Characterising the Ill-posedness of PROSAIL Inversion for Biophysical Parameter Retrieval

Laurens Arp[a], Peter M. van Bodegom[b], Holger H. Hoos[c,a] and Mitra Baratchi[a]

[a]Leiden Institute of Advanced Computer Science, Leiden University, The Netherlands
[b]Institute of Environmental Sciences, Leiden University, The Netherlands
[c]Chair for Artificial Intelligence Methodology, RWTH Aachen University, Germany

ARTICLE HISTORY
Compiled February 20, 2026

ABSTRACT
The inversion of PROSAIL, a radiative transfer model widely used for vegetation parameter retrieval, is often considered to be an ill-posed problem. However, beyond this general label, the ill-posedness and its characteristics are not well understood. This study aims to theoretically characterise the inversion problem and its ill-posedness in terms of its loss landscapes, and to analyse the impact of strategies reducing ill-posedness in terms of these landscapes. We carried out five empirical experiments from the perspective of numerical optimisation, allowing for a rigorous analysis of the inversion loss landscape. We found that PROSAIL inversion, on its own, does not meet a strict definition of ill-posedness, because i) there exists a solution to the problem, ii) the solution is unique, and iii) the outputs (parameter configurations) are continuous with respect to the inputs (spectral observations). However, our experiments show that both random spectral noise and spectral mixing, combined with a 'plateau'-like spectral loss landscape, can explain ill-posed behaviour for the parameter retrieval tasks. We recommend a focus on data-centric approaches aimed at exploiting information complementarity in future work, as it appears unlikely that a model-centric focus alone will be able to overcome this type of ill-posedness in parameter retrieval.

KEYWORDS
model inversion; PROSAIL; parameter retrieval; ill-posedness; radiative transfer model; vegetation

1. Introduction

The estimation of biophysical parameters is an important task for the monitoring of ecosystems, planning interventions where appropriate, and modelling their impacts on other predictive tasks. For some parameters it is possible to perform in-situ measurement campaigns to directly measure the values of these parameters (Brown et al., 2020; National Ecological Observatory Network (NEON)). However, such campaigns are costly and time-intensive, and can only cover smaller, individual areas at a particular time, leading to scalability and representability issues for regular, global monitoring applications. This necessitates indirect retrieval methods to estimate these parameters from remote sensing data instead (Verrelst et al., 2019a). The retrieval performance can then be validated using available in-situ data.

CONTACT Laurens Arp. Email: l.r.arp@liacs.leidenuniv.nl

Among retrieval efforts, those reliant on airborne spectral data are larger in scale than in-situ missions, but are expensive and difficult to scale up further, while the retrieved variables may be less reliable. Due to the cost- and scale constraints, retrieval from spaceborne sensors is preferred for regular, global monitoring applications, but performing the parameter estimation required for this retrieval is not a trivial problem.

A common method of estimating these parameters is through the use of radiative transfer model (RTM) inversion (Bacour et al., 2006). RTMs are physical models that simulate a light spectrum based on physical input parameters. Possibly the most commonly used RTM for vegetation applications is PROSAIL (Jacquemoud et al., 1995, 2009), consisting of the leaf- and canopy vegetation parameters from its constituent models PROSPECT (Jacquemoud and Baret, 1990) and 4SAIL (Verhoef, 1984), used to estimate soil and vegetation parameters. Since the PROSAIL model is based on the causal relationship between biophysical parameters and light spectra, while light spectra can be readily observed through remote sensing technologies, these models must be inverted to perform retrieval. This RTM inversion task is widely considered to be an ill-posed problem: multiple solutions may fit the observations equally well (Combal et al., 2003; Quan et al., 2015; Sun et al., 2022).

RTM inversion can be performed based on two main approaches. Traditionally, numerical optimisation techniques have been used (Jacquemoud et al., 1994, 2000; Richter et al., 2009). More recently, hybrid modelling approaches, where a machine learning model is trained on a look-up table (LUT) of simulations, have gained popularity (Berger et al., 2021; Binh et al., 2022; Caballero et al., 2022; de Sa et al., 2021; Ranghetti et al., 2023; Rossi et al., 2022), in part because they can always provide an estimation for the parameters to retrieve, even for ill-posed problems where traditional inversion methods may fail to provide a prediction.

However, there is a danger that hybrid models obscure the ill-posedness. The reason is that, when the underlying problem is ill-posed, there may be more than one valid result to the inversion problem, which may be disjoint from the area around a point prediction, while the metrics for validating the performance of these models only evaluate performance based on (the confidence interval of) one of these many solutions.

