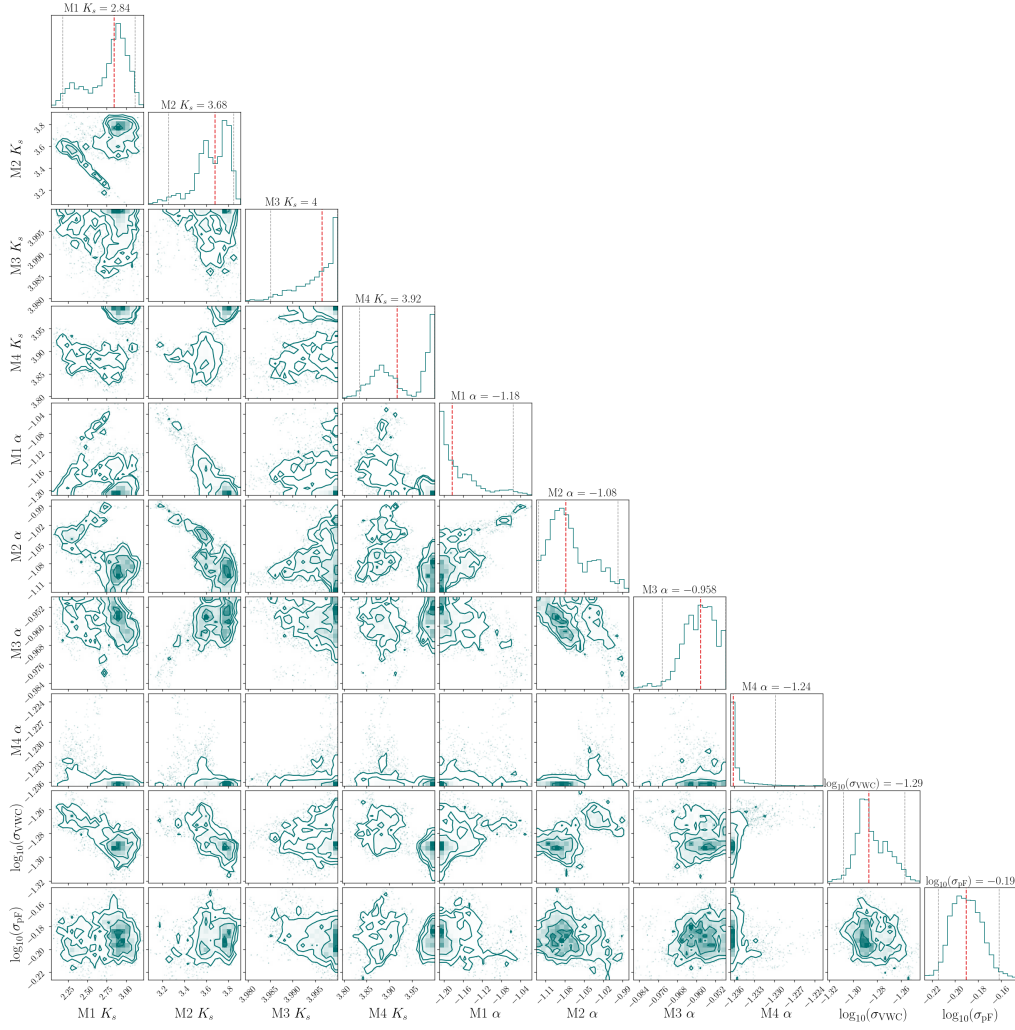


Figure 20: Posterior marginal distributions and pairwise correlations of the ten calibrated parameters at R40. Dashed lines on the marginals indicate the median and the 2.5th and 97.5th percentiles. K_s and α are shown on a $\log_{10}$ scale. The marginals of $M3 K_s$ and $M4 K_s$ pile against the upper conductivity bound, $M4 \alpha$ collapses to a near-point posterior, and the within-layer K_s - α pairs are correlated.

Parameter α moves modestly and stays interior to its bounds, so its posteriors are genuine rather than truncated. The surface and subsurface values (α_1, α_2) remain close to their laboratory estimates, leaving the near-surface retention essentially as measured, whereas the deeper α_3 is pulled above its core value and α_4 collapses to a tight value just below it. The calibration thus preserves the shallow retention shape and reworks only the deep air-entry, plausibly compensating for the deep conductivities that are themselves pinned at the bound. The two error standard deviations, $\sigma_{\text{vwc}} = 0.052 \text{ m}^3 \text{ m}^{-3}$ and $\sigma_{\text{pF}} = 0.65 \text{ pF}$, are both larger than the instruments alone would produce (σ_{vwc} against a sensor precision of $\pm 0.03 \text{ m}^3 \text{ m}^{-3}$, Table 1), so they absorb structural model error as well as measurement noise. The much larger σ_{pF} indicates that matric potential is reproduced less faithfully than water content.

Pasture

At Pasture the sampler mixed less freely, with a mean acceptance fraction of 0.23, but still converged. $\hat{R}$ stays below 1.030 for every parameter (Table 12) and the chains stabilize after

burn-in. Effective sample sizes are again modest, between 21 and 74. The traces and the full prior-to-posterior updating are in Appendix G (Figures 37 and 38).

The contraction ratios (Table 12) broadly echo R40, with one difference at the surface: the material 1 K_s is identified, contracting to 0.52 and shifting cleanly upward from its core value, and α_1 is also informed with 0.30. Among the α parameters, α_2 is the most tightly determined (0.05) and α_3 barely departs from its prior (0.82). As at R40, the deep conductivities only appear constrained: $K_{s,2}$ and $K_{s,3}$ (0.04 and 0.15) are pinned against the upper bound, and $K_{s,4}$ (0.32) is informed but broad, its 95 % interval reaching the same ceiling.

Table 12: Posterior estimates of the eight van Genuchten–Mualem parameters and the two measurement-error hyperparameters for Pasture, obtained from 4200 DREAM(ZS) samples. The median (p50) and the 95 % credible interval (p2.5–p97.5) are reported, together with the Gelman–Rubin statistic $\hat{R}$, the effective sample size (ESS) and the prior-to-posterior contraction ratio $\sigma_{\text{post}}/\sigma_{\text{prior}}$. K_s and α were sampled in $\log_{10}$ space.

Parameter	Mat.	Median	95 % CI	$\hat{R}$	ESS	$\sigma_{\text{post}}/\sigma_{\text{prior}}$
K_s [cm d^{-1}]	1	785	[720, 1333]	1.025	74	0.52
	2	9189	[8503, 9867]	1.011	22	0.04
	3	9534	[7923, 10000]	1.025	38	0.15
	4	4573	[3042, 9994]	1.002	23	0.32
α [cm^{-1}]	1	0.0377	[0.0377, 0.0414]	1.005	45	0.30
	2	0.0704	[0.0699, 0.0711]	1.005	33	0.05
	3	0.0447	[0.0336, 0.0458]	1.001	30	0.82
	4	0.0366	[0.0317, 0.0465]	1.030	26	0.28
σ_{VWC} [$\text{m}^3 \text{m}^{-3}$]	—	0.0343	[0.0301, 0.0364]	1.011	21	0.05
σ_{pF} [pF]	—	0.207	[0.184, 0.231]	1.023	22	0.05

The same artifact recurs, more strongly. The deep layers again run to the 10^4 cm d^{-1} ceiling for the reason given for R40, being wet, weakly dynamic and clamped to the prescribed boundary. Because the Pasture cores are very low below 50 cm (13 and 25 cm d^{-1} for materials 3 and 4, Table 6), the upward revision spans two to nearly three orders of magnitude, with materials 2 and 3 pinned at the ceiling and material 4 reaching it within its interval. The surface conductivity, though genuinely identified, still settles near 785 cm d^{-1} , an order of magnitude above its core. The deep K_s are again bound-limited and carry no weight for the propagation.

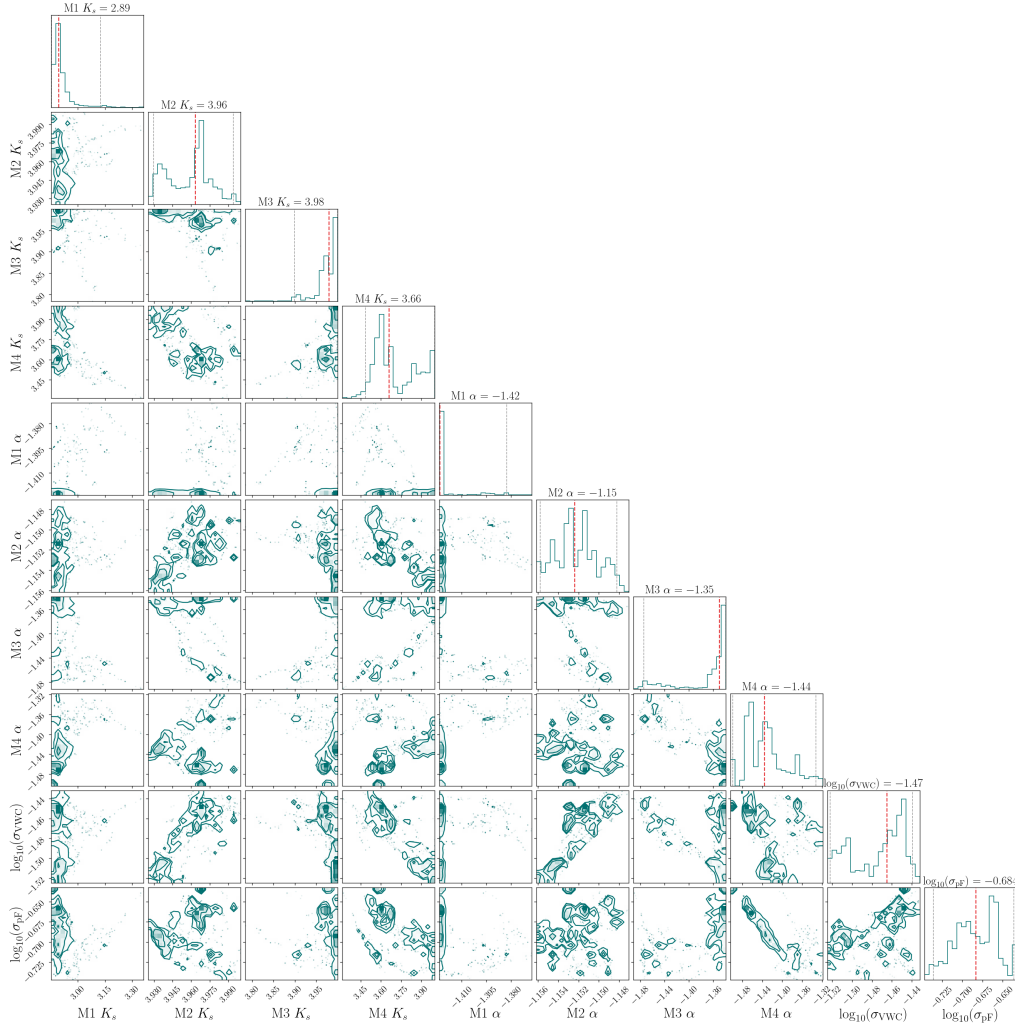


Figure 21: Posterior marginal distributions and pairwise correlations of the ten calibrated parameters at Pasture. Dashed lines on the marginals indicate the median and the 2.5th and 97.5th percentiles. K_s and α are shown on a $\log_{10}$ scale. The marginals of $M2 K_s$ and $M3 K_s$ pile against the upper conductivity bound and $M4 K_s$ reaches it within its credible interval, whereas the surface conductivity $M1 K_s$ forms a well-defined interior peak.

Parameter α again stays interior to its bounds, so its posteriors are genuine rather than truncated, but the updating shows no clear shallow-to-deep gradient as at R40. Settling above its core value, α_2 is the one parameter the data pin down tightly, while α_3 barely departs from its prior and α_1 and α_4 shift only modestly downward from their cores. The two error standard deviations, $\sigma_{\text{vwc}} = 0.034 \text{ m}^3 \text{ m}^{-3}$ and $\sigma_{\text{pf}} = 0.21 \text{ pF}$, are both smaller than at R40, with σ_{vwc} close to the sensor precision (Table 1). These smaller residuals follow from the narrower dynamic range of the Pasture series, which never reaches R40's dry extremes, rather than from a better calibration.

5 Model validation

For each site, the 200 parameter sets drawn from the joint posterior (Section V.5.4) were propagated through HYDRUS-1D over the full simulation period. The 2.5th, 50th and 97.5th percentiles of the ensemble were assembled into a posterior-median trajectory bracketed by a

95 % predictive interval. The resulting VWC and MP time series are shown depth by depth in Appendix H. Goodness of fit is quantified by the Pearson R^2 of the posterior median against the in-situ records, both per depth and pooled across all depths (Table 13), and by simulated against observed scatter plots (Figures 22 and 23). For R40, the calibration window is distinguished from the subsequent record, which was never seen by the sampler and therefore constitutes a genuine out-of-sample validation. For Pasture the posterior propagation coincides with the calibration window, so only a calibration fit is reported and no independent validation window is evaluated.

Table 13: Coefficients of determination R^2 of the posterior-median simulation against the in-situ VWC and MP records, by depth and pooled (Overall). R40 is split into calibration (days 343–1073) and validation (days 1074–2534). Pasture is calibration only, evaluated over days 366–1006. The 200 cm row reflects the prescribed lower boundary condition rather than predictive skill. *n/a*: no validation-period matric-potential observations at that depth.

Depth [cm]	R40				Pasture	
	VWC		pF		VWC	pF
	Calibration	Validation	Cal	Val	Cal	Cal
10	0.71	0.50	0.59	0.56	0.38	0.05
20	0.39	0.09	0.53	<i>n/a</i>	0.45	0.07
30	0.54	0.41	0.59	0.61	0.57	0.32
50	0.75	0.70	0.79	0.66	0.58	0.43
100	0.72	0.56	0.92	0.57	0.61	0.55
150	0.63	0.69	0.78	0.64	0.56	0.02
200	0.53	0.02	1.00	0.99	0.00	1.00
Overall	0.56	0.48	0.60	0.59	0.11	0.16

R40

The model reproduces the observed dynamics well at this site. The overall calibration R^2 reaches 0.56 for VWC and 0.60 for pF, and the validation scores remain close to them (0.48 and 0.59 respectively). This near-parity between the calibration and the out-of-sample windows is the most informative single result. The parameters inferred from the 2019–2021 record continue to reproduce the 2022–2025 record they were never fitted to, so the fit is not an artifact of overfitting the calibration window. The agreement is strongest in the mid-profile, where the 50 and 100 cm layers exceed $R^2 = 0.7$ in calibration and remain above 0.55 in validation. It is weaker at the surface and at 20 cm, where the faster, sharper near-surface oscillations are only partially resolved. The much larger σ_{pF} inferred during calibration (Section VI.4) is consistent with this picture as MP is reproduced less faithfully than water content over most of the profile.

The two 200 cm rows require care. The near-perfect pF score at that depth (1.00 / 0.99) is mechanical rather than predictive as the 200 cm pressure head is the prescribed lower boundary.

It is reproduced by construction. It nonetheless enters the pooled R^2 and inflates the overall pF figure slightly, so the pooled pF should be read with that caveat. The corresponding VWC at 200 cm tells the opposite and more telling story. It collapses from 0.53 in calibration to 0.02 in validation, because the deep water content is slaved to the prescribed pressure head through the retention curve, and that head is pinned near the wet sensor limit ($\approx$ pF 2). The bottom of the profile is therefore held in a near-saturated, almost stationary state that cannot follow the genuine variability recorded by the deep VWC probe.

This points to the principal structural limitation of the calibration. The prescribed 200 cm pressure head is of limited quality. The MP sensors having saturated near pF 2 during the wet season and being unable to record the wetter states the field likely reached. Because the deep nodes are clamped to this near-constant boundary, the inversion is pulled toward parameter combinations that satisfy the imposed head rather than the interior dynamics, which is the same mechanism that drove the deep K_s estimates onto their upper bound (Section VI.4) and that caps the upward capillary flux the lower boundary can sustain. The poor boundary thus simultaneously biases the deep parameter estimates and degrades the deep-layer fit, and part of that error plausibly propagates upward into the overlying layers.

The Figure 22 confirms these tendencies. For VWC, most depths cluster tightly around the 1:1 line across the full observed range. The 10 cm layer reveals a systematic underestimation of peak wet states as observations reach $\approx 0.50 \text{ m}^3 \text{ m}^{-3}$ while the model rarely exceeds $0.35 \text{ m}^3 \text{ m}^{-3}$ at the same conditions, consistent with the post rainfall saturation near the surface being incompletely captured. For pF, the scatter follows the 1:1 line tightly across the full dynamic range ($\approx 1.5\text{--}4.0$ pF), with the characteristic hysteretic loops visible as closed curves for each depth, indicating that the model reproduces not only the mean seasonal trajectory but also the wetting–drying asymmetry.

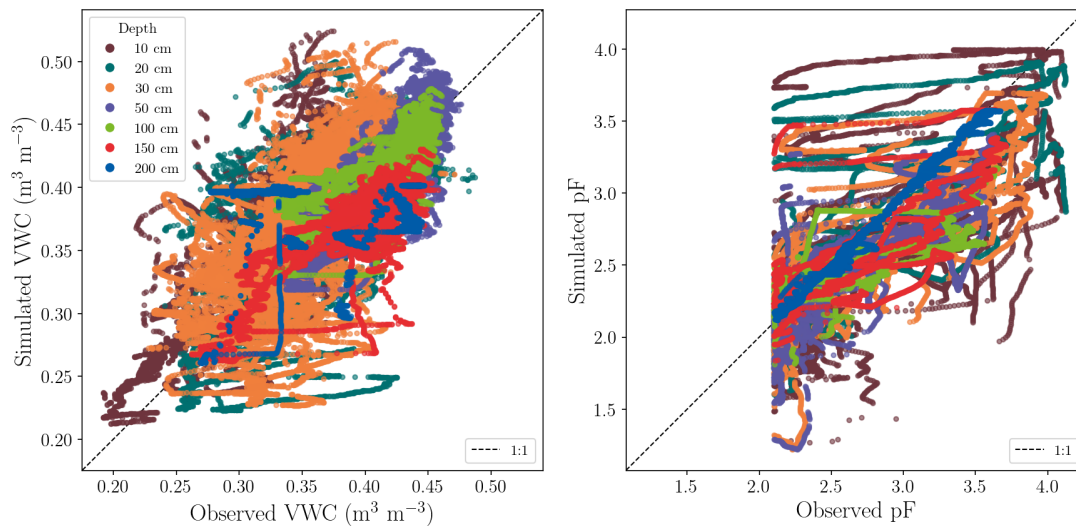


Figure 22: Simulated versus observed VWC (left) and matric potential in pF (right) for R40, pooling all observation depths over the full simulation period. Points are colored by depth. The dashed line is the 1:1 reference. Tight scatter around the 1:1 line is visible for most depths; the 10 cm layer shows underestimation of peak wet states, and the hysteretic loops in the pF panel confirm that the model captures the wetting–drying asymmetry.