Similarly, applying conventional machine learning approaches (as opposed to RTM inversion), such as training regression models on spectral observations and in-situ measurements, will be difficult because the training data are necessarily limited to a specific study area (see, e.g., (Bai et al., 2019; Liangyun, Liu and Bowen, Song, 2021; NASA, 2008)). Moreover, these regression models may be affected by the same ill-posedness as model inversion methods, if this is an inherent property of the spectral information.

Given the challenges it causes, much work on PROSAIL inversion has incorporated measures to reduce the ill-posedness. Most commonly, the ill-posedness is reduced by adding prior knowledge (priors) to the model, representing domain knowledge on, for example, certain types of vegetation known to be dominant in a study area (Bacour et al., 2002; Guo et al., 2024; Jacquemoud et al., 1995; Meroni et al., 2004; Wang et al., 2022). Based on this prior domain knowledge, the ranges of some key parameters can be reduced, which has previously been found to improve retrieval performance by addressing ill-posedness (Combal et al., 2003). Similarly, in the case of hybrid models, the ill-posedness can be reduced by training specialised models on simulated training data with reduced parameter ranges that are statistically probable for a specific study area (Binh et al., 2022; Candiani et al., 2022; Sahoo et al., 2023). Numerous active learning heuristics have been proposed to improve hybrid models (Binh et al., 2022), such as approaches sampling points with the highest uncertainty or automatically

matching an expected distribution (Candiani et al., 2022; Sahoo et al., 2023).

While such methods may have succeeded in improving performance on ill-posed problems for specific study areas, little work has been done on characterising the ill-posedness underlying the problem in more depth, investigating possible causes and analysing the impact of the mechanisms through which approaches such as constraining parameter ranges could reduce ill-posedness. Without a clear understanding of the critical characteristics of the ill-posed problem and its causes, it will be difficult to establish how to overcome ill-posedness, or even if it is a problem of PROSAIL inversion specifically, or a symptom of the overarching parameter retrieval problem. Such knowledge is critical to design scalable biophysical parameter retrieval models that are generalisable to a diverse range of environments.

In this work, we aim to address the gap in the knowledge about PROSAIL inversion for vegetation parameter retrieval, through a thorough empirical study of the theoretical properties of the ill-posedness of the problem. We do this through the lens of the inversion *loss landscape*. This landscape quantifies the goodness-of-fit (similarity between the simulated and observed spectra) of all possible combinations of parameters, when comparing their simulated output to observed spectra. By empirically studying the properties of this loss landscape, we can verify whether PROSAIL inversion meets the formal definition of an ill-posed problem, and if not, what could be other possible causes of ill-posedness for parameter retrieval.

With this knowledge, we hope to enable practitioners to focus their efforts to alleviate ill-posedness on the factors that contribute strongly to the ill-posedness of biophysical parameter retrieval. Through our experiments, we found that PROSAIL inversion is *not* ill-posed; however, parameter retrieval as a whole is. These results encourage future work to focus not only on further improving the effectiveness of finding the optimal solution for the PROSAIL inversion, which is actually a well-posed problem, but rather on addressing key limitations of the data in parameter retrieval as a whole, such as noise or spectral mixing.

Through our analyses, we aimed to answer the following research questions (RQs):

- **RQ1:** Does PROSAIL inversion meet the formal requirements of a well-posed problem?
- **RQ2:** What are the possible causes of ill-posedness in biophysical parameter retrieval through PROSAIL inversion?
- **RQ3:** How does adding priors to the parameter ranges impact the ill-posedness of biophysical parameter retrieval through PROSAIL inversion?

2. Materials and methods

In the following, we will explain the details of our methods and experimental setup. To this end, we will first introduce the terms and formal definitions that will be used throughout this section.

Definition 2.1: (biophysical) parameter. A variable describing a component of a biophysical system (vegetation in the case of PROSAIL), often a target variable to retrieve.

We denote an individual parameter as $p \in P$, where P is the set of all free parameters under study.

Definition 2.2: configuration. A set of concrete value assignments for all parameters $p \in P$.

We denote a configuration as $\mathbf{c}$, represented by a vector containing concrete value assignments for the parameters $p \in P$, where every jth value $\mathbf{c}^j$ in $\mathbf{c}$ corresponds to the jth parameter p^j in P. We denote the true configuration as $\mathbf{c}^+$, with the true values $\mathbf{c}^{j+}$ for individual parameters p^j.

Definition 2.3: parameter space (search space). A $|P|$-dimensional space of all possible configurations (also known as the search space in optimisation contexts), where every point corresponds to a specific configuration.

We denote the parameter space as $\mathcal{C}$, and for all configurations $\mathbf{c}$ it holds that $\mathbf{c} \in \mathcal{C}$.

For spectral information, we use S to denote an observed spectrum, consisting of individual measurements s^b at multiple spectral bands b. Our experiments contained 12 bands b, corresponding to the spectral bands of the popular Sentinel-2 satellite (level 2A data products; cirrus band B10 dropped after atmospheric correction).

Definition 2.4: instance. One specific problem scenario to solve, consisting of an observed spectrum S_i, and an unknown true configuration $\mathbf{c}_i^+$ to approximate.

A simulated spectrum can be obtained by running PROSAIL, denoted as M, on a configuration $\mathbf{c}$, which performs a simulation to generate the simulated spectrum $\hat{S} = M(\mathbf{c})$. This allows us to define a spectral loss function for a configuration $\mathbf{c}$, given the observed spectral information S:

Definition 2.5: (spectral) loss function. A function measuring the distance between a simulated spectrum and an observed spectrum, used to measure the goodness-of-fit of candidate solutions for optimisation purposes.

We denote the loss function as $\mathcal{L}(M(\mathbf{c}), S) = d(\hat{S}, S)$, where d can be any distance metric between $\hat{S}$ and S. In our experiments, we used the proportional mean absolute error (PMAE) and the spectral angle mapper (SAM) functions. PMAE is a variant of the mean absolute error (MAE) whose proportionality results in an equal weighting between the lower spectral bands (with lower intensities) and the infrared bands (with higher intensities). SAM focuses on the relationship between bands rather than absolute values, and may therefore measure complementary properties compared to standard error metrics like MAE (for details, see Appendix A.1).

If the loss function $\mathcal{L}(M(c), S)$ were evaluated for every point $\mathbf{c}$ in $\mathcal{C}$, a $|P|$-dimensional manifold in $|P|+1$-dimensional space would emerge, such that all possible coordinates (representing parameter configurations $\mathbf{c} \in \mathcal{C}$) are associated with a goodness-of-fit value quantified by the loss function. This manifold is known as a loss landscape:

Definition 2.6: loss landscape. A $|P|$-dimensional manifold in $|P|+1$-dimensional space, measuring loss function values for every possible configuration $\mathbf{c} \in \mathcal{C}$.

Finally, we denote an optimum (or minimum) minimising the loss function $\mathcal{L}$ for an observed spectrum S as $\mathbf{c}^*$.

Definition 2.7: optimum. A configuration $\mathbf{c}^* \in \mathcal{C}$ for which the loss function value is minimised compared to its direct neighbourhood (local optimum) or for the entire loss landscape (global optimum).

The objective of PROSAIL inversion, therefore, is to find this optimum $\mathbf{c}^*$:

$$\mathbf{c}^* \in \underset{\mathbf{c} \in \mathcal{C}}{\operatorname{argmin}} \mathcal{L}(M(\mathbf{c}), S) \tag{1}$$

It is possible for more than one configuration $\mathbf{c}$ to minimise the loss function, if

multiple configurations share the same loss function value.

2.1. Preliminaries

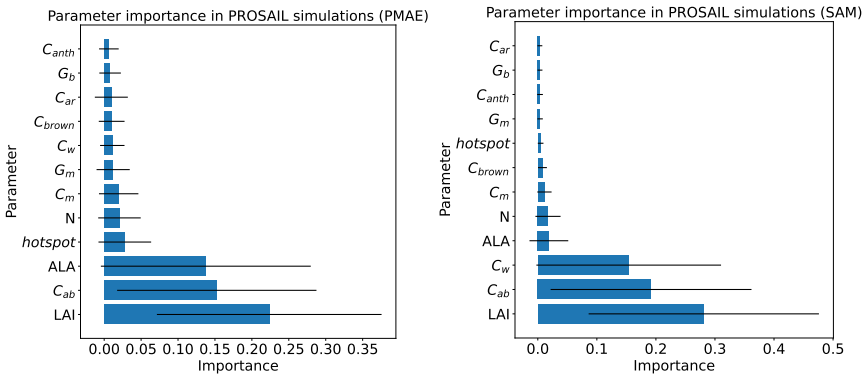
Our general study design was as follows. First, we used PROSAIL to generate simulated instances i. The sampling strategy used to select initial generating configurations $\mathbf{c}^+$ took the prior distributions of individual parameters (see Table 1 for details) into account. We generated a simulated look-up table (LUT) D, consisting of 1000 instances i combining the true configurations $\mathbf{c}_i^+$ and the associated simulated spectra S_i. We mapped the raw PROSAIL outputs, which contain 1300 hyperspectral bands (wavelengths of 400–2700nm in steps of 1nm), to the multispectral format of the popular Sentinel-2 satellite using the spectral responses provided by the European Space Agency (ESA), thereby conforming to a realistic application setup.