Pasture

The headline scores are far weaker, with a pooled R^2 of only 0.11 for VWC and 0.16 for pF, yet these figures are misleading on their own. The individual depths between 10 and 150 cm score between 0.38 and 0.61 for VWC, comparable to R40. So the model does track the temporal dynamics at each depth. The collapse of the pooled metric therefore reflects a systematic between-depth offset rather than a failure to capture the dynamics. The predicted bands for the deeper layers sit consistently above the observations (Appendix H), and pooling all depths against a single mean turns that vertical bias into a large residual. Averaged over the six shallower depths, thus excluding the 200 cm depth, the per-depth VWC R^2 is 0.53, comparable to R40 and far above the pooled figure. This confirms that the 0.11 pooled value is dominated by the boundary depth and the between-depth offset, not by poor temporal tracking at any individual depth.

The MP record at Pasture is the weakest input of the whole study. The in situ measurements remain confined to a narrow band near the wet sensor limit, with almost no dynamic range over the period. The shallow pF scores are near zero (0.05 and 0.07 at 10 and 20 cm) not because the simulation is wrong but because there is hardly any observed variance left to explain. Since these nearly flat, sensor-limited series still enter the likelihood (Equation 28) with a small inferred $\sigma_{\text{pF}} = 0.21$ pF (Section VI.4), the calibration was driven to reproduce data that carry little information and are probably biased toward the wet end, distorting the inferred van Genuchten–Mualem parameters. The Pasture inversion is consequently constrained mainly by the VWC record, and the absolute matric potentials it returns should be treated as indicative.

The Figure 23 makes both issues immediately visible. In the VWC panel, the 100, 150 and 200 cm depths form distinct, near-horizontal clusters well above the 1:1 line, quantifying the systematic upward bias that dominates the pooled R^2 . The 200 cm layer is essentially a horizontal band, confirming that the model holds that depth at a near-constant water content regardless of what the probe records. The shallower layers scatter more symmetrically around the 1:1 line. In the pF panel, the observed axis barely exceeds a range of 1.0 pF units (≈ 2.0 – 3.3 pF), whereas the simulated axis spans a much wider range, making the near-zero shallow R^2 immediately understandable: the sensor data provide almost no signal for the model to fit against, yet the tight inferred σ_{pF} forced the calibration to treat them as reliable.

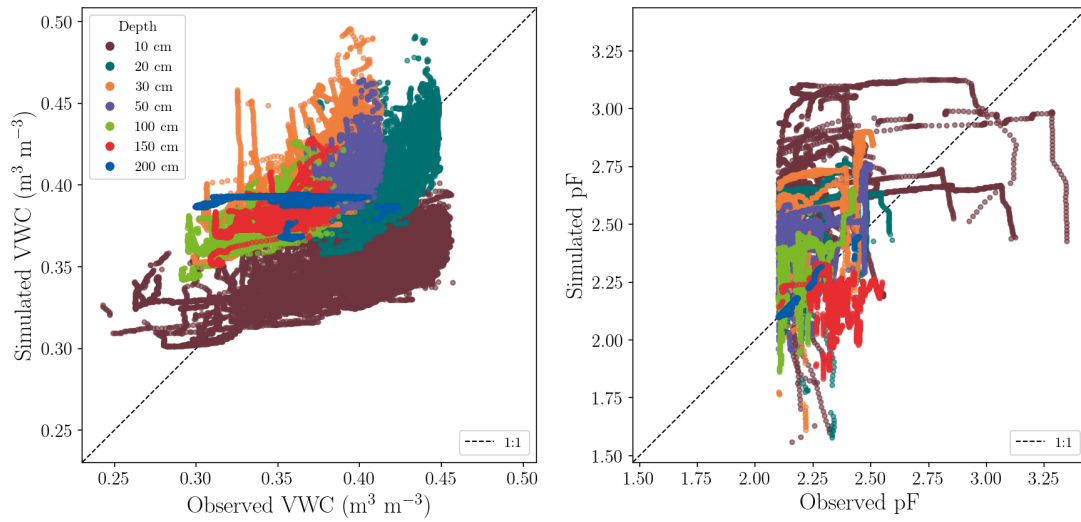


Figure 23: Simulated versus observed VWC (left) and matric potential in pF (right) for Pasture, pooling all observation depths over the calibration period. Points are colored by depth. The 100, 150 and 200 cm VWC clusters above the 1:1 line illustrate the systematic between-depth offset that dominates the pooled R^2 . The compressed observed pF range (≈ 2.0 – 3.3 pF) reflects the Teros 21 sensor saturation near the wet limit, rendering the shallow pF scores near zero by construction rather than by model failure.

Overall assessment

Despite these limitations, chiefly the low-quality prescribed lower boundary at both sites and the sensor-limited pF record at Pasture, the simulations reproduce the dominant features of the observed soil water behavior. The seasonal alternation of wetting and drying, the downward damping of amplitude with depth, and the mid-profile storage dynamics are all captured. At R40 they are captured out-of-sample at $R^2 \approx 0.5$ – 0.6 . The Pasture simulation, although penalized by a biased boundary and a poor pF record, still follows the core water-content dynamics of the profile and can reasonably be expected to represent its essential behavior once the systematic deep offset is set aside. Both models are therefore considered fit for their intended purpose, namely the quantification of the water-balance terms in the following sections, provided the deep-layer estimates and the absolute Pasture potentials are interpreted with the corresponding caution.

6 Simulated water balance of each plot

The posterior parameter ensembles validated in Section VI. were propagated through HYDRUS-1D to extract the complete water balance of each plot. This section describes each plot on its own terms, first flux by flux and then as a closed column balance, the plot-to-plot comparison being drawn in Section VI.7. To keep both plots strictly comparable, all fluxes reported here are evaluated over the common window of three complete hydrological years, 1 October 2021 to 30 September 2024 (1096 days), and are given as posterior-median rates [cm yr^{-1}]. The first hydrological year, however, falls within the Pasture spin-up and is therefore less directly comparable than the later two. The fluxes and storage compared here are obtained from a

single deterministic run driven by the posterior-median parameters, extended over the complete Pasture record. Calibration and posterior propagation were restricted to the first 1006 days for tractability, but this median run spans the full record, so the comparison window lies entirely within it. The defining values are collected in Table 14.

A note on sign and definition is needed before the terms are read. The infiltration reported here is the *net* flux across the soil surface, that is the water that effectively enters the profile after soil evaporation has been returned to the atmosphere across the same boundary. Soil evaporation is therefore already embedded in this net infiltration and is not a separate line of the balance, which closes as $\Delta S = \text{infiltration} - T_a + C_r - D$, where T_a is actual transpiration, C_r capillary rise and D drainage. The difference between gross precipitation and net infiltration corresponds to the water lost at or above the surface, namely canopy interception together with soil evaporation, before it can reach the rest of the profile.

6.1 Term-by-term analysis

Gross precipitation over the common window is identical at the two adjacent plots, 227.2 cm yr⁻¹. Of this input, the net infiltration into the soil is 175.2 cm yr⁻¹ at R40 and 180.6 cm yr⁻¹ at the Pasture. The remaining 52.0 cm yr⁻¹ at R40 and 46.6 cm yr⁻¹ at the Pasture is the combined loss to canopy interception and soil evaporation. These two surface losses are reached by contrasting routes at the two plots. At R40 the dense, high-LAI canopy intercepts a large fraction of gross rainfall but shades the ground, so its surface loss is dominated by interception. At the Pasture the sparse herbaceous cover intercepts little but exposes the soil to radiation, so its surface loss is dominated by soil evaporation. The near-equality of the two net infiltrations is thus not a coincidence but the outcome of interception at the forest being largely replaced by soil evaporation at the Pasture. Infiltration is strongly concentrated in the wet-season rainfall pulses and falls to near zero between them. Surface runoff is null over the whole window at both sites, the calibrated near-surface conductivities being high enough that infiltration capacity exceeds rainfall intensity on almost every event, as expected of a well-structured tropical soil column (Section III.1).

Actual transpiration averages 41.5 cm yr⁻¹ at R40 and 37.8 cm yr⁻¹ at the Pasture, the deeper-rooted, higher-LAI forest drawing slightly more water from the profile than the herbaceous cover. The period-mean ET_a/ET_p ratio is 0.942 at R40 against 0.996 at the Pasture, and the daily ratio (Figure 24) shows that R40 stress is episodic and intensifies through the record, the dry-season ratio falling from 0.944 in 2022 to 0.799 in the severe 2024 dry season, whereas the Pasture ratio stays at or above 0.987 in every dry season. This is the signature of a canopy sustaining an evaporative demand the soil cannot always meet, against a shallow cover whose lower demand the profile satisfies almost without restriction. Deep drainage at the base of the profile is the dominant output at both sites, averaging 137.1 cm yr⁻¹ at R40 and 161.5 cm yr⁻¹ at the Pasture, in both cases by far the largest loss of the balance ($\text{Drain}/P = 0.60$ and 0.71 respectively). In the cumulative water-flux curves (Figure 25), where inputs accumulate above zero and losses below, drainage grows almost monotonically through every wet sea-

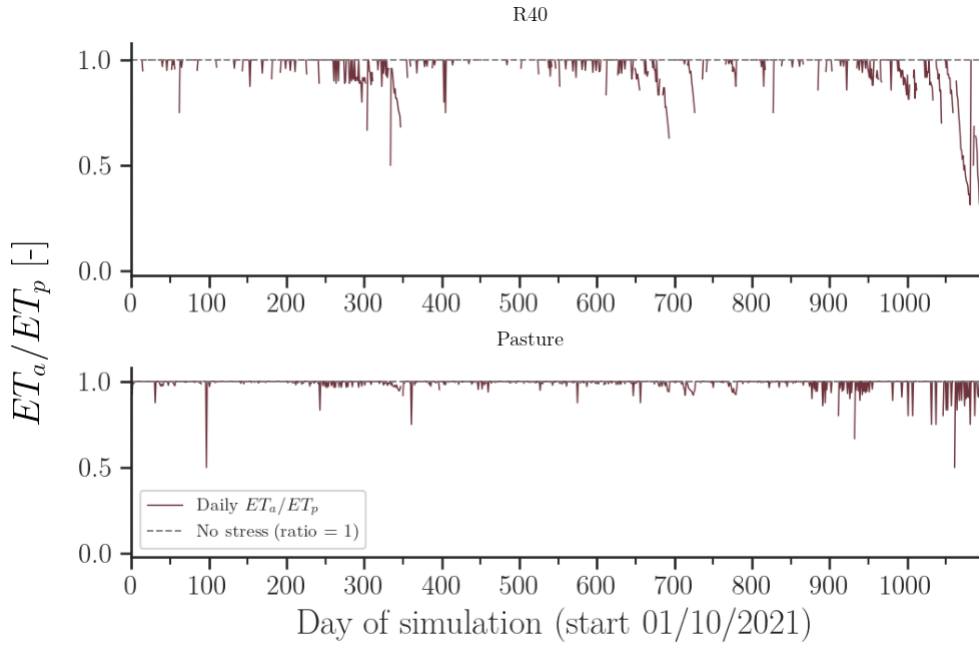


Figure 24: Daily ET_a/ET_p ratio for R40 (top) and Pasture (bottom) over the common window (start 1 October 2021). The dashed line marks the unstressed condition (ratio = 1). The ratio falls to 0.3 to 0.5 during the late-record dry-season drawdowns at R40, most markedly in 2024, while it rarely drops below 0.9 at the Pasture.

son and stalls through the dry season, mirroring the infiltration pulses with a short lag. That drainage dominates the losses is consistent with the humid regime of the site, where wet-season rainfall greatly exceeds the storage capacity and evaporative demand of the profile, so most infiltrated water percolates below 200 cm. Upward capillary flux across the lower boundary is the smallest term of the balance and shows the widest plot-to-plot contrast. It averages only 1.8 cm yr^{-1} at R40 but 18.9 cm yr^{-1} at the Pasture, so that relative to the transpiration it sustains it is negligible at R40 ($Cap/T_a = 0.04$) and substantial at the Pasture ($Cap/T_a = 0.50$). In the cumulative water-flux curves (Figure 25) it forms the thin input band that grows almost entirely during the dry seasons, when the depleted root zone establishes an upward gradient toward the wetter base of the profile (Section III.1), and it is visibly thicker at the Pasture than at R40. Both estimates, and the large Pasture value in particular, are strongly conditioned by the prescribed pressure-head lower boundary at 200 cm. Because the upward flux is set by this imposed head rather than by an interactive water table, a Pasture boundary held close to the wet sensor limit can sustain a large upward supply that may be an artifact of the boundary condition rather than a physical feature of the profile. It cannot, however, be excluded that the Pasture genuinely sits above a shallower water table, in which case part of this flux would be real. The two interpretations cannot be separated with the present setup, and the consequences are examined in Sections VI.5 and VIII.. The storage of the 0 cm to 200 cm profile (Figure 26) oscillates seasonally about a quasi-stationary mean, recharging in the wet season and depleting through the dry season. The oscillation is far wider at R40, whose storage swings between roughly 55 and 90 cm over the window, than at the Pasture, which remains in a narrow band near 75 cm. The wider R40 swing reflects both its stronger dry-season drawdown by transpiration and its larger wet-season recharge, the same dynamic

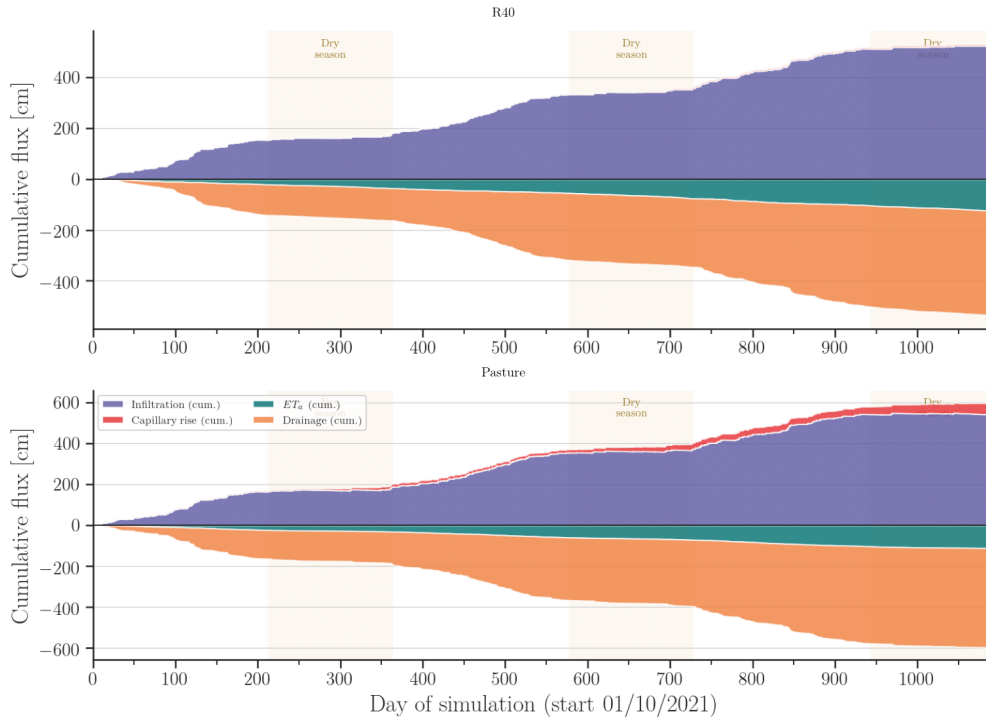


Figure 25: Cumulative water fluxes for R40 (top) and Pasture (bottom) over the common window (start 1 October 2021), inputs plotted above zero (net infiltration, capillary rise) and losses below (transpiration, drainage). Drainage dominates the losses at both plots, while the cumulative capillary rise is far larger at the Pasture.

range seen in the in-situ observations (Section VI.1.2), and it is most pronounced in the deep 2024 dry-season drawdown that brings the profile to its lowest state. The Pasture, by contrast, is markedly more stable.

6.2 Full-column balance

Assembling the terms over the common window closes the column balance of Equation (1). At R40 the net infiltration of 525.6 cm is removed by a transpiration of 124.4 cm and a drainage of 411.3 cm, and supplemented by a capillary rise of 5.5 cm, for a net storage change of -2.8 cm over the three hydrological years. At the Pasture the net infiltration of 541.9 cm is removed by a transpiration of 113.5 cm and a drainage of 484.4 cm, and supplemented by a capillary rise of 56.6 cm, for a net storage change of 4.4 cm. Soil evaporation, already contained in the net infiltration, accounts together with canopy interception for the difference between gross precipitation and net infiltration noted above and does not enter this closure as a separate term. The seasonal decomposition (Figure 27) shows the same structure at both plots, a large wet-season turnover of net infiltration inputs and drainage losses against a quiet dry season in which transpiration and capillary rise lead while infiltration and drainage nearly vanish. The seasonal ΔS alternates in sign accordingly, positive recharge in most wet seasons and negative draw down in most dry seasons, the net outcome being the small residual of two large opposing seasonal movements. This near-cancellation of large gross fluxes is why the plot-to-plot comparison of the next section is resolved into individual terms rather than read

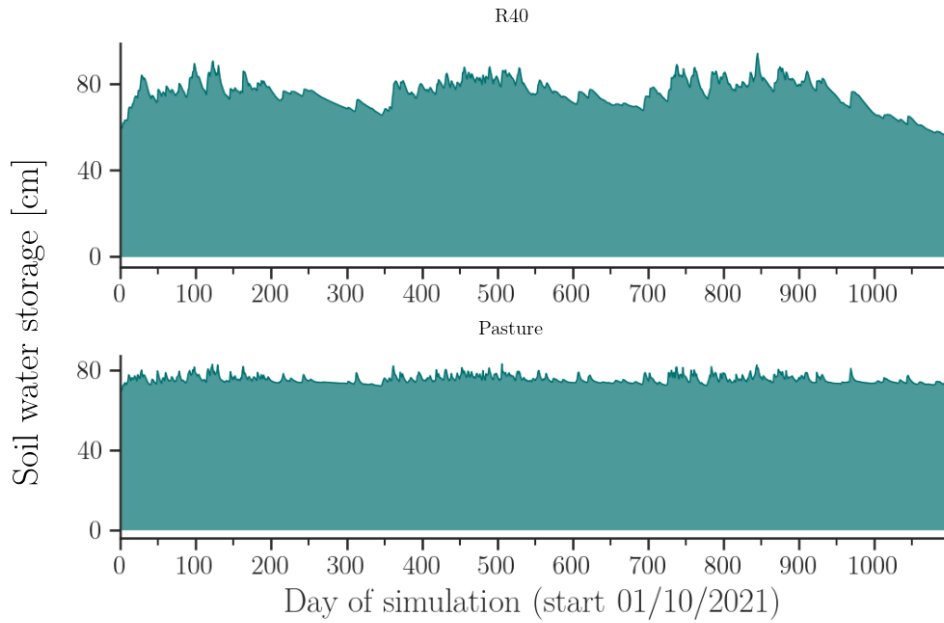


Figure 26: Soil water storage of the 0 cm to 200 cm profile for R40 (top) and Pasture (bottom) over the common window (start 1 October 2021). Both oscillate about a quasi-stationary seasonal mean with no sustained trend, the amplitude being markedly wider at R40.

from the totals.

7 Impact of reforestation on the soil water balance

This section answers the central question of the thesis by comparing the two plots over the common window of three complete hydrological years, 1 October 2021 to 30 September 2024. The hydrological year is taken to run from the onset of the wet season (1 October) to the end of the following dry season (30 September), following the seasonal definition of Section II.2, so that both endpoints share the same phase of the seasonal oscillation and the storage change is not contaminated by where the window is cut. The section isolates the sign and magnitude of the difference in net storage change and attributes it to the individual fluxes of Equation (1).

Over the three hydrological years the two land uses diverge in the sign of their net storage change. The Pasture *gains* water, with hydrological-year ΔS of +8.8, +0.9 and -5.2 cm for 2022, 2023 and 2024 respectively, summing to 4.4 cm (Figure 27). Over the same years R40 *loses* water, with hydrological-year ΔS of +22.0, -4.0 and -20.9 cm, summing to -2.8 cm. The reforested plot is therefore a net consumer of soil water relative to the pasture over the comparison window, the difference in three-year storage change amounting to about 7.2 cm in favor of the pasture. This difference is small in absolute terms, of the order of a single wet-season recharge event, and both plots remain close to a seasonally stationary regime. The result should therefore be read as a modest tendency rather than a large divergence. Under the soil and climate of São Francisco Xavier, the 40-year regenerated forest does not build up additional soil water relative to the adjacent pasture, and over the years sampled it draws its profile down slightly faster than it recharges it. The forest is also far more reactive

hydrologically, its hydrological-year ΔS ranging from +22.0 to -20.9 cm between a wet and a dry year while the Pasture stays within a few centimeters, a contrast relevant to the water-security framing of Section I..

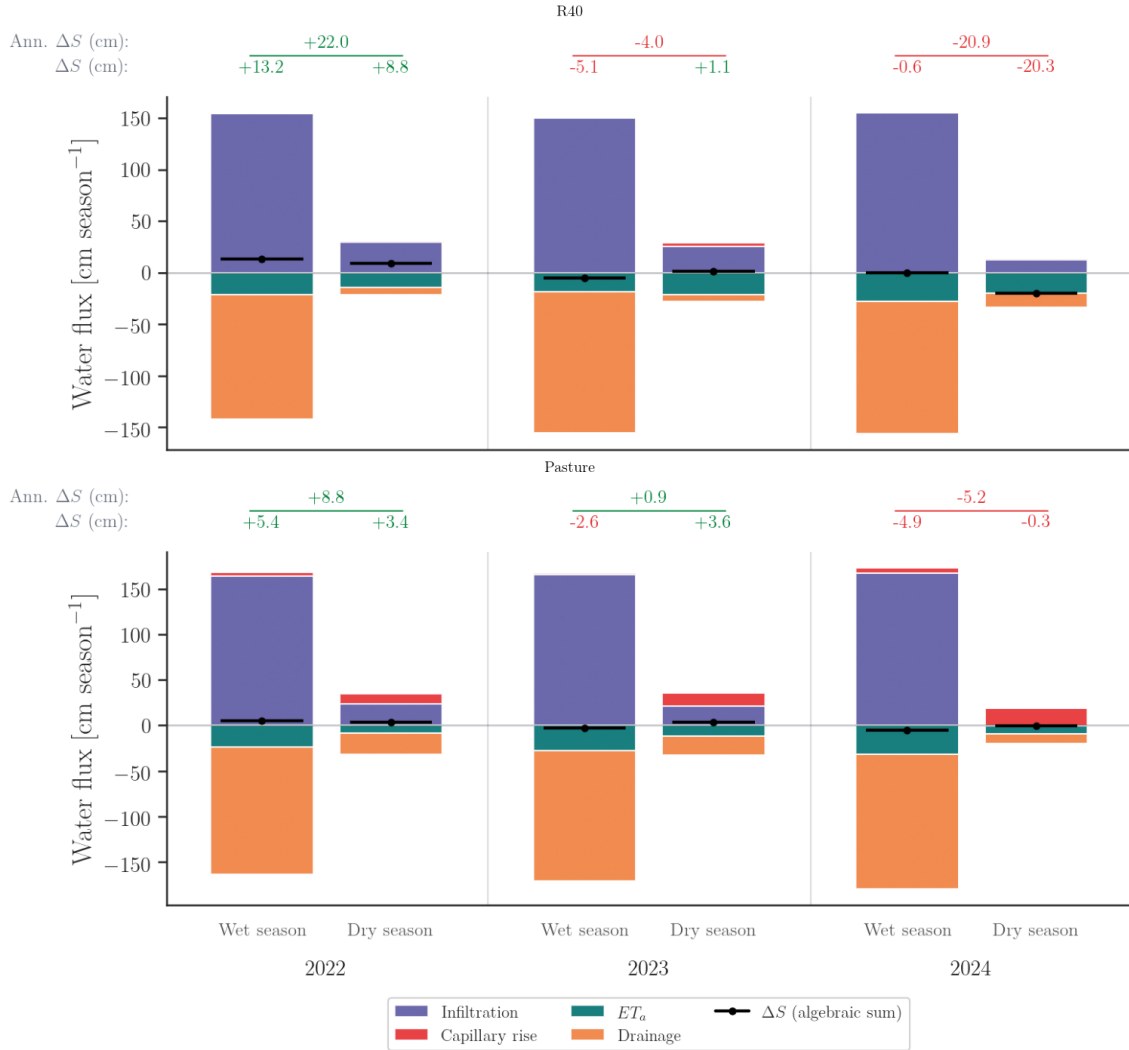


Figure 27: Seasonal soil water balance and net storage change (ΔS) for R40 (top) and Pasture (bottom), by wet and dry season over the three common hydrological years. Bars show the seasonal net infiltration, capillary rise (inputs) and transpiration, drainage (losses), the marker giving the resulting ΔS as their algebraic sum, with seasonal and annual totals above. Over the three years the Pasture gains 4.4 cm while R40 loses -2.8 cm.

7.1 Attribution to individual fluxes

The last column of Table 14 decomposes the 2.4 cm yr^{-1} difference in annual ΔS between the plots into the four balance terms. The decisive feature is that this small difference is the residual of much larger, partly offsetting contrasts. Three terms favor the Pasture, a larger capillary rise, a larger net infiltration and a lower transpiration, while one term works strongly against it, a larger drainage. These contributions nearly cancel, and the small surplus that remains is what tips the Pasture into net gain and R40 into net loss. No single flux dominates the outcome, and the largest single contrast, drainage, is itself a consequence of the larger total water input to the Pasture profile rather than an independent driver.

Table 14: Water-balance terms for R40 and Pasture over the common window 1 October 2021 to 30 September 2024 (three hydrological years), the fluxes as posterior-median annual rates [cm yr^{-1}] and the net storage change as a three-year total [cm]. Infiltration is the net surface flux, soil evaporation being already embedded in it; the interception-plus-evaporation line is the difference between gross precipitation and net infiltration, given for information. The last column is the Pasture-minus-R40 difference of each rate. A positive capillary rise is an input, a positive drainage a loss.

Term	R40	Pasture	Pasture–R40
Gross precipitation [cm yr^{-1}]	227.2	227.2	0.0
Net infiltration [cm yr^{-1}]	175.2	180.6	+5.4
Interception + soil evap. [cm yr^{-1}]	52.0	46.6	–5.4
Transpiration [cm yr^{-1}]	41.5	37.8	–3.7
Capillary rise [cm yr^{-1}]	1.8	18.9	+17.0
Drainage [cm yr^{-1}]	137.1	161.5	+24.4
Net storage change, 3-yr total [cm]	–2.8	+4.4	+7.2
ET_a/ET_p [–]	0.942	0.996	
Cap/T_a [–]	0.044	0.499	

7.2 Net storage change over the common window

The one difference that is both clearly attributable to land use and free of boundary-condition uncertainty is transpiration, higher at R40 (41.5 against 37.8 cm yr^{-1}). The forest sustains this demand into the dry season, when its ET_a/ET_p ratio falls as low as 0.80 while the Pasture stays near unity (Section VI.6). This is the larger root-zone withdrawal anticipated in Section V.5.5 from the higher LAI and deeper root system of the forest, and it acts toward a lower storage at R40, consistent with the trade-off hypothesis of Bruijnzeel (2004). Its magnitude, about 3.7 cm yr^{-1} , is modest relative to other flux contrasts.

The capillary contribution is the largest single term favoring the Pasture and at the same time the least trustworthy. It contradicts the working hypothesis of Section V.5.5, which anticipated a larger capillary rise at R40 from the stronger dry-season drawdown of the forest root zone. The simulation shows the opposite, with capillary rise an order of magnitude smaller at R40 (1.8 cm yr^{-1}) than at the Pasture (18.9 cm yr^{-1}). Because the model’s lower boundary is the prescribed pressure head at 200 cm rather than an interactive water table, this upward flux is dictated by that imposed head. The Pasture MP record stays close to the wet sensor limit ($\approx \text{pF } 2$) and never reaches the dry states of R40 (Section VI.), so the imposed Pasture boundary is held wetter and can sustain a large upward supply that is plausibly an artifact of the boundary condition. The alternative, that the Pasture genuinely overlies a shallower water table and that part of this capillary rise is real, cannot be ruled out, and a shallow water table is common in the humid tropics (Section III.1). The two explanations cannot be separated with the present setup.

Net infiltration is marginally higher at the Pasture, and drainage much higher. These two are

not independent, the larger input to the Pasture profile, augmented by its large capillary rise, must ultimately leave as deeper drainage, so most of the drainage contrast is the downstream consequence of the input contrast rather than a separate land-use effect. The near-equality of net infiltration despite very different canopy interception was explained in Section VI.6, interception at the forest being largely offset by soil evaporation at the Pasture, so infiltration carries little independent land-use signal. The near-surface saturated conductivity measured in the laboratory is in fact markedly higher at R40 than at the Pasture (Table 6), so the structural capacity to infiltrate is greater at the forest, in agreement with the recovery of macroporosity reported by Hassler et al. (2011) and Ilstedt et al. (2007). The calibration does not sharpen this contrast, the surface conductivity being poorly identified at both plots (Section VI.4). With the calibrated parameters the infiltration is supply-limited rather than capacity-limited, the surface conductivity exceeding rainfall intensity on almost every event at both plots, so the higher forest capacity does not translate into more infiltrated water.

7.3 Synthesis and answer to the research question

Confronting these results with the expectations of Section V.5.5, no single process dominates as cleanly as the infiltration-evapotranspiration trade-off of Bruijnzeel (2004) would suggest. The plots end the comparison window only 7.2 cm apart in three-year storage change, the residual of four much larger and largely offsetting fluxes. The only contrast that is both robustly attributable to the land cover and independent of the lower boundary is the higher forest transpiration, which pushes R40 toward depletion but is modest in size. The term that most tips the balance the other way, the Pasture's large capillary rise, is precisely the term least to be trusted, being governed by the prescribed pressure-head boundary rather than by an interactive water table. The infiltration advantage that might, following Krishnaswamy et al. (2013), have let the restored forest offset its higher water use is not expressed in the realized balance, because infiltration is supply-limited in this humid regime and the forest's higher interception is offset by the Pasture's higher soil evaporation.

To the general question, *what is the impact of reforestation on the soil water balance in the Atlantic Forest?*, the present modeling study answers, for this specific 40-year regenerated plot and its adjacent pasture, that reforestation produces only a marginal net drying of the profile relative to the pasture, about 7.2 cm of storage over three years, and that this small difference is the residual of large, nearly cancelling fluxes rather than the expression of a single dominant process. The forest also renders the profile hydrologically far more dynamic, with a much wider seasonal and interannual storage amplitude.

These conclusions are subject to the limitations established in Sections VI.5 and VIII., and one of them is central rather than peripheral here. The prescribed pressure-head lower boundary controls the capillary term directly, and that term is the single largest contributor keeping the Pasture in surplus. Whether through a boundary held artificially wet near the sensor limit or, if the Pasture truly overlies a shallow water table, through a genuine but unconstrained groundwater contribution, the capillary rise is the least physical term of the balance. Because

it is also the term that most determines the sign of the storage difference, the sign itself must be regarded as uncertain, and the conclusion that the pasture gains while the forest loses should be read as conditional on the lower boundary rather than as firmly established. The more robust statement the data support is the narrower one, that the forest transpires more and is hydrologically more dynamic, while the net storage difference is small and boundary-sensitive. The recovery of infiltration capacity documented for restored Atlantic Forest soils by Lozano-Baez et al. (2018, 2020) and Lozano-Baez et al. (2019) is consistent with the higher laboratory forest conductivity found here, but not for the calibrated one.

VII. Conclusion

This thesis set out to quantify the hydrological consequences of reforestation in the Atlantic Forest by asking, over a complete hydrological year, whether a 40-year regenerated forest gains or loses soil water relative to an adjacent degraded pasture, and through which fluxes any difference arises. To answer it, the one-dimensional soil water balance of both plots was simulated with HYDRUS-1D, solving the Richards equation under the van Genuchten–Mualem constitutive relations with Lenhard hysteresis, and the hydraulic parameters were constrained by Bayesian inverse estimation against the in-situ water-content and matric-potential records.

The simulation rests on a substantial body of field and laboratory data whose quality is uneven and ultimately conditions the confidence that can be placed in the results. The meteorological forcing, reconstructed by quantile mapping from a neighboring station to fill the 20.4 % of missing records, is reliable for the variables that dominate the energy budget, relative humidity, temperature and pressure all being reproduced with $R^2 > 0.97$, while precipitation, the single most important input to the water balance, is the least faithfully imputed in its event structure even though its totals remain close to the true amounts. The soil hydraulic characterization combined three independent laboratory devices into depth-grouped retention and conductivity curves that capture the broad behavior of the profile. The vegetation parameters, the fine-root profiles and the radiative-transfer LAI retrievals, are well resolved. Whereas the canopy storage parameter calibrated against the throughfall campaign returned an interception fraction of about 30 %, well above the 10 % to 20 % range expected for such a canopy, an overestimate traceable to period-level aggregation and to the systematic undercatch of the meteorological gauge. The weakest input by far is the MP record, whose sensors saturated near pF 2 during the wet season and so could not capture the wettest states the profiles likely reached. A deficiency whose consequences propagate through the entire modeling chain because the 200 cm pressure head is prescribed as the lower boundary condition.

The Bayesian inverse estimation proved both indispensable and partially limited. By inferring the full posterior distribution of the saturated conductivity K_s and the air-entry parameter α across the four material layers, rather than a single optimum, it propagated parameter uncertainty directly into the simulated fluxes and adapted the core-scale laboratory estimates to the effective field-scale behavior they cannot represent. The laboratory measurements that anchor the priors show a markedly higher near-surface saturated conductivity at R40 than at the Pasture, consistent with four decades of structural recovery of macroporosity under forest cover (Section III.3), The surface layers being only weakly identified. The calibration did not sharpen this contrast, but it confirmed, through the joint estimation of the measurement-error hyperparameters, that MP is reproduced less faithfully than water content. Its main limitation is structural rather than algorithmic as the deep conductivities are not identifiable. Clamped to the prescribed wet lower boundary, the likelihood rewards ever higher deep K_s with no interior optimum, so the chains climb until the bound truncates them and the apparently tight posteriors are bound artifacts rather than recovered properties. The deep estimates must therefore be read as bound-limited, but this affects the physical interpretation of the deep

layers more than the propagated balance, since those layers are weakly dynamic and their fluxes are governed by the imposed boundary in any case.

The validation supports the use of the calibrated models for their intended purpose, with an important asymmetry between the two plots. At R40, the parameters inferred from the 2019–2021 record reproduced the unseen 2022–2025 record with little loss of skill ($R^2 \approx 0.5$ – 0.6 for both water content and pF in calibration and validation alike), so the fit is not an artifact of overfitting and the model can be trusted to represent the forest profile out of sample. At the Pasture, where calibration and simulation periods necessarily coincided owing to the shorter record, the pooled R^2 collapses to about 0.11. Yet this figure is misleading as the per-depth scores between 10 and 150 cm are comparable to R40, and the low pooled value reflects a systematic between-depth offset and a near-flat, sensor-limited pF series rather than a failure to track the temporal dynamics. Both simulations capture the dominant features of the observed behavior, the seasonal alternation of wetting and drying, the downward damping of amplitude, and the mid-profile storage dynamics, and are therefore fit to quantify the water balance, provided the deep-layer estimates and the absolute Pasture potentials are treated with caution.

The central result follows from the term-by-term comparison of the two balances over three complete hydrological years (Table 14). Over this window the Pasture gains 4.4 cm of soil water while R40 loses 2.8 cm, a difference of about 7.2 cm in favor of the pasture. The 40-year regenerated forest is therefore a marginally larger net consumer of soil water than the adjacent pasture. The difference is small, of the order of a single wet-season recharge event, and both plots remain close to a seasonally stationary regime, so the result is best read as a modest but consistent tendency rather than a large divergence. Its attribution is unambiguous on one side and uncertain on the other. The forest transpires more, sustaining an evaporative demand into the dry season that the soil cannot always meet, and although its dense canopy intercepts a larger fraction of gross rainfall, in HYDRUS-1D this interception is removed from the evapotranspiration demand rather than from the throughfall, so it does not reduce the water entering the soil. The depletion is therefore driven by the higher transpiration itself rather than by reduced infiltration. This confirms the evapotranspiration side of the infiltration–evapotranspiration trade-off of Bruijnzeel (2004). The infiltration side, by contrast, leaves no imprint on the realized balance. Although the near-surface conductivity measured in the laboratory is markedly higher at R40, with the calibrated parameter, infiltration in this humid regime is supply-limited rather than capacity-limited, the surface conductivity exceeding rainfall intensity on almost every event. This is the principal contribution of the study, that under a humid highland rainfall regime the infiltration recovery widely documented for restored Atlantic Forest soils (Hassler et al., 2011; Ilstedt et al., 2007; Lozano-Baez et al., 2018; Lozano-Baez et al., 2019) does not translate into a soil-water gain. Unlike the Western Ghats catchments of Krishnaswamy et al. (2013), where recovered infiltration offset the higher forest water use, here it cannot, because rainfall is rarely limiting.