2.1.1. Parameter importance and correlation

The PROSAIL model contains 15 numerical parameters that could be explored, although 3 of these concern observer- and solar geometry, which are known in practice, resulting in 12 potential free parameters. Not all of these parameters are equally impactful to the spectral loss, making some more appropriate to retrieve through PROSAIL inversion than others, while exploring all parameters also makes the problem prohibitively computationally expensive. In the interest of efficiency, we selected the most impactful parameters based on their importance in a preliminary experiment, determined by functional ANOVA (fANOVA) (Hutter et al., 2014) and Sobol indices (Sobol, 2001).

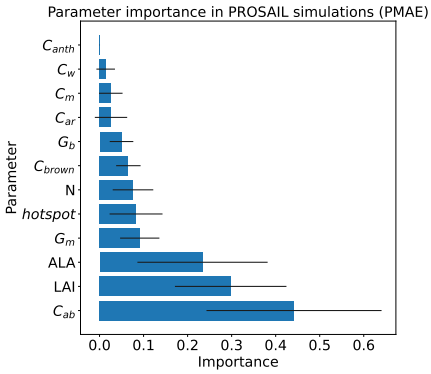
In fANOVA, the variance of a surrogate model approximating a target function (in our case PROSAIL forward simulations) is partitioned per input parameter. By using a surrogate model, the multiple function evaluations necessary to determine the variance of the model can be computed efficiently for potentially complex, high-dimensional functions. The change in output can be computed as a function of changes in the input, resulting in sensitivity indices that give an indication of parameter importance. Similarly, Sobol indices can be computed by decomposing the total variance of the output into the fraction of the total variance explained by every parameter. These indices can also be computed for higher-order indices; we included both individual contributions and parameter interactions. Parameters with a high sensitivity have a large impact on simulated spectra, while parameters with a low sensitivity have a low impact on the spectra. Therefore, we performed experiments on parameters with a high impact on the spectra, since these are the most suitable parameters to retrieve via PROSAIL inversion.

Concretely, we generated 200 additional instances, keeping all 12 parameters as free parameters. For fANOVA, we sampled 200 possible configurations with their spectral loss for every instance. For Sobol indices, we sampled 2048 configurations with their spectral loss, since this approach likely needed a larger sample size for reliable results. For both methods, this approach allowed us to compute the importance of different parameters for specific instances, which we then aggregated over all 200 additional instances. We performed this preliminary experiment for both the PMAE and SAM loss functions, and the resulting distributions can be found in Figure 1. Some parameters that were very important for PMAE (such as average leaf angle – ALA), were far less important for SAM, and vice versa (such as leaf water content – C_w). The results for fANOVA and Sobol indices largely overlapped, although the

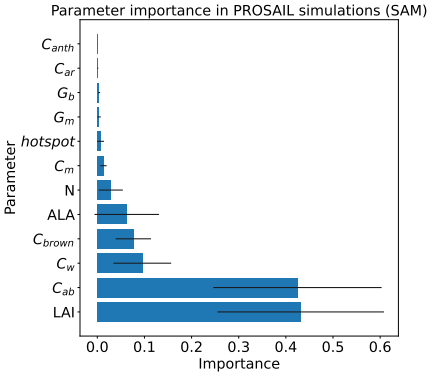


(a) fANOVA, PMAE

(b) fANOVA, SAM



(c) Sobol, PMAE



(d) Sobol, SAM

Figure 1.: Parameter importance for the spectral loss landscape determined using functional ANOVA (Hutter et al., 2014) (1a, 1b) and Sobol indicies (Sobol, 2001) (1c, 1d). The PROSAIL parameters shown in these plots are leaf area index (LAI), chlorophyll $a + b$ content (C_{ab}), average leaf angle (ALA), hotspot, structural N, leaf dry matter content (C_m), leaf water content (C_w), leaf brown pigment content (C_{brown}), carotenoid concentration (C_{ar}), anthocyannins (C_{anth}), soil moisture (G_m) and soil brightness (G_b). The relative importance of parameters can be heavily reliant on the spectral loss function used: proportional mean absolute error (PMAE) or spectral angle mapper (SAM).

order of LAI and C_{ab} was switched for PMAE. The results mainly started to differ starting from the fourth-most important parameter (e.g., soil moisture G_m). The most impactful parameters tended to have the highest impact on specific spectral bands; details can be found in Appendix D.1.