The one flux that contradicted the initial expectation is capillary rise, anticipated to be larger at the forest because of its stronger root-zone draw down but found to be an order of magnitude

smaller at R40 (1.8 against 18.9 cm yr⁻¹). This contrast is governed by the prescribed lower boundary rather than by root-zone dynamics, the Pasture boundary being held near the wet sensor limit and so able to sustain a large upward supply that may be a boundary artifact, although a genuinely shallower water table beneath the pasture cannot be excluded. The sign of the storage difference is robust to this uncertainty, resting on the well-constrained evapotranspiration contrast, but its precise magnitude is not. Finally, beyond the sign of the balance, the forest renders the profile hydrologically far more dynamic, its annual storage change ranging from 22.0 to -20.9 cm between a wet and a dry year while the Pasture stays within a few centimeters, a contrast in reactivity that is itself relevant to the regional water-security stakes that motivated this work.

In sum, for this specific 40-year regenerated plot and its adjacent pasture, reforestation in the highland Atlantic Forest does not build up soil water but slightly depletes it relative to pasture, a difference driven by higher evaporative demand rather than by any loss of infiltration, and accompanied by a markedly greater hydrological reactivity of the forested profile. These conclusions are necessarily conditioned by the limitations examined in the following section, the most consequential being the low-quality prescribed lower boundary, and they rest on a single pair of plots, so their extension to the wider biome must be made with the corresponding caution.

VIII. Limits and perspectives

1 Limitations

The results above are conditioned by a set of limitations that span the input data, the structural choices of the model, and the experimental design. They are grouped here from the most to the least consequential for the central question.

The single most consequential limitation is the prescribed pressure-head lower boundary at 200 cm. The MP sensors saturated near pF 2 during the wet season and could not record the wetter states the profiles likely reached (Section VI.), so the boundary that drives the model is itself of poor quality, especially at the Pasture. Three effects follow. First, the deep nodes are clamped to this near-constant head, which is the mechanism that drove the deep K_s estimates onto their upper bound and left them non-identifiable (Section VI.4). Second, the deep water-content fit degrades sharply because the simulated profile cannot follow the genuine deep variability recorded by the probes. Third, and most directly for the research question, the upward capillary flux is dictated by the imposed head rather than by an interactive water table, so the large Pasture capillary rise of 18.9 cm yr^{-1} cannot be cleanly separated into a boundary artifact and a possible genuine contribution from a shallow water table. The capillary term, and hence the precise size of the storage gap between the plots, must therefore be regarded as provisional.

The canopy interception of R40 carries its own uncertainty. The calibrated storage capacity returned an interception fraction near 30 %, above the literature range, for two compounding reasons. First the throughfall campaign was read at roughly two-week intervals, so calibration was performed at the period rather than the event level, which inflates the apparent canopy storage. Secondly the on-site pluviometers over-caught the meteorological station by a mean factor of about 1.5, pushing the fitted parameter further upward (Section 1.3.3.). Calibrating on dimensionless interception fractions limited but did not remove this bias. Stem flow was not measured and was omitted from the model, a small but systematic missing input in the net-rainfall partition. Although the scenario analysis showed the simulated soil water state to be almost insensitive to the interception parameter, because HYDRUS-1D subtracts the intercepted amount from the potential transpiration demand and the two opposing effects largely cancel in the drainage budget (Šimůnek et al., 2024), the absolute interception loss is still likely overestimated, which matters for any future use of these fluxes in isolation.

The LAI input was retrieved by radiative transfer modeling from Sentinel-2 and never validated against direct field measurements, so the absolute magnitudes, and through them the evaporation–transpiration partition and the interception storage, rest on the assumption that the retrieval is accurate for these canopies. The frequently cloudy wet season further thinned the usable retrievals precisely when the canopy is most dynamic, and the gap-filling assumes a smooth seasonal trajectory between observations that may miss short-term variation.

The one-dimensional column approach neglects all lateral exchange. At R40 the slope reaches

10 %, yet run-in is set to zero and surface runoff is permitted only above a ponding threshold. This is justified for a uniform slope by the tendency of overland flow to re-infiltrate down slope (Bonell, 1998), and surface runoff proved negligible in any case, but lateral subsurface flow on the forested slope is not represented and may matter during intense wet-season events.

Several forest-specific processes are present in the soil but cannot be isolated by the chosen model. Macroporosity and water repellency (Doerr et al., 2000; Lozano-Baez et al., 2020) are represented only implicitly, through the effective van Genuchten–Mualem parameters, so their individual contributions cannot be quantified. The microclimatic feedback of the forest canopy, in particular the cooler, more humid near-surface atmosphere under closed cover, was not accounted for. Both plots were forced with the same open-field meteorological series, which likely overstates the evaporative demand applied to the forest floor.

Two methodological limitations should also be noted. The hysteretic tied-parameter set caused widespread convergence failures, with only 41 % of the Morris runs completing at R40 and 10 % at the Pasture (Section VI.2). The surviving subset may not sample the parameter space uniformly, so the Pasture screening in particular is indicative rather than quantitative. The structural scenario analyses, moreover, necessarily preceded the calibration and so used the non-hysteretic laboratory parameters, had the optimized parameters been available, the relative ranking of the configurations might have differed.

Finally, the study is site-specific in the strongest sense. It compares a single forest plot with a single pasture plot, with no mature primary-forest reference against which to gauge how far recovery has progressed, and the soil, slope and highland climate of São Francisco Xavier differ from the lower, flatter Atlantic Forest sites where much of the comparable infiltration-recovery evidence was gathered (Section III.4). The conclusions should accordingly be read as a quantified result for this setting rather than a property of the biome.

2 Perspectives

The limitations above point to a clear order of priority for future work, and several of the highest-value improvements are also the cheapest to implement.

The most rewarding upgrade is to replace the prescribed-head lower boundary with a better-constrained one. Installing a wider-range matric-potential sensor or, better, a piezometer to monitor the actual water-table depth at both plots would remove the pF 2 saturation that currently caps the boundary, restore the identifiability of the deep conductivities, and allow the capillary term, the one flux that contradicted the working hypothesis, to be resolved as a physical quantity rather than a boundary artifact. This single change would directly sharpen the magnitude of the central result. A free-drainage lower boundary could be tested as a complementary configuration where no water-table influence is expected, isolating how much of the Pasture capillary supply is genuine.

On the vegetation side, the interception estimate would be improved most simply by an event-level throughfall campaign with co-located, tree-sheltered gauges cross-checked against the

meteorological station, together with explicit stem-flow collection. This would let the canopy-storage parameter be identified physically rather than as an effective value. Direct LAI measurements, by hemispherical photography or a canopy analyzer, would validate or correct the radiative-transfer retrieval on which the partitioning and interception currently depend. Adding a mature primary-forest reference plot would place R40 on the full deforestation–recovery trajectory and indicate how much of the structural and hydrological recovery remains incomplete.

Beyond these targeted fixes, this work also provides a reusable foundation for the different work that will follow it at this site. The gap-filled multi-year meteorological series, the depth-grouped laboratory hydraulic database, and above all the calibrated parameter posteriors and the tested HYDRUS-1D–Bayesian inverse pipeline could be carried forward directly, providing a validated baseline against which new configurations can be compared.

With this base in place, several avenues open naturally. The slope of R40 motivates extending the model to two or three dimensions to capture the lateral flow the column approach omits. The calibrated profiles can be driven by perturbed climate forcings to ask how the modest forest-versus-pasture storage difference would respond to a drier or more variable rainfall regime, a question with direct bearing on the water-security framing of this work. And the same framework could be applied to plots of differing regeneration age to turn the single-point comparison made here into a recovery chronosequence. In continuity with the successive studies at São Francisco Xavier (Delvoie, 2023; Dermauw, 2024; Oufini, 2025; Parmentier, 2024), the present contribution is thus both an answer, that reforestation might not yield a soil-water gain under a humid highland regime, and a tested instrument for the questions that remain.

References

- Allen, R., Pereira, L., Smith, M., & Raes, D. (1998, January). *Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56* (Vol. 56).
- Alvares, C. A., Stape, J., Sentelhas, P., Gonçalves, J., & Sparovek, G. (2013). Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, 22, 711–728. <https://doi.org/10.1127/0941-2948/2013/0507>
- Beven, K., & Germann, P. (1982). Macropores and Water Flow in Soils. *Water Resources Research - WATER RESOUR RES*, 18, 1311–1325. <https://doi.org/10.1029/WR018i005p01311>
- Bonell, M. (1998). Possible Impacts of Climate Variability and Change on Tropical Forest Hydrology. *Climatic Change*, 39, 215–272. <https://doi.org/10.1023/A:1005343820004>
- Braden, H. (1985). Ein Energiehaushalts- und Verdunstungsmodell für Wasser- und Stoffhaushaltsuntersuchungen landwirtschaftlich genutzter Einzugsgebiete. *Mitteilungen Deutsche Bodenkundliche Gesellschaft*, 42, 294–299.
- Brooks, R. H., & Corey, A. (1964). Hydraulic properties of porous media.
- Bruijnzeel, L. A. (2004). Hydrological functions of tropical forests: Not seeing the soil for the trees? *Agriculture, Ecosystems & Environment*, 104(1), 185–228. <https://doi.org/10.1016/j.agee.2004.01.015>
- Calmon, M., Brancalion, P., Paese, A., Aronson, J., Castro, P., Silva, S., & Rodrigues, R. (2011). Emerging Threats and Opportunities for Large-Scale Ecological Restoration in the Atlantic Forest of Brazil. *Restoration Ecology*, 19, 154–158. <https://doi.org/10.1111/j.1526-100X.2011.00772.x>
- Calvin, K., Dasgupta, D., Krinner, G., Mukherji, A., Thorne, P. W., Trisos, C., Romero, J., Aldunce, P., Barrett, K., Blanco, G., Cheung, W. W., Connors, S., Denton, F., Diongue-Niang, A., Dodman, D., Garschagen, M., Geden, O., Hayward, B., Jones, C., ... Ha, M. (2023, July). *IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]*. IPCC, Geneva, Switzerland. (P. Arias, M. Bustamante, I. Elgizouli, G. Flato, M. Howden, C. Méndez-Vallejo, J. J. Pereira, R. Pichs-Madruga, S. K. Rose, Y. Saheb, R. Sánchez Rodríguez, D. Ürges-Vorsatz, C. Xiao, N. Yassaa, J. Romero, J. Kim, E. F. Haites, Y. Jung, R. Stavins, ... C. Péan, Eds.; First, tech. rep.). Intergovernmental Panel on Climate Change (IPCC). <https://doi.org/10.59327/IPCC/AR6-9789291691647>
- Chazdon, R., Broadbent, E., Rozendaal, D., Bongers, F., Almeyda Zambrano, A., Aide, T. M., Balvanera, P., Becknell, J., Boukili, V., Brancalion, P., Craven, D., Almeida, J., Cabral, G., Jong, B., Denslow, J., Dent, D., DeWalt, S., Dupuy, J., Durán, S., & Poorter, L. (2016). Carbon sequestration potential of second-growth forest regeneration in the Latin American tropics. *Science Advances*, 2. <https://doi.org/10.1126/sciadv.1501639>
- Cuartas, L. A., Tomasella, J., Nobre, A. D., Hodnett, M. G., Waterloo, M. J., & Múnera, J. C. (2007). Interception water-partitioning dynamics for a pristine rainforest in Central

- Amazonia: Marked differences between normal and dry years. *Agricultural and Forest Meteorology*, 145(1), 69–83. <https://doi.org/10.1016/j.agrformet.2007.04.008>
- Delvoie, B. (2023). *Quel est l'impact de la reforestation sur l'équilibre des ressources en eau du sol ?* [Master's thesis, UCLouvain].
- Dermauw, C. (2024). *Comment le changement de couverture végétale, après une déforestation, affecte-t-il le carbone, la structure et les propriétés hydrauliques du sol ?* [Master's thesis, UCLouvain].
- Diaz, S., Pascual, U., Stenseke, M., Martín-López, B., Watson, R., Molnár, Z., Hill, R., Chan, K., Baste, I., Brauman, K., Polasky, S., Church, A., Lonsdale, M., Larigauderie, A., Leadley, P., van Oudenhoven, A., Plaat, F., Schröter, M., Lavorel, S., & Shirayama, Y. (2018). Assessing nature's contributions to people. *Science*, 359, 270–272. <https://doi.org/10.1126/science.aap8826>
- Doerr, S. H., Shakesby, R. A., & Walsh, R. P. D. (2000). Soil water repellency: Its causes, characteristics and hydro-geomorphological significance. *Earth-Science Reviews*, 51(1), 33–65. [https://doi.org/10.1016/S0012-8252\(00\)00011-8](https://doi.org/10.1016/S0012-8252(00)00011-8)
- Don, A., Schumacher, J., & Freibauer, A. (2011). Impact of tropical land use change on soil organic carbon stocks – A meta-analysis. *Global Change Biology*, 17, 1658–1670. <https://doi.org/10.1111/j.1365-2486.2010.02336.x>
- Durner, W. (1994). Hydraulic conductivity estimation for soils with heterogeneous pore structure. *Water Resources Research - WATER RESOUR RES*, 30, 211–223. <https://doi.org/10.1029/93WR02676>
- Feddes, R. A., Kowalik, P. J., & Zaradny, H. (1978). *Simulation of field water use and crop yield*. Centre for Agricultural Publishing and Documentation (PUDOC). OCLC: 476268904.
- Ferreira, M., Fernandes, B., & Curi, N. (1999). Influência da mineralogia da fração argila nas propriedades físicas de latossolos da região sudeste do Brasil. *Revista Brasileira de Ciência do Solo*, 23, 515–524. <https://doi.org/10.1590/S0100-06831999000300004>
- Filoso, S., Ometto Bezerra, M., Weiss, K., & Palmer, M. (2017). Impacts of forest restoration on water yield: A systematic review. *PLOS ONE*, 12, e0183210. <https://doi.org/10.1371/journal.pone.0183210>
- Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4), 457–472. <https://doi.org/10.1214/ss/1177011136>
- Germer, S., Neill, C., Krusche, A. V., & Elsenbeer, H. (2010). Influence of land-use change on near-surface hydrological processes: Undisturbed forest to pasture. *Journal of Hydrology*, 380(3), 473–480. <https://doi.org/10.1016/j.jhydrol.2009.11.022>
- Germer, S., Neill, C., Vetter, T., Chaves, J., Krusche, A., & Elsenbeer, H. (2009). Implications of long-term land-use change for the hydrology and solute budgets of small catchments in Amazonia. *Journal of Hydrology*, 364, 349–363. <https://doi.org/10.1016/j.jhydrol.2008.11.013>
- Grillakis, M., Polykretis, C., Manoudakis, S., Seiradakis, K., & Alexakis, D. (2020). A Quantile Mapping Method to Fill in Discontinued Daily Precipitation Time Series. *Water*, 12, 2304. <https://doi.org/10.3390/w12082304>

- Gudmundsson, L., Bremnes, J. B., Haugen, J., & Skaugen, T. (2012). Technical Note: Downscaling RCM Precipitation to the Station Scale Using Quantile Mapping – A Comparison of Methods. *Hydrology and Earth System Sciences Discussions*, 9, 6185–6201. <https://doi.org/10.5194/hessd-9-6185-2012>
- Haines, W. B. (1930). Studies in the physical properties of soil. V. The hysteresis effect in capillary properties, and the modes of moisture distribution associated therewith. *The Journal of Agricultural Science*, 20(1), 97–116. <https://doi.org/10.1017/S002185960008864X>
- Hassler, S. K., Zimmermann, B., van Breugel, M., Hall, J. S., & Elsenbeer, H. (2011). Recovery of saturated hydraulic conductivity under secondary succession on former pasture in the humid tropics. *Forest Ecology and Management*, 261(10), 1634–1642. <https://doi.org/10.1016/j.foreco.2010.06.031>
- Herman, J., & Usher, W. (2017). SALib: An open-source Python library for Sensitivity Analysis. *The Journal of Open Source Software*, 2, 97. <https://doi.org/10.21105/joss.00097>
- Homaee, M., Feddes, R. A., & Dirksen, C. (2002). A Macroscopic Water Extraction Model for Nonuniform Transient Salinity and Water Stress. *Soil Science Society of America Journal*, 66(6), 1764–1772. <https://doi.org/10.2136/sssaj2002.1764>
- Hosseini, R., Kumar, K., & Delenne, J.-Y. (2024). Investigating the source of hysteresis in the soil–water characteristic curve using the multiphase lattice Boltzmann method. *Acta Geotechnica*, 19, 7577–7601. <https://doi.org/10.1007/s11440-024-02295-y>
- Hudson, B. D. (1994). Soil organic matter and available water capacity. *Journal of Soil and Water Conservation*, 49(2), 189–194. <https://doi.org/10.1080/00224561.1994.12456850>
- Ilstedt, U., Malmer, A., Verbeeten, E., & Murdiyarso, D. (2007). The effect of afforestation on water infiltration in the tropics: A systematic review and meta-analysis. *Forest Ecology and Management*, 251(1), 45–51. <https://doi.org/10.1016/j.foreco.2007.06.014>
- Instituto Nacional de Meteorologia - INMET. (n.d.). Retrieved June 1, 2026, from <https://portal.inmet.gov.br/normais>
- IPCI-Canada. (2010). *About Pernambuco IPCI - Canada*. Retrieved May 31, 2026, from <https://www.ipci-canada.org/pernambuco.html>
- Jackson, R. B., Canadell, J., Ehleringer, J. R., & Mooney, H. A. (1996). A global analysis of root distributions for terrestrial biomes. *Oecologia*. <https://doi.org/10.1007/BF00333714>
- 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(3), 523–546. <https://doi.org/10.1111/j.1365-2389.2007.00915.x>
- Jobbágy, E., & Jackson, R. (2000). The Vertical Distribution of Soil Organic Carbon and Its Relation to Climate and Vegetation. *ECOLOGICAL APPLICATIONS*, 10, 423–436. <https://doi.org/10.2307/2641104>
- Joly, C., Metzger, J., & Tabarelli, M. (2014). Experiences from the Brazilian Atlantic Forest: Ecological findings and conservation initiatives. *New Phytologist*, 204. <https://doi.org/10.1111/nph.12989>

- Kool, J. B., & Parker, J. C. (1987). Development and evaluation of closed-form expressions for hysteretic soil hydraulic properties. *Water Resources Research*, 23(1), 105–114. <https://doi.org/10.1029/WR023i001p00105>
- Krishnaswamy, J., Bonell, M., Venkatesh, B., Purandara, B. K., Rakesh, K. N., Lele, S., Kiran, M. C., Reddy, V., & Badiger, S. (2013). The groundwater recharge response and hydrologic services of tropical humid forest ecosystems to use and reforestation: Support for the “infiltration-evapotranspiration trade-off hypothesis”. *Journal of Hydrology*, 498, 191–209. <https://doi.org/10.1016/j.jhydrol.2013.06.034>
- Lenhard, R. J., Parker, J. C., & Kaluarachchi, J. J. (1991). Comparing Simulated and Experimental Hysteretic Two-Phase Transient Fluid Flow Phenomena. *Water Resources Research*, 27(8), 2113–2124. <https://doi.org/10.1029/91WR01272>
- Likos, W., Lu, N., & Godt, J. (2014). Hysteresis and Uncertainty in Soil Water-Retention Curve Parameters. *Journal of Geotechnical and Geoenvironmental Engineering*, 140, 04013050. [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0001071](https://doi.org/10.1061/(ASCE)GT.1943-5606.0001071)
- Lozano-Baez, S. E., Cooper, M., Ferraz, S., Rodrigues, R., Castellini, M., & Di Prima, S. (2020, March). Assessing Water Infiltration and Soil Water Repellency in Brazilian Atlantic Forest Soils. <https://doi.org/10.5194/egusphere-egu2020-22580>
- Lozano-Baez, S. E., Cooper, M., Ferraz, S. F. B., Ribeiro Rodrigues, R., Pirastru, M., & Di Prima, S. (2018). Previous Land Use Affects the Recovery of Soil Hydraulic Properties after Forest Restoration. *Water*, 10(4), 453. <https://doi.org/10.3390/w10040453>
- Lozano-Baez, S. E., Cooper, M., Frosini de Barros Ferraz, S., Ribeiro Rodrigues, R., Castellini, M., & Di Prima, S. (2019). Recovery of Soil Hydraulic Properties for Assisted Passive and Active Restoration: Assessing Historical Land Use and Forest Structure. *Water*, 11(1), 86. <https://doi.org/10.3390/w11010086>
- Meireles, L., Shepherd, G., & Kinoshita, L. (2008). Variations in floristic and phytosociological structure of an upper montane forest in Mantiqueira Range, Monte Verde, MG. *Brazilian Journal of Botany*, 31, 559–574. <https://doi.org/10.1590/S0100-84042008000400003>
- METER group. (n.d.-a). *HYPROP - METER Group*. Retrieved May 26, 2026, from <https://metergroup.com/products/hyprop/>
- METER group. (n.d.-b). *KSAT - METER Group*. Retrieved May 29, 2026, from <https://metergroup.com/products/ksat/>
- METER group. (n.d.-c). *TEROS 10 - METER Group*. Retrieved May 24, 2026, from <https://metergroup.com/products/teros-10/>
- METER group. (n.d.-d). *TEROS 21 - METER Group*. Retrieved May 24, 2026, from <https://metergroup.com/products/teros-21/>
- METER group. (n.d.-e). *WP4C - METER Group*. Retrieved May 26, 2026, from <https://metergroup.com/products/wp4c/>
- Miguez-Macho, G., & Fan, Y. (2012). The role of groundwater in the Amazon water cycle: 2. Influence on seasonal soil moisture and evapotranspiration. *Journal of Geophysical Research: Atmospheres*, 117(D15). <https://doi.org/10.1029/2012JD017540>

- Morris, M. D. (1991). Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*, 33(2), 161–174. <https://doi.org/10.1080/00401706.1991.10484804>
- Mualem, Y. (1974). A conceptual model of hysteresis. *Water Resources Research*, 10(3), 514–520. <https://doi.org/10.1029/WR010i003p00514>
- Mualem, Y. (1976). A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, 12(3), 513–522. <https://doi.org/10.1029/WR012i003p00513>
- Myers, N., Mittermeier, R., Mittermeier, C., Fonseca, G., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403, 853–8. <https://doi.org/10.1038/35002501>
- Nepstad, D. C., de Carvalho, C. R., Davidson, E. A., Jipp, P. H., Lefebvre, P. A., Negreiros, G. H., da Silva, E. D., Stone, T. A., Trumbore, S. E., & Vieira, S. (1994). The role of deep roots in the hydrological and carbon cycles of Amazonian forests and pastures. *Nature*, 372(6507), 666–669. <https://doi.org/10.1038/372666a0>
- Neumann-Cosel, L., Zimmermann, B., Hall, J., van Breugel, M., & Elsenbeer, H. (2011). Soil carbon dynamics under young tropical secondary forests on former pastures—A case study from Panama. *Forest Ecology and Management*, 261, 1625–1633. <https://doi.org/10.1016/j.foreco.2010.07.023>
- Oliveira, R. S., Dawson, T. E., Burgess, S. S. O., & Nepstad, D. C. (2005). Hydraulic redistribution in three Amazonian trees. *Oecologia*, 145(3), 354–363. <https://doi.org/10.1007/s00442-005-0108-2>
- Oliveira-Filho, A., & Fontes, M. (2000). Patterns of floristic differentiation among Atlantic Forests in Southeastern Brazil and the influence of climate. *Biotropica*, 32, 793–810. <https://doi.org/10.1111/j.1744-7429.2000.tb00619.x>
- Oufini, M. (2025). *Dynamique des flux du bilan hydrique et propriétés hydrodynamiques d'une parcelle forestière tropicale* [Master's thesis, UCLouvain].
- Parmentier, L. (2024). *What is the impact of reforestation on evapotranspiration?* [Master's thesis, UCLouvain].
- Philip, J. R. (1964). Similarity hypothesis for capillary hysteresis in porous materials. *Journal of Geophysical Research (1896-1977)*, 69(8), 1553–1562. <https://doi.org/10.1029/JZ069i008p01553>
- Pinheiro-Dick, D., & Schwertmann, U. (1996). Microaggregates from Oxisols and Inceptisols: Dispersion through selective dissolutions and physicochemical treatments. *Geoderma*, 74(1), 49–63. [https://doi.org/10.1016/S0016-7061\(96\)00047-X](https://doi.org/10.1016/S0016-7061(96)00047-X)
- Poorter, L., Bongers, F., Aide, T. M., Almeyda Zambrano, A., Balvanera, P., Becknell, J., Boukili, V., Brancalion, P., Broadbent, E., Chazdon, R., Craven, D., Almeida, J., Cabral, G., Jong, B., Denslow, J., Dent, D., DeWalt, S., Dupuy, J., Durán, S., & Rozendaal, D. (2016). Biomass resilience of Neotropical secondary forests. *Nature*, 530. <https://doi.org/10.1038/nature16512>
- Raes, D., & Deproost, P. (2003). Model to assess water movement from a shallow water table to the root zone. *Agricultural Water Management*, 62(2), 79–91. [https://doi.org/10.1016/S0378-3774\(03\)00094-5](https://doi.org/10.1016/S0378-3774(03)00094-5)

- Rawls, W. J., Pachepsky, Y. A., Ritchie, J. C., Sobecki, T. M., & Bloodworth, H. (2003). Effect of soil organic carbon on soil water retention. *Geoderma*, 116(1), 61–76. [https://doi.org/10.1016/S0016-7061\(03\)00094-6](https://doi.org/10.1016/S0016-7061(03)00094-6)
- Rezende, C., Scarano, F., Assad, E., Joly, C., Metzger, J., Strassburg, B., Tabarelli, M., Fonseca, G., & Mittermeier, R. (2018). From hotspot to hopespot: An opportunity for the Brazilian Atlantic Forest. *Perspectives in Ecology and Conservation*, 16. <https://doi.org/10.1016/j.pecon.2018.10.002>
- Ribeiro, M. C., Metzger, J. P., Martensen, A. C., Ponzoni, F. J., & Hirota, M. M. (2009). The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. *Biological Conservation*, 142(6), 1141–1153. <https://doi.org/10.1016/j.biocon.2009.02.021>
- Ribeiro Filho, J. C., Lemos Filho, L. C. D. A., Santos, J. C. N. D., Brasil, J. B., & Palácio, H. A. D. Q. (2019). Partitioning of rainfall into throughfall, stemflow, and interception loss in the Brazilian Northeastern Atlantic Forest. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 23(1), 21–26. <https://doi.org/10.1590/1807-1929/agriambi.v23n1p21-26>
- Richards, L. A. (1931). Capillary conduction of liquids through porous mediums. *Physics*, 1(5), 318–333. <https://doi.org/10.1063/1.1745010>
- Ritchie, J. (1972). A Model for Predicting Evaporation From a Low Crop With Incomplete Cover. *Water Resources Research - WATER RESOUR RES*, 8. <https://doi.org/10.1029/WR008i005p01204>
- Rodrigues, R., Lima, R., Gandolfi, S., & Nave, A. (2009). On the restoration of high diversity forests: 30 years of experience in the Brazilian Atlantic Forest. *Biological Conservation*, 1242–1251. <https://doi.org/10.1016/j.biocon.2008.12.008>
- Ross, J., & Moroz, I. C. (1996). MAPA GEOMORFOLÓGICO DO ESTADO DE SÃO PAULO. *Geography Department, University of Sao Paulo*, 41–58. <https://doi.org/10.7154/RDG.1996.0010.0004>
- Rossi, M. (2017, July). *Mapa Pedológico do Estado de São Paulo: Revisado e ampliado*.
- Salemi, L. F., Groppo, J. D., Trevisan, R., de Moraes, J. M., de Barros Ferraz, S. F., Villani, J. P., Duarte-Neto, P. J., & Martinelli, L. A. (2013). Land-use change in the Atlantic rainforest region: Consequences for the hydrology of small catchments. *Journal of Hydrology*, 499, 100–109. <https://doi.org/10.1016/j.jhydrol.2013.06.049>
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008, January). *Global Sensitivity Analysis. The Primer* (Vol. 304). <https://doi.org/10.1002/9780470725184.ch6>
- Schad, P., van Huyssteen, C., & Micheli, E. (2015, January). *World Reference Base for Soil Resources 2014, update 2015*.
- Scharnagl, B., Vrugt, J. A., Vereecken, H., & Herbst, M. (2011). Inverse modelling of in situ soil water dynamics: Investigating the effect of different prior distributions of the soil hydraulic parameters. *Hydrology and Earth System Sciences*, 15(10), 3043–3059. <https://doi.org/10.5194/hess-15-3043-2011>

- Simunek, J., Jirka, Van Genuchten, M., & Šejna, M. (1999). The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media. *3*, 1–240.
- Šimůnek, J., Brunetti, G., Jacques, D., Van Genuchten, M. T., & Šejna, M. (2024). Developments and applications of the HYDRUS computer software packages since 2016. *Vadose Zone Journal*, *23*(4), e20310. <https://doi.org/10.1002/vzj2.20310>
- Šimůnek, J., Kodešová, R., Gribb, M., & Van Genuchten, M. (1999). Estimating hysteresis in the soil water retention function from cone permeameter experiments. *Water Resources Research - WATER RESOUR RES*, *35*, 1329–1345. <https://doi.org/10.1029/1998WR900110>
- Soylu, M. E., Istanbuluoglu, E., Lenters, J. D., & Wang, T. (2011). Quantifying the impact of groundwater depth on evapotranspiration in a semi-arid grassland region. *Hydrology and Earth System Sciences*, *15*(3), 787–806. <https://doi.org/10.5194/hess-15-787-2011>
- ter Braak, C., & Vrugt, J. (2008). Differential Evolution Markov Chain with snooker updater and fewer chains. *Statistics and Computing*, *18*, 435–446. <https://doi.org/10.1007/s11222-008-9104-9>
- Tisdall, J. M., & Oades, J. M. (1982). Organic matter and water-stable aggregates in soils. *Journal of Soil Science*, *33*(2), 141–163. <https://doi.org/10.1111/j.1365-2389.1982.tb01755.x>
- Tolimir, M., Kresović, B., Životić, L., Dragović, S., Dragović, R., Sredojević, Z., & Gajić, B. (2020). The conversion of forestland into agricultural land without appropriate measures to conserve SOM leads to the degradation of physical and rheological soil properties. *Scientific Reports*, *10*(1), 13668. <https://doi.org/10.1038/s41598-020-70464-6>
- Topp, G. C. (1971). Soil Water Hysteresis in Silt Loam and Clay Loam Soils. *Water Resources Research*, *7*(4), 914–920. <https://doi.org/10.1029/WR007i004p00914>
- Van Genuchten, M. (1980). A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils1. *Soil Science Society of America Journal*, *44*. <https://doi.org/10.2136/sssaj1980.03615995004400050002x>
- van Dam, J. C., Huygen, J., Wesseling, J. G., Feddes, R. A., Kabat, P., van Walsum, P. E. V., Groenendijk, P., & van Diepen, C. A. (1997). Theory of SWAP version 2.0; simulation of water flow, solute transport and plant growth in the soil-water-atmosphere-plant environment.
- Vercambre, G., Doussan, C., Pagès, L. L., Habib, R., & Pierret, A. (2002). Influence of xylem development on axial hydraulic conductance within *Prunus* root systems. *Trees - Structure and Function*, *16*, 479.
- Vogel, T., & Cislerova, M. (1988). On the reliability of unsaturated hydraulic conductivity calculated from the moisture retention curve. *Transport in Porous Media*, *3*, 1–15. <https://doi.org/10.1007/BF00222683>
- Von Hoyningen-Hüne, J. (1983). Die Interzeption des Niederschlages in landwirtschaftlichen Pflanzenbeständen. *Die Interzeption des Niederschlages in landwirtschaftlichen Pflanzenbeständen*, (57), 1–53.

- Wesseling, J. G. (1991). Meerjarige simulatie van grondwaterstroming voor verschillende bodemprofielen, grondwatertrappen en gewassen met het model SWATRE.
- Zhang, L., Hu, Z., Fan, J., Zhou, D., & Tang, F. (2014). A meta-analysis of the canopy light extinction coefficient in terrestrial ecosystems. *Frontiers of Earth Science*, 8(4), 599–609. <https://doi.org/10.1007/s11707-014-0446-7>
- Zimmermann, B., Elsenbeer, H., & De Moraes, J. M. (2006). The influence of land-use changes on soil hydraulic properties: Implications for runoff generation. *Forest Ecology and Management*, 222(1), 29–38. <https://doi.org/10.1016/j.foreco.2005.10.070>

A Detailed Penman–Monteith equation for the hourly reference evapotranspiration

The hourly version of the Penman–Monteith formula adapted by the FAO is detailed below:

$$ET_0 = \frac{0.408 \Delta (R_n - G) + \gamma \frac{37}{T + 273} u_2 (e^\circ(T) - e_a)}{\Delta + \gamma (1 + 0.34 u_2)}$$

with ET_0 the reference evapotranspiration [mm h^{-1}], R_n the net radiation at the grass surface [$\text{MJ m}^{-2} \text{h}^{-1}$], G the soil heat flux density [$\text{MJ m}^{-2} \text{h}^{-1}$], Δ the slope of the saturation vapour pressure curve [$\text{kPa } ^\circ\text{C}^{-1}$], γ the psychrometric constant [$\text{kPa } ^\circ\text{C}^{-1}$], T the hourly mean air temperature [$^\circ\text{C}$], $e^\circ(T)$ the saturation vapour pressure at air temperature T [kPa], e_a the hourly mean actual vapour pressure [kPa], and u_2 the hourly mean wind speed at 2 m height [m s^{-1}].

where:

- The slope of the saturation vapour pressure curve (Δ) is calculated as follows:

$$\Delta = \frac{4098 e^\circ(T)}{(T + 237.3)^2}$$

- The saturation vapour pressure at air temperature ($e^\circ(T)$) is calculated as follows:

$$e^\circ(T) = 0.6108 \exp\left(\frac{17.27 T}{T + 237.3}\right)$$

- The hourly mean actual vapour pressure is calculated as follows:

$$e_a = e^\circ(T) \frac{HR}{100}$$

- The psychrometric constant (γ) is calculated as follows:

$$\gamma = 0.665 \cdot 10^{-3} P_{\text{mean}}$$

with P_{mean} : Mean measured pressure [kPa]

- The net radiation at the grass surface is calculated as follows:

$$R_n = R_{ns} - R_{nl}$$

where:

$$R_{ns} = (1 - a) R_s \quad \text{and} \quad R_{nl} = \sigma T^4 (0.34 - 0.14 \sqrt{e_a}) \left(1.35 \frac{R_s}{R_{so}} - 0.35 \right)$$

where:

$$R_{so} = (0.75 + 2 \cdot 10^{-5} z) R_a$$

where:

$$R_a = \frac{12 \times 60}{\pi} G_{so} d_r [(\omega_2 - \omega_1) \sin \varphi \sin \delta + \cos \varphi \cos \delta (\sin \omega_2 - \sin \omega_1)]$$

where:

$$d_r = 1 + 0.033 \cos \frac{2\pi}{365} J$$

and:

$$\delta = 0.409 \sin \left(\frac{2\pi}{365} J - 1.39 \right)$$

and:

$$\omega_1 = \omega - \frac{\pi t_1}{24} \quad \text{and} \quad \omega_2 = \omega - \frac{\pi t_2}{24}$$

where:

$$\omega = \frac{\pi}{12} [(t + 0.06667(L_z - L_m) + S_c) - 12]$$

where:

$$S_c = 0.1645 \sin 2b - 0.1255 \cos b - 0.025 \sin b$$

where:

$$b = \frac{2\pi (J - 81)}{364}$$

with R_{ns} the net shortwave solar radiation [$\text{MJ m}^{-2} \text{h}^{-1}$], R_{nl} the net longwave radiation [$\text{MJ m}^{-2} \text{h}^{-1}$], a the albedo of a reference grass surface ($= 0.23$) $[-]$, σ the Stefan–Boltzmann constant ($= 2.0429 \cdot 10^{-10}$) [$\text{MJ m}^{-2} \text{h}^{-1}$], R_s the measured solar radiation [$\text{MJ m}^{-2} \text{h}^{-1}$], R_{so} the calculated clear-sky radiation [$\text{MJ m}^{-2} \text{h}^{-1}$], z the altitude of the weather station ($= 1100$) [m], R_a the extraterrestrial radiation [$\text{MJ m}^{-2} \text{h}^{-1}$], G_{so} the solar constant ($= 0.082$) [$\text{MJ m}^{-2} \text{min}^{-1}$], d_r the inverse relative Earth–Sun distance [rad], J the day number of the year (between 1 (01/01) and 365 or 366 (31/12)) [days], δ the solar declination [rad], φ the latitude [rad], ω_2 the solar time angle at the end of the hour [rad], ω_1 the solar time angle at the beginning of the hour [rad], ω the solar time angle at the midpoint of the hour [rad], t the time at the midpoint of the hour [h], L_z the longitude of the center of the local time zone [$^\circ$], L_m the longitude of the measurement area [$^\circ$], and S_c the seasonal correction for solar time [h].