Based on the information in Figure 1, we opted to include all parameters that were in the top 3 most impactful parameters for either of the loss functions and either sensitivity analysis method. This resulted in a parameter selection of leaf area index (LAI), chlorophyll $a+b$ content (C_{ab}), average leaf angle (ALA), and leaf water content (C_w). A description of these parameters, along with and their ranges, prior distributions and default values, can be found in Table 1. All other parameters were kept at default values (as in the study by de Sa et al., 2021).

Table 1.: PROSAIL parameters with their names, ranges, values, and distributions, for the parameters we kept variable in our experiments. A distribution of $\mathcal{N}(\mu, \sigma)$ refers to a normal distribution with mean μ and standard deviation σ, and a distribution of $\mathcal{U}(min, max)$ refers to a uniform distribution within the specified bounds. The ranges were determined through the parameter ranges specified in the documentation of the PROSAIL implementation[1].

Parameter	Abbreviation	Distribution	Range	Default	Unit
Leaf area index	LAI	$\mathcal{N}(3.88, 1.98)$	(0.1 – 10.0)	3.88	m^2 leaf $/m^2$ soil
Chlorophyll a+b	C_{ab}	$\mathcal{N}(32.81, 18.87)$	(0.3 – 106.72)	32.81	$\mu g/cm^2$
Average leaf angle	ALA	$\mathcal{U}(0, 90)$	(0 – 90)	45	$\circ$
Leaf water content	C_w	$\mathcal{N}(0.0129, 0.0073)$	(0.0043 – 0.07)	0.0129	cm

2.1.2. PROSAIL inversion approach

Since we are interested in the characteristics of the inversion loss landscape, our experiments were based on numerical optimisation. Numerical optimisation methods, such as black-box optimisation techniques (e.g., stochastic local search procedures (Hoos and Stützle, 2018), evolutionary algorithms (Bäck, 2018) and swarm-based metaheuristic algorithms (Chakraborty and Kar, 2017)), can provide richer insight into the underlying loss landscape of the inversion problem compared to the point predictions (or distribution parameterisation) of hybrid models, because they sample along the loss landscape.

We used a greedy local search algorithm with a budget of 10 000 function evaluations (simulations with loss function value computations) as an optimisation algorithm. Greedy local search can converge quickly to the global optimum in unimodal settings, though it may get stuck in local optima in multimodal settings. In our experiments, the downside of local optima worked to our advantage, because it allowed us to check for the number of optima in loss landscape by determining whether the optimisation algorithm converged to different local minima. It also enabled us to cover a larger part of the parameter space than the 1000 instances included in our dataset, as part of the 10 000 function evaluation budget used for finding the optimum was spent on exploring the search space. This further increased the probability that, if there are local irregularities in the search space, they would be encountered along the way. For further details on the optimisation algorithm, we refer to Appendix A. The budget was sufficient for our experiments; this can be validated through the plots in Appendix B.

A summary of our experiments and the research questions they correspond to can be found in Table 2. We will describe the experiments in detail in the following.

2.2. Ill-posedness characteristics (RQ1)

Ill-posed problems are problems that do not meet the requirements of well-posedness as defined below:

Definition 2.8: well-posed/ill-posed. A problem is well-posed if and only if i) there is a solution to the problem, ii) this solution is unique, and iii) the appropriate solution changes continuously with changes in the observations (no sudden jumps in the parameter space) (Kabanikhin, 2008). If one of these conditions is not met, the

[1] `https://github.com/jgomezdans/prosail`

Table 2.: Summary of experiments and their target research questions.

Experiment	Description	Research question
E1	Testing for unimodality VS multimodality.	RQ1
E2	Testing for continuity of the relationship between input spectra and the identified optimum.	RQ1
E3	Testing for the average shift in optimum found for various levels of Gaussian noise on the spectral observations.	RQ2
E4	Testing for the impact of spectral mixing on the retrieval estimations.	RQ2
E5	Testing for the mechanism through which range constraint priors alleviate ill-posedness in retrieval problems.	RQ3

problem is considered ill-posed.