B Labros fit parameter's distribution

Table 15: Parameter's 2.5% and 97.5% distribution for R40 and Pasture. Obtained by fitting the laboratories measurements with Labros. Since K_s was obtained by another method, the prior information for it are from the measurements used for its geometric mean.

Depth group [cm]	R40				Pasture			
	10	20+30	50+100	150+200	10	20+30+50	100+150	200
θ_r [$\text{cm}^3 \text{cm}^{-3}$]	0.171	0.107	0.160	0.007	0.050	0.050	0.123	0.093
2.50%	0.086	0.074	0.040	0	0.0377	0	0.053	0
97.50%	0.155	0.140	0.179	0.075	0.0555	0.086	0.193	0.251
θ_s [$\text{cm}^3 \text{cm}^{-3}$]	0.646	0.625	0.591	0.525	0.427	0.547	0.500	0.501
2.50%	0.635	0.617	0.534	0.520	0.422	0.541	0.494	0.484
97.50%	0.664	0.633	0.548	0.530	0.432	0.552	0.505	0.519
α [cm^{-1}]	0.0721	0.0903	0.0869	0.0699	0.0459	0.0587	0.0349	0.0417
2.50%	0.0617	0.0751	0.0669	0.0580	0.0377	0.0472	0.0266	0.0193
97.50%	0.1061	0.1085	0.1128	0.0843	0.0555	0.0731	0.0458	0.0902
n [-]	1.387	1.230	1.153	1.131	1.103	1.123	1.18	1.167
2.50%	1.259	1.201	1.118	1.107	1.083	1.097	1.132	1.083
97.50%	1.378	1.263	1.199	1.160	1.128	1.155	1.246	1.334
K_s [cm d^{-1}]	1962	456	336	99	65	424	14	25
1st value	6677	320	715	38	79	250	12	25
2nd value	454	799	546	13	54	1345	8	/
3rd value	2493	442	217	126	/	649	15	/
4th value	/	506	156	653	/	148	23	/
5th value	/	164	327	93	/	/	/	/
6th value	/	722	/	253	/	/	/	/
l [-]	1.390	-0.450	-2.117	-5.000	2.359	0.766	-5.000	-5.000
2.50%	0.315	-1.161	-3.860	-5.000	0.144	-0.998	-5.000	-5.000
97.50%	2.023	0.260	-0.374	-3.861	4.533	2.534	-3.174	1.438

C Gap filled meteorological series at São Francisco Xavier for the six measured variables.

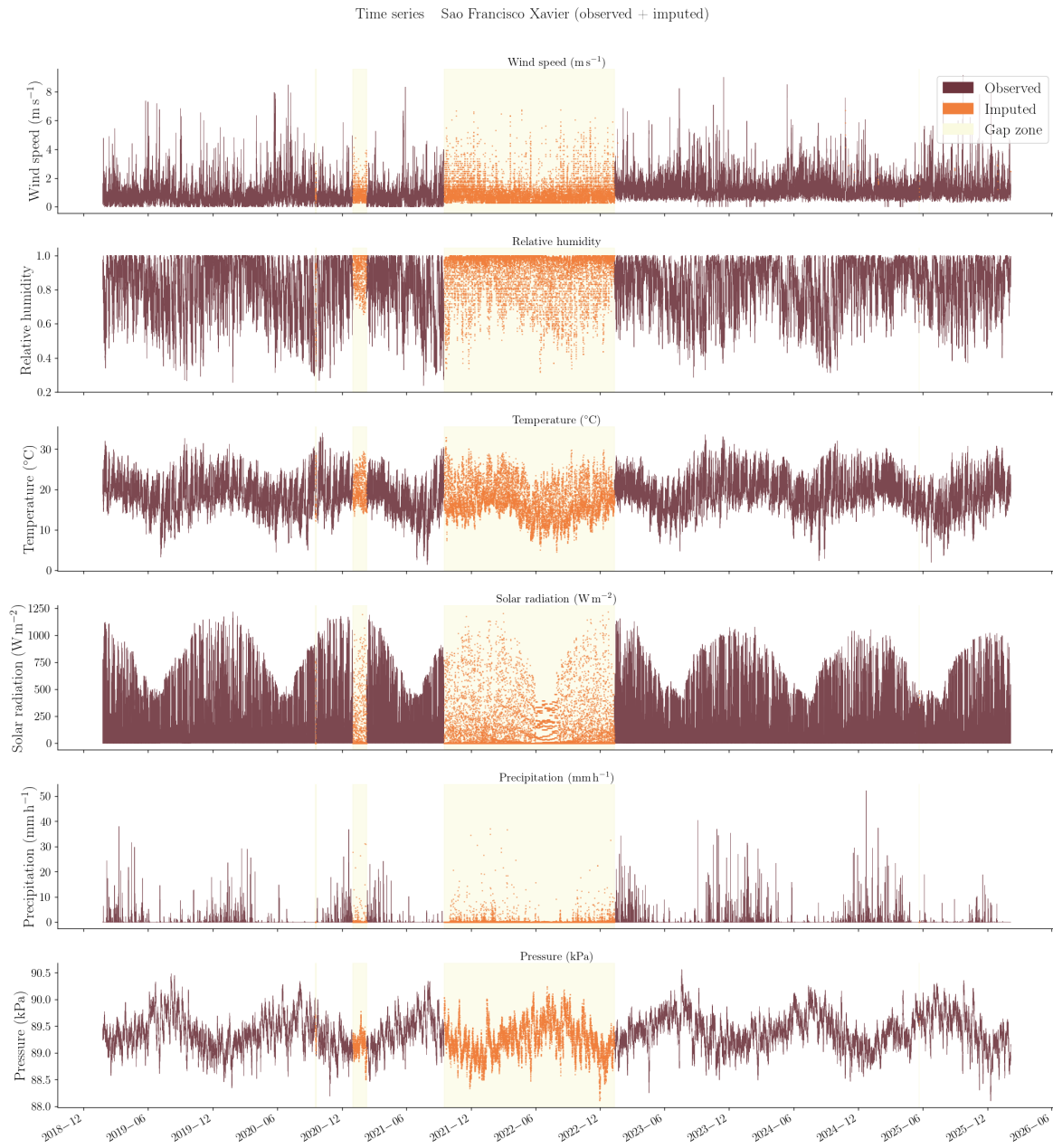


Figure 28: Gap filled meteorological series at São Francisco Xavier for the six measured variables. Observed INPE data, quantile-mapped imputed values and remaining gap hours are distinguished by color. The solar radiation different appearance from June to July 2022 is due to the residual filling of this period using the mean of the corresponding week for the whole evaluated period.

D Leaf area index monthly box plots for both sites

The box plots of monthly LAI (Figure 29) show the seasonal variation but it especially shows the lower number of observations for the wet period. It is caused by the cloudier conditions of that season, hence the variability for those months is higher. It also affects the quality of the linear interpolations for the time series, having longer gaps between observations.

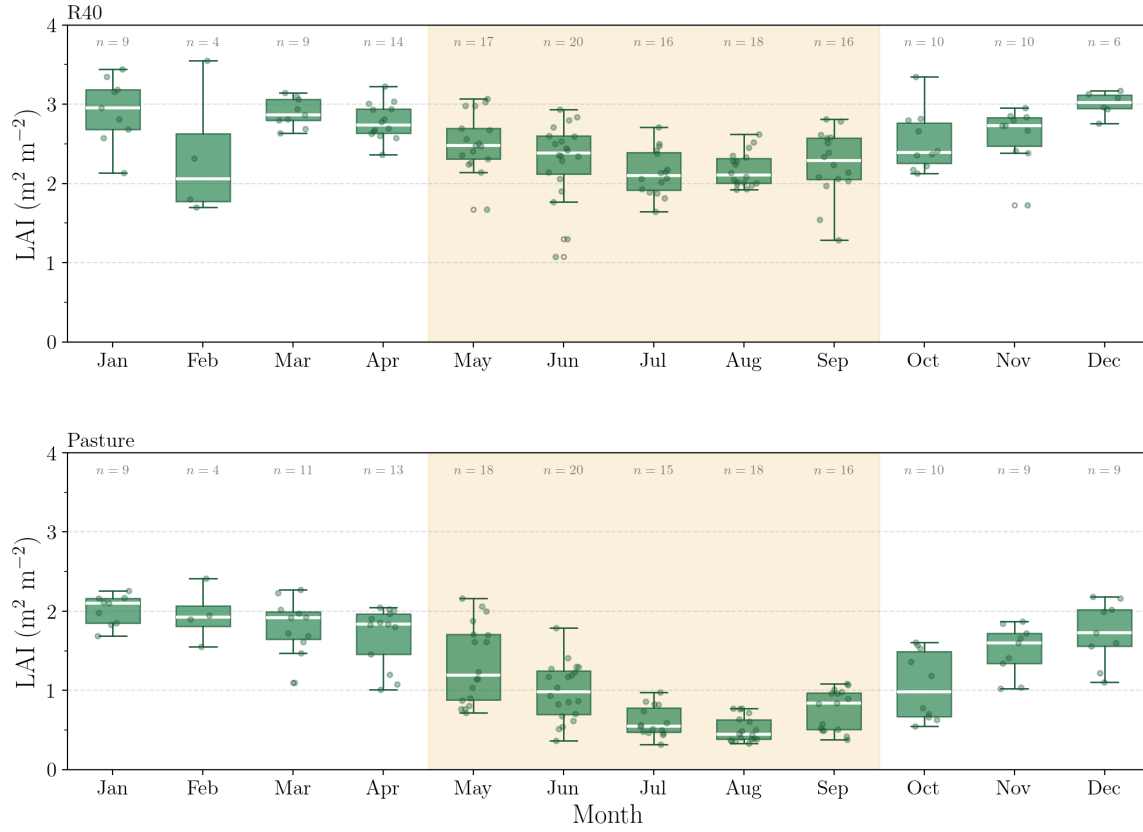


Figure 29: Box plots of the monthly leaf area index for R40 (top) and Pasture (bottom). n is the number of valid Sentinel-2 retrievals per month. The wet and dry theoretical seasons are also displayed on the graph. The dry season ranges from May to September while the wet season occupies the rest of the months.

E Canopy interception visualization for both sites

Figure 30 shows the behavior of Von Hoyningen-Hüne Braden model. At the beginning the canopy intercepts a maximum of rainfall according to bP and reach saturation later at a value of $a \cdot \text{LAI}$. The variation of the saturated values depends solely on the LAI at the time of the rainfall. The mean LAI differs from the observation mean since this one is the mean of the LAI time series, thus including all the interpolated values.

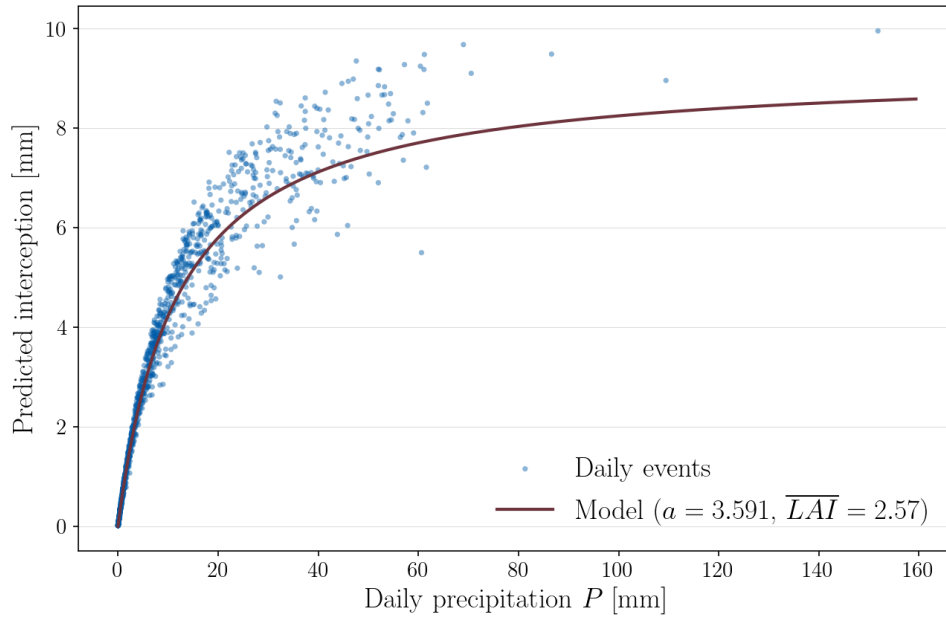


Figure 30: Daily interception against daily precipitation at R40, with the fitted Von Hoyningen-Hüne Braden model. Canopy storage parameter $a = 3.59 \text{ mm d}^{-1}$ and average leaf area index $\overline{\text{LAI}} = 2.57$.

Figure 31 and 32 represents the interception but for the Pasture site. The value of the canopy interception is $a = 0.25 \text{ mm d}^{-1}$ and thus produces way lower interception than R40.

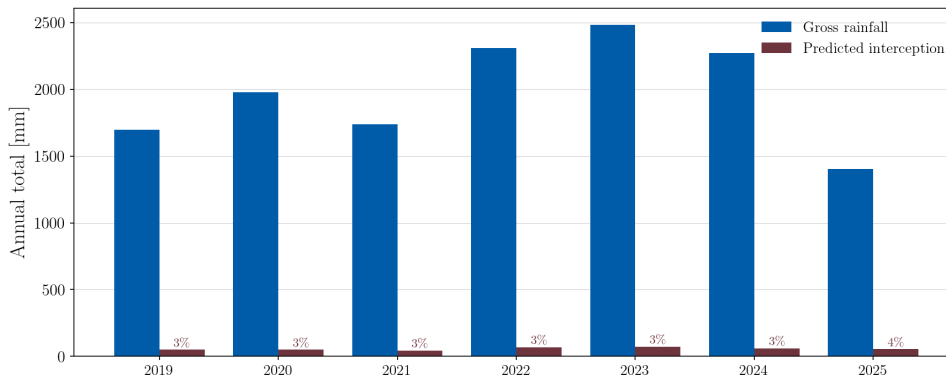


Figure 31: Gross rainfall and predicted canopy interception at Pasture, 2019–2025, canopy storage parameter $a = 0.25 \text{ mm d}^{-1}$. Labels give the annual interception fraction I/P .

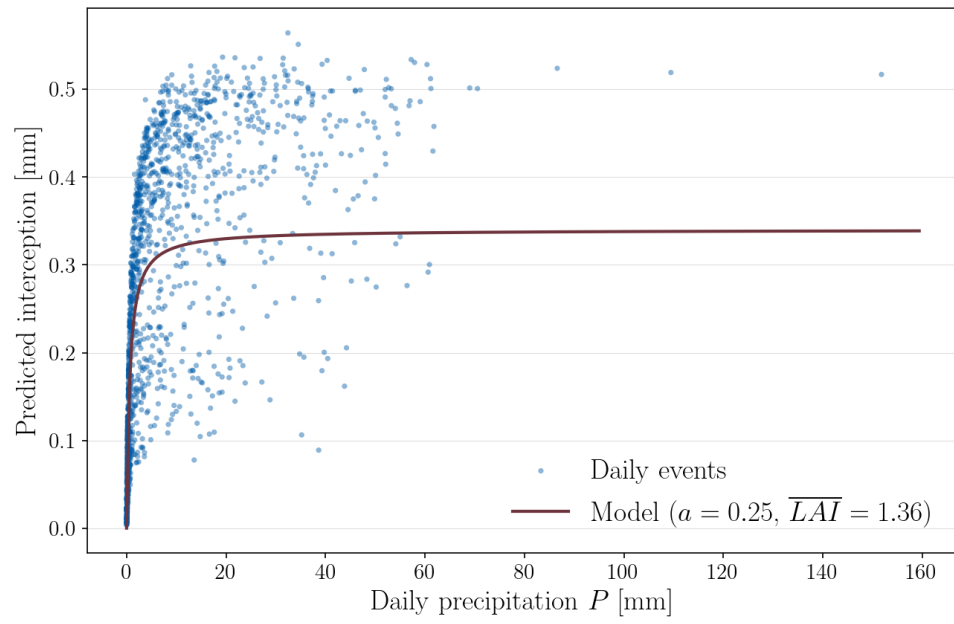


Figure 32: Daily interception against daily precipitation at Pasture, with the fitted Von Hoyningen-Hüne Braden model. Canopy storage parameter $a = 0.25 \text{ mm d}^{-1}$ and average leaf area index $\overline{LAI} = 1.36$.

F Morris sensitivity analysis heat maps for both sites

Figure 33 and 34 show the mean absolute elementary effect, μ^* , for every parameter across all outputs for R40 and Pasture respectively. Across outputs, μ^* values for each parameter seems consistent. This is less the case for Pasture, which could be the result of the less successful number of runs. For both sites, l and K_{sW} does not impact the simulations at all.

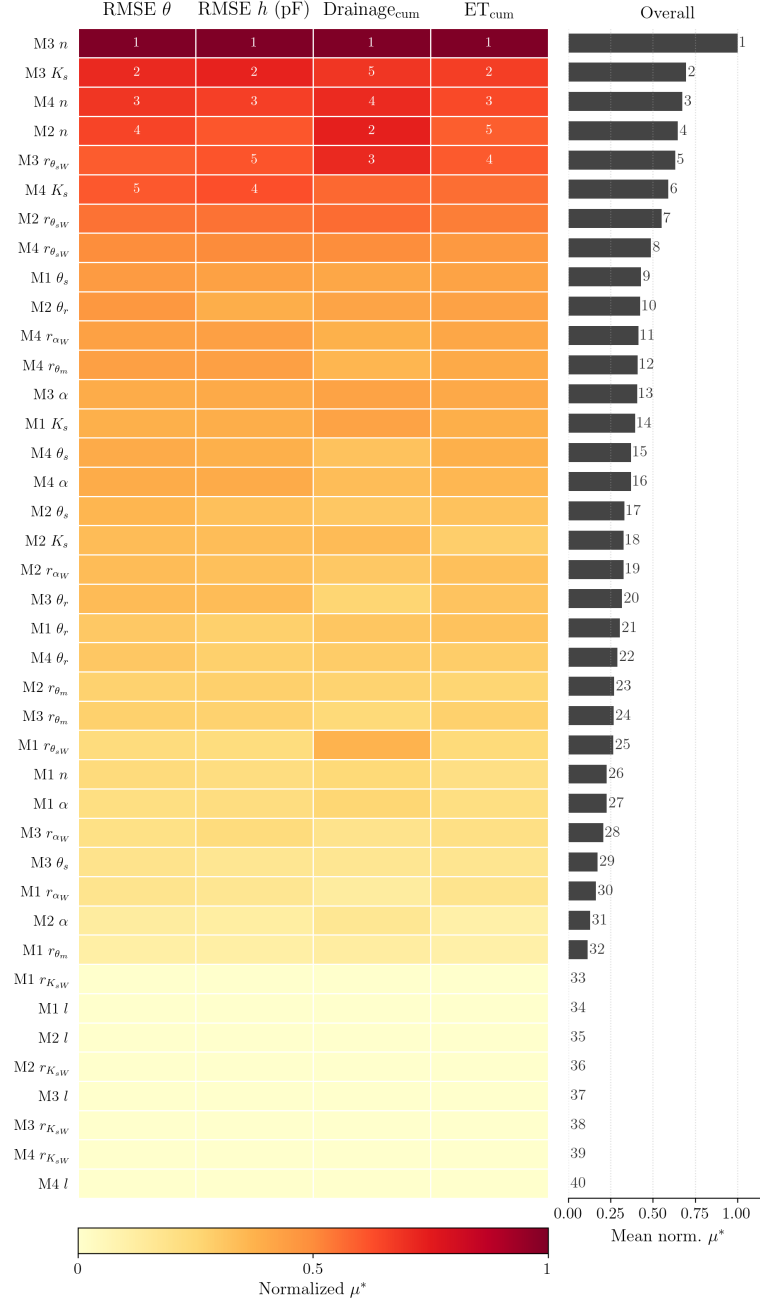


Figure 33: Heat map of the normalized μ^* of the Morris screening for every parameter across the four evaluated outputs (RMSE _{θ} , RMSE_{pF}, D_{cum} and ET_{cum}) for R40. The rank of the parameters for each output is based on μ^* values, the mean absolute elementary effect. The overall rank is based on the overall normalized μ^* , which is shown on the right of the figure.

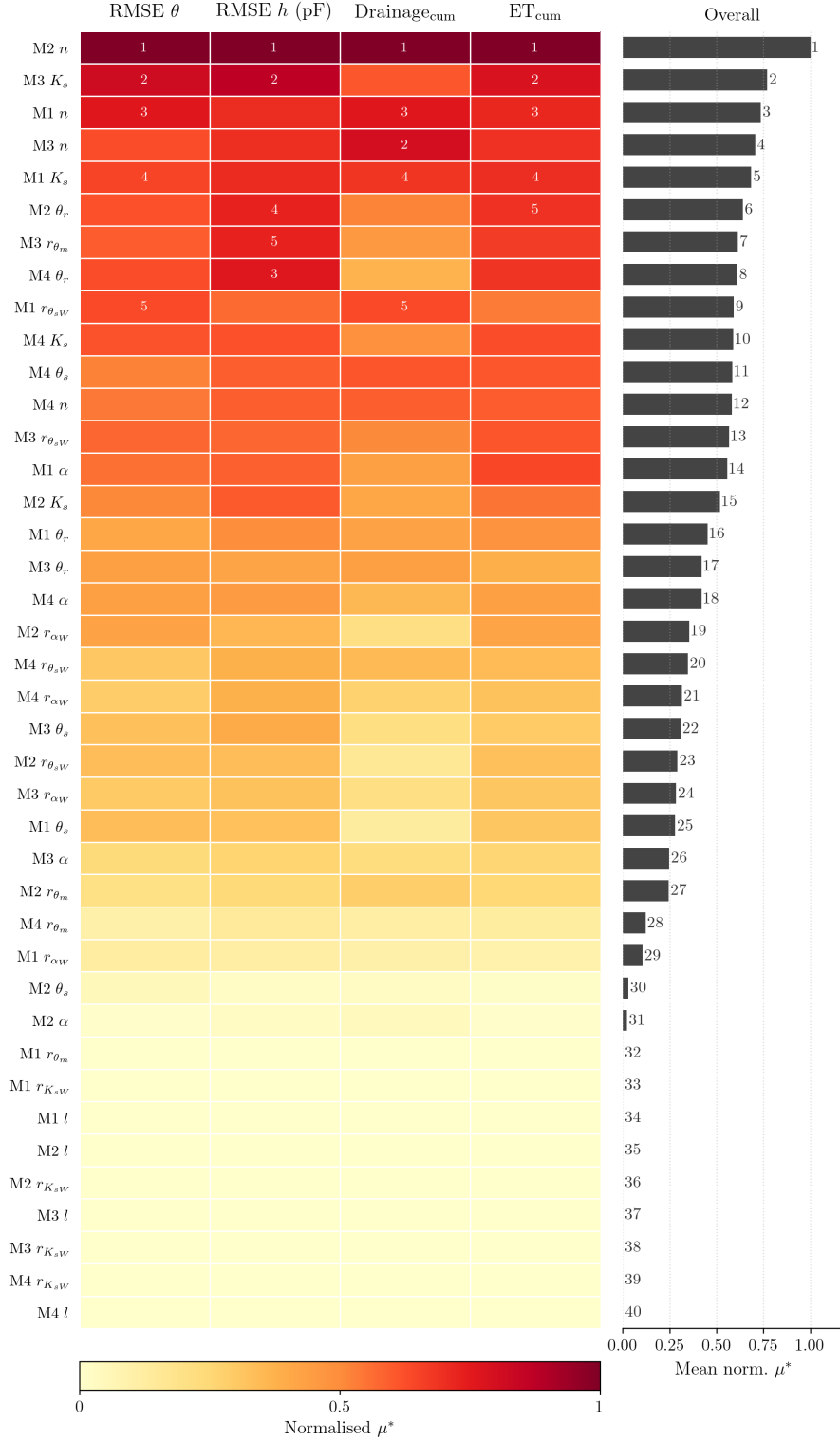


Figure 34: Heat map of the normalized μ^* of the Morris screening for every parameter across the four evaluated outputs (RMSE $_{\theta}$, RMSE $_{pF}$, D_{cum} and ET $_{cum}$) for Pasture. The rank of the parameters for each output is based on μ^* values, the mean absolute elementary effect. The overall rank is based on the overall normalized μ^* , which is shown on the right of the figure.

G Bayesian inference post-to-posterior distribution

The sampling traces of the five chains and the prior-to-posterior updating of every parameter are shown below for R40 (Figures 36 and 35) and for Pasture (Figures 37 and 38).

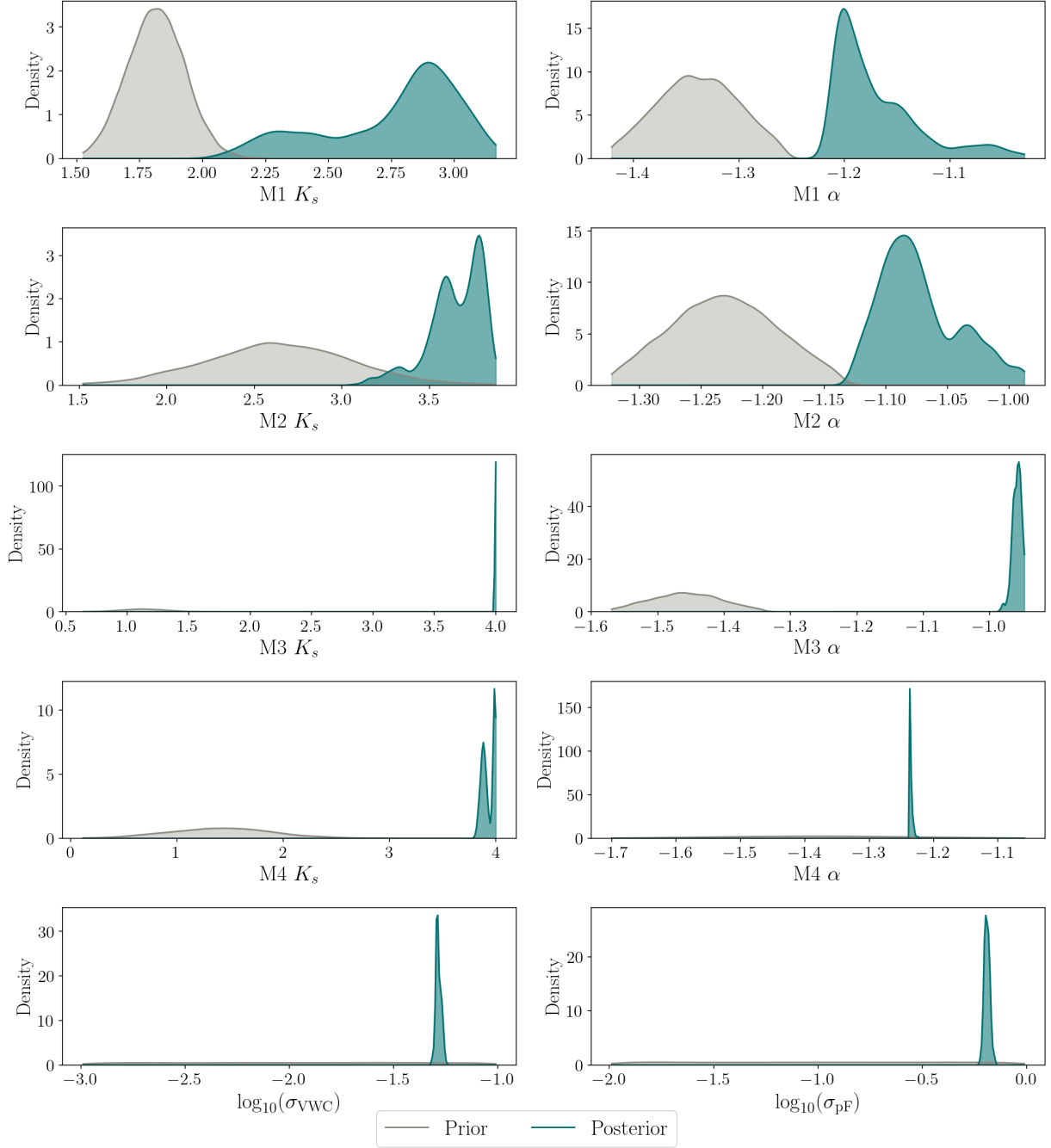


Figure 35: Prior (grey) and posterior (blue) marginal distributions for the ten calibrated parameters at R40. The contraction of the posterior relative to the prior measures the information gained from the in-situ observations. The deeper α parameters are sharply updated, while M1 K_s and M1 α remain close to, or broader than, their priors. The deep conductivities M3 K_s and M4 K_s collapse onto the upper bound rather than onto an interior value.

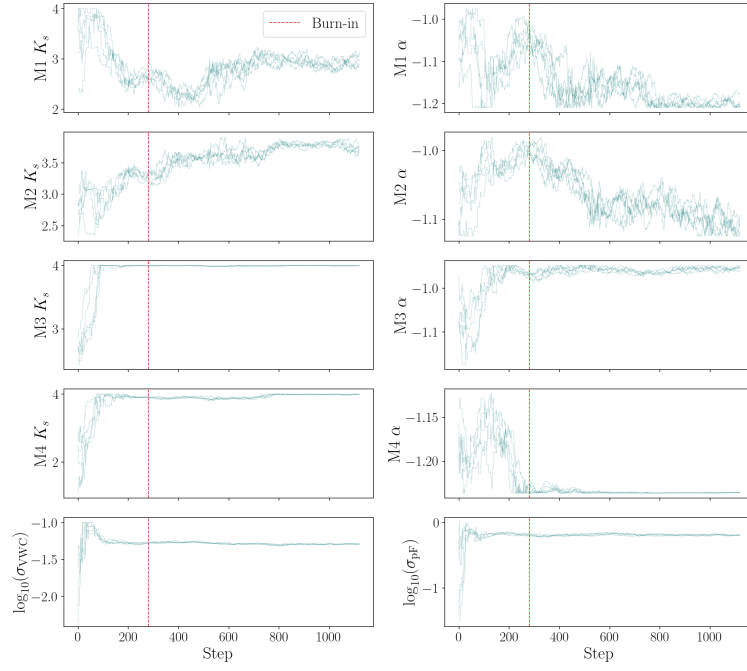


Figure 36: DREAM(ZS) sampling traces for the ten calibrated parameters at R40, five chains over 1120 steps. The red dashed line marks the end of the burn-in (step 280). After burn-in the chains fluctuate around stable levels for the well-identified parameters, while other parameters continue to wander, reflecting their weaker identifiability.

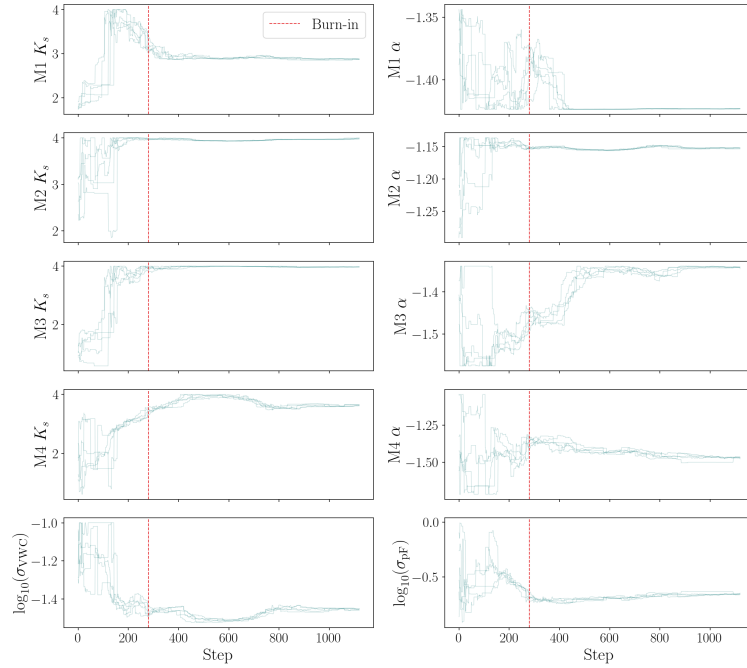


Figure 37: DREAM(ZS) sampling traces for the ten calibrated parameters at Pasture, five chains over 1120 steps. The red dashed line marks the end of the burn-in (step 280). The chains stabilize after burn-in, M1 K_s mixing more tightly than at R40, while the deeper K_s settle on or near the upper bound.

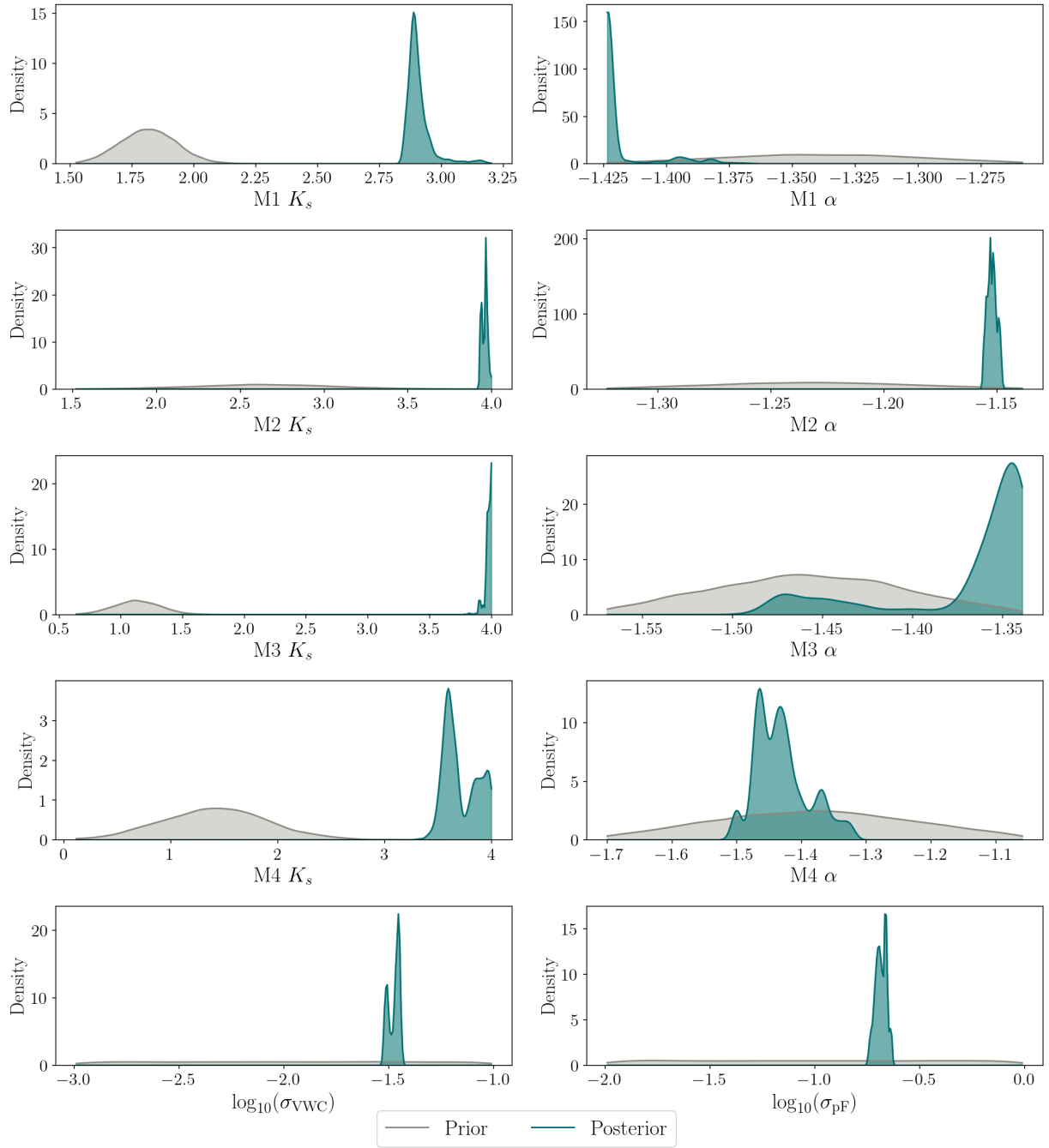


Figure 38: Prior (grey) and posterior (blue) marginal distributions for the ten calibrated parameters at Pasture. The surface conductivity M1 K_s shifts cleanly upward to a well-defined posterior, while the deeper M2–M4 K_s collapse onto or near the upper bound. The parameters M3 α and M4 α barely depart from their prior.

H Posterior predictive simulations of VWC and matric potential

The posterior-predictive simulations summarized in Section VI.5 are shown in full below for both sites. Each panel corresponds to one observation depth and overlays the in-situ observations, the posterior median and the 95 % predictive interval assembled from the 200 propagated draws. For R40 the observations are split into the calibration set (used by the sampler) and the validation set (the unseen record); for Pasture only the calibration set is available. The per-panel and overall R^2 values reported here are those collected in Table 13.

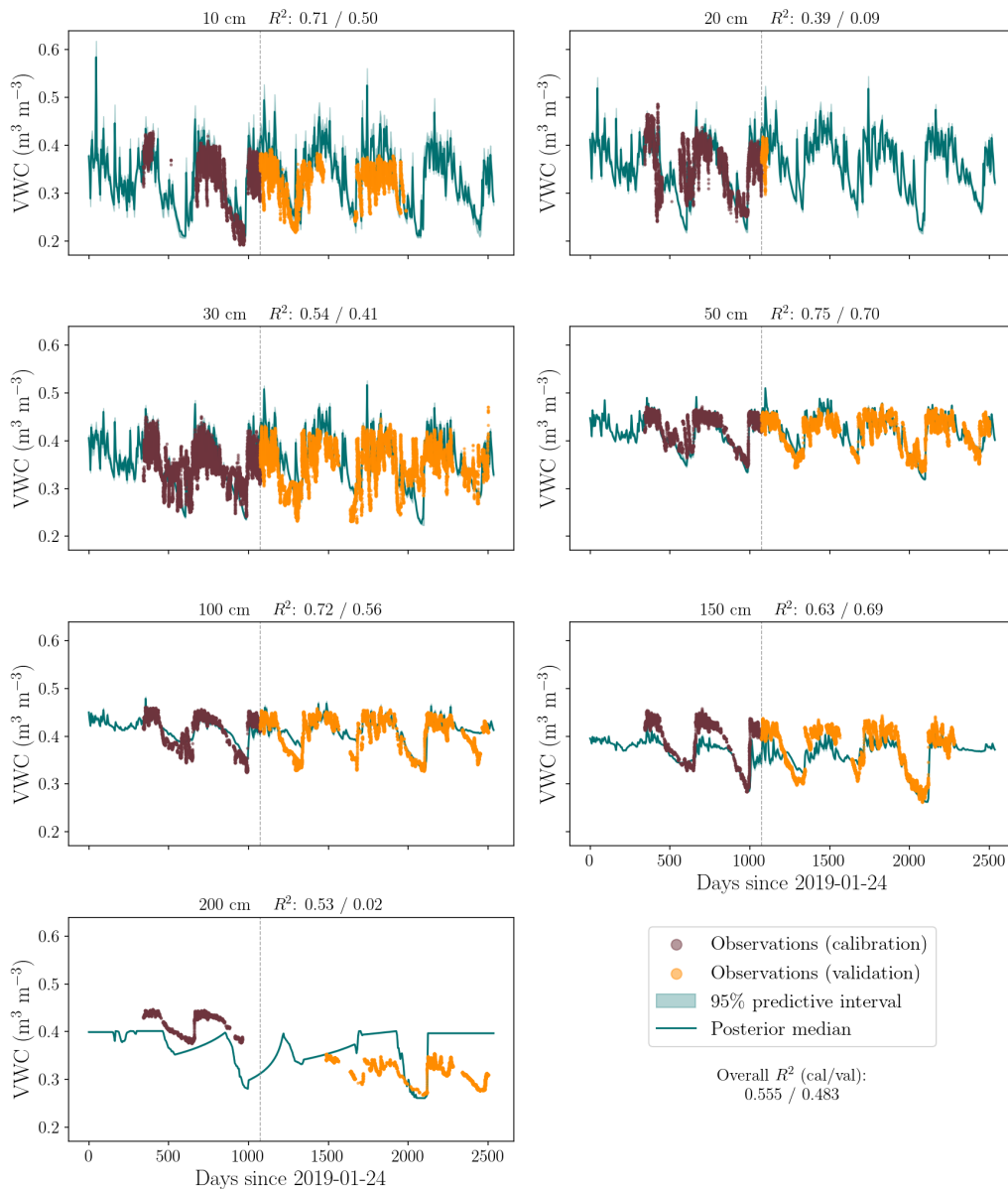


Figure 39: Posterior predictive simulation of volumetric water content at R40, by depth. Calibration and validation observations, the posterior median and the 95 % predictive interval are distinguished by color. Per-panel and overall R^2 (calibration / validation) are annotated.

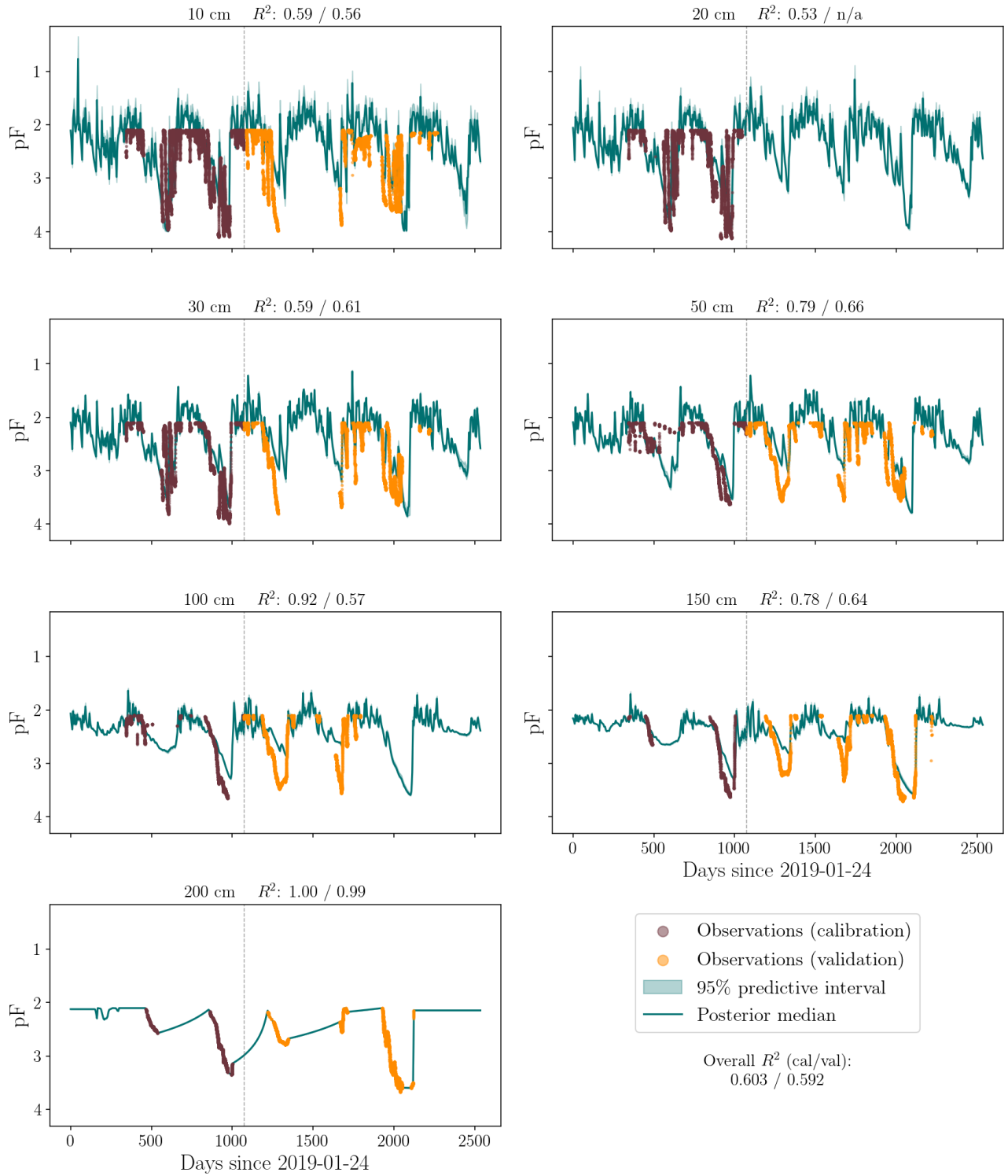


Figure 40: Posterior predictive simulation of matric potential (pF) at R40, by depth. The near-perfect score at 200 cm reflects the prescribed lower boundary condition rather than predictive skill. Per-panel and overall R^2 (calibration/validation) are annotated.

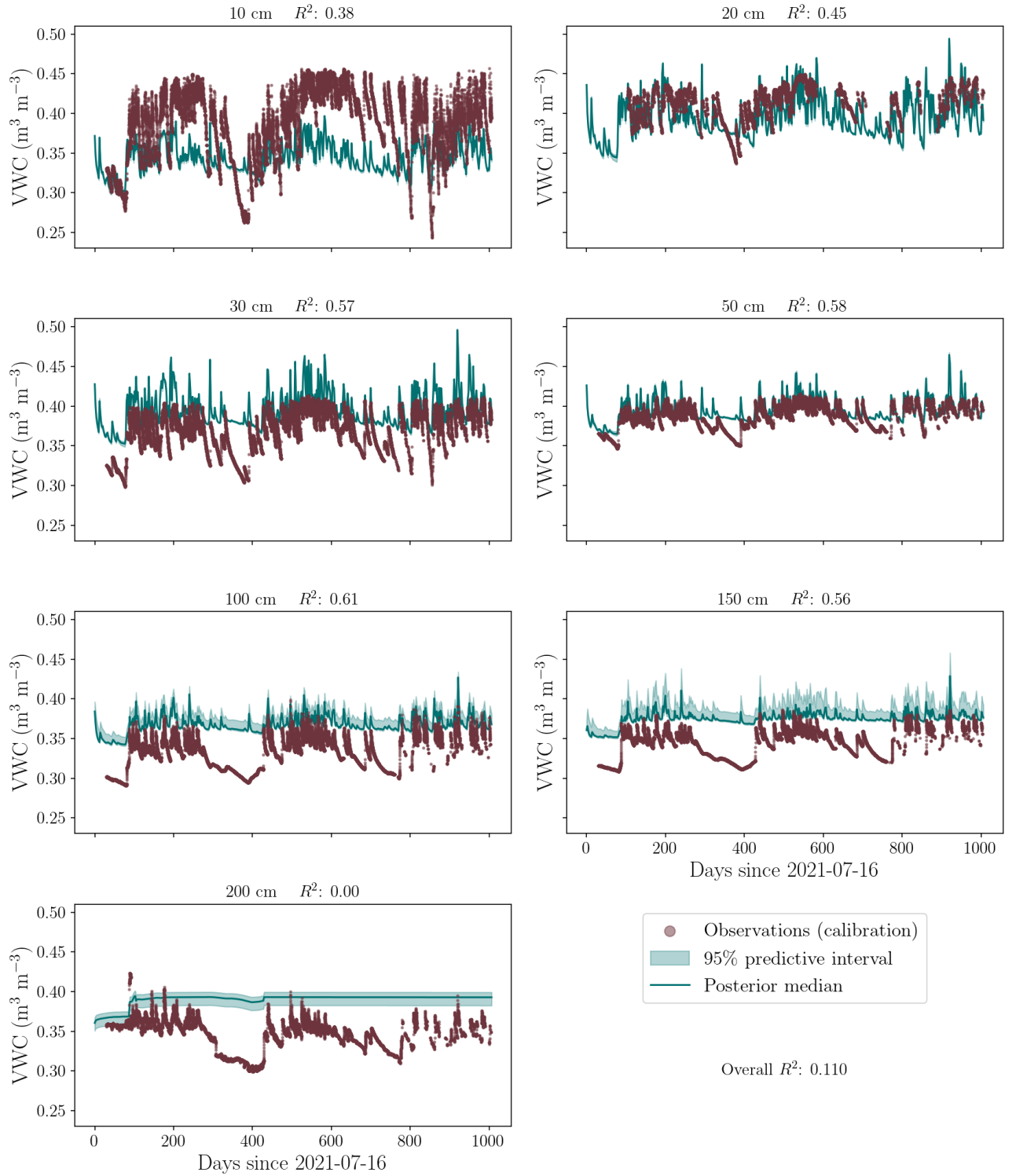


Figure 41: Posterior predictive simulation of volumetric water content at Pasture, by depth. Only calibration observations are available. The deeper layers show a systematic offset between the predicted band and the observations, which dominates the pooled R^2 .

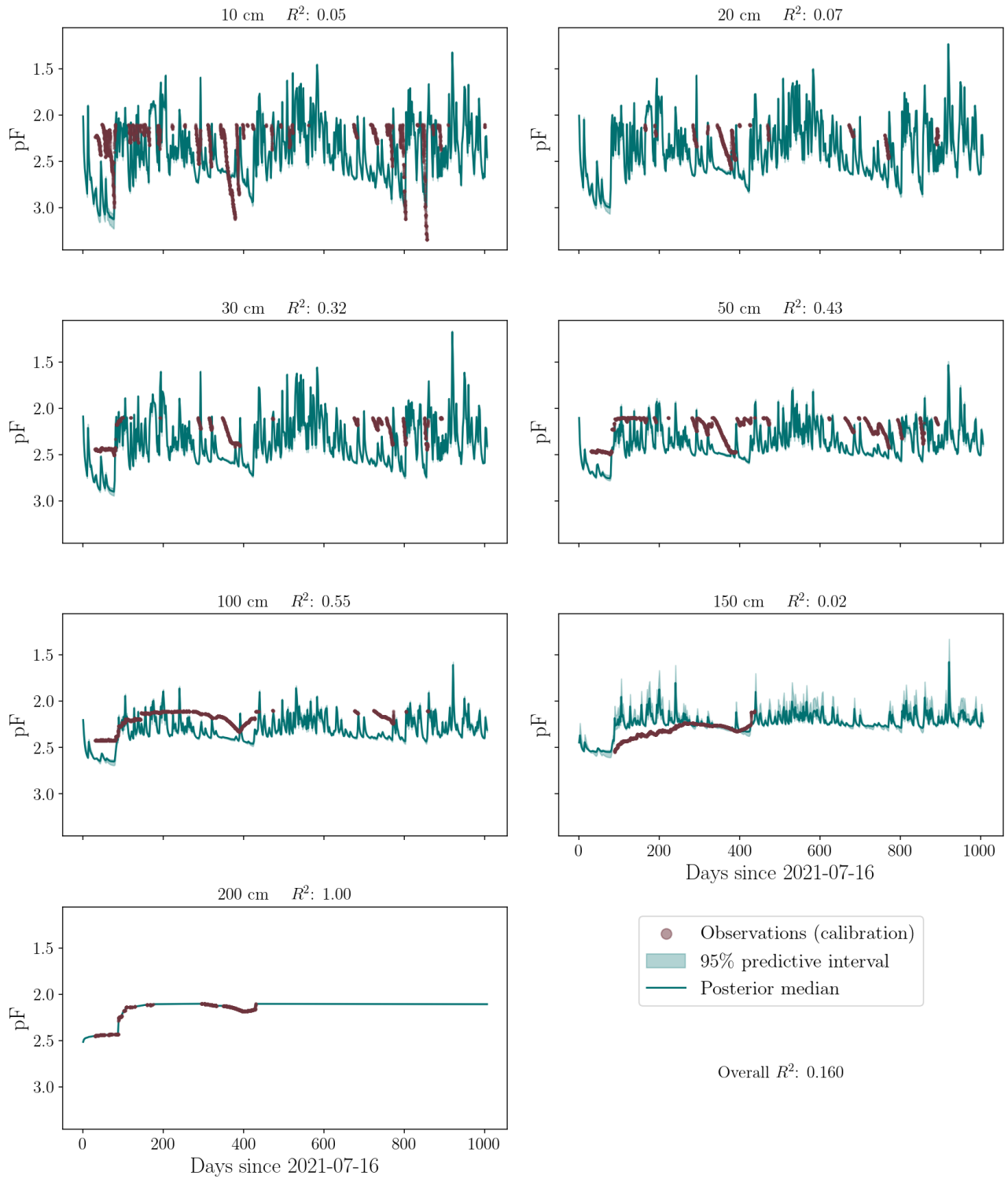


Figure 42: Posterior predictive simulation of matric potential (pF) at Pasture, by depth. The observations remain confined to a narrow band near the wet sensor limit, leaving little dynamic range to score against; the near-perfect 200 cm panel reflects the prescribed lower boundary condition.

Does reforestation increase or deplete soil water storage? Quantifying the effect of reforestation on the soil water balance: a HYDRUS-1D case study comparing a regenerated plot and an adjacent pasture in the Atlantic Forest, Brazil.

The Atlantic Forest has been heavily reduced by centuries of deforestation, and forest restoration is now a central conservation policy. Yet the consequences of regrowth for water resources remain unsettled, since a recovering forest both restores the soil structure that governs infiltration and raises the evaporative demand returned to the atmosphere, so whether a reforested plot is a net consumer or contributor of water must be quantified rather than assumed.

This thesis was carried out in collaboration with the INPE through the co-supervisor of this thesis, Laura Borma. It continues the series of studies conducted at the São Francisco Xavier site (São Paulo State). This present work compares two adjacent plots that differ only in land-use history, a 40-year regenerated forest and a degraded pasture, asking over a complete hydrological year whether the forest gains or loses soil water relative to the pasture and through which fluxes any difference arises. The one-dimensional water balance of each plot was simulated with HYDRUS-1D under the van Genuchten–Mualem relations with Lenhard hysteresis, the hydraulic parameters being constrained by Bayesian inverse estimation against in-situ records measured by the INPE.

Over three common hydrological years the pasture gains 4.4 cm of soil water while the reforested site loses 2.8 cm, a difference of about 7.2 cm in favor of the pasture, so the regenerated forest is a marginally larger net consumer of soil water. This difference is small and emerges as the residual of much larger, largely offsetting fluxes rather than from a single dominant process. The contrast clearly attributable to land use is the higher transpiration of the forest, whereas the recovered near-surface infiltration capacity had no effect, and the larger capillary supply that keeps the pasture in surplus is governed by the prescribed lower boundary rather than by an interactive water table. The forested profile is also far more hydrologically dynamic between wet and dry years. The study thus confirms the evapotranspiration side of the infiltration–evapotranspiration trade-off, while the sign and magnitude of the storage difference remain conditioned by the poorly constrained lower boundary and the single plot design, the main avenue identified for future work.

Lucas Jeangout