Therefore, the first characteristics to test for are whether the conditions of well-posedness are met in PROSAIL inversion. Our experimental setup for these tests are as follows.

2.2.1. A solution exists.

The objective of PROSAIL inversion is to find a configuration $\mathbf{c}^*$ that minimises the spectral loss $\mathcal{L}(\hat{S}, S)$ (see Equation 1).

If there exists any solution $\mathbf{c}^*$ minimising the loss function $\mathcal{L}$, this property is satisfied. In continuous problems without constraints or undefined operations (e.g., zero divisions), this property is trivially satisfied: if any configuration $\mathbf{c}$ has a valid output $M(\mathbf{c})$, there will be at least one configuration minimising $\mathcal{L}(M(\mathbf{c}), S)$. In PROSAIL inversion, assuming the ranges of the paramters have been set up correctly, this will always be the case.

2.2.2. The solution is unique.

There is no guarantee that the optimum to a PROSAIL inversion problem is unique. If there are multiple optima (modalities) in the parameter space that would explain the observed spectra equally well, or perfectly flat areas with exactly the same optimal loss, the inversion problem is ill-posed by violating property 2 of well-posed problems (unique solution). Therefore, we tested for multimodality in our first experiment.

Experiment 1 (E1): In this experiment, we iterated over instances in D, and performed iterated greedy local search (see Appendix A.2 for details) (Hoos and Stützle, 2018) 5 times with a random initialisation sampled uniformly (overriding the normal distributions described in Table 1, as this could bias the experiment), resulting in a new local optimum for every iteration. For every instance, we computed the largest distance between any pair of optima out of the 5 optima in the set. If the landscape is unimodal, the optimisers should all converge to the same point in the parameter space regardless of their initialisation, resulting in a low maximum distance. If the landscape is multimodal, the optimisers can converge to different points in the space, resulting in a large maximum distance.

If more instances contain greater distances than can be explained through minor approximation inaccuracies of the optimisation algorithm (because the optimum for continuous optimisation problems will generally not match the true configuration exactly: $\mathbf{c}^* \approx \mathbf{c}^+$), we can conclude that the landscape is likely multimodal, and the PROSAIL inversion problem is ill-posed.

2.2.3. Solution continuous with observations.

Unlike the previous two characteristics, which focused on the optimal solutions to the problem, this characteristic describes the underlying loss landscape. Intuitively, well-posed problems should not have sudden 'jumps' in their solution space: if the inputs (in this case: spectra) shift by a certain amount, the shift observed in the outputs (in this case: optimal solution) should be proportional to this shift. Concretely, for any input spectrum S and its associated optimum $\mathbf{c}^*$, given another spectrum S' and its associated optimum $\mathbf{c}'^*$, the new optimum $\mathbf{c}'^*$ should converge to the original optimum $\mathbf{c}^*$ as the new spectrum S' approaches the original spectrum S. If this property is not satisfied, errors in some parts of the parameter space could become unpredictable, as negligible inaccuracies could still result in a large error due to a 'jump' through the space.

Experiment 2 (E2). We tested for this property by (deterministically) mapping the observed spectrum S into a perturbed spectrum S' by perturbation levels β of -10%, -1%, -0.1%, 0%, 0.1% 1% and 10% of the mean band value ($S'^b = S^b + \beta * \overline{S^b}$), and computing the optimum. We compared the change between perturbations and their associated change in the optima. If the distance between the perturbed optima and the original optima converges to 0 as the perturbation intensity approaches 0, PROSAIL inversion likely meets the continuous input-output relationship requirement of well-posedness.

2.3. Causes of ill-posedness (RQ2)

When aiming to understand the source of ill-posedness in parameter retrieval through PROSAIL inversion, it can be beneficial to disentangle the parameter retrieval problem from the (PROSAIL) model inversion problem, where solving the latter can be used as a proxy to solving the former. In parameter retrieval, the objective is to obtain an estimate $\mathbf{c}^*$ of the true configuration that is as close as possible to the real configuration values $\mathbf{c}^+$, such that a retrieval loss function $\mathcal{L}_R(\mathbf{c}^*, \mathbf{c}^+)$ (e.g., mean squared error) is minimised. This problem can be addressed through, e.g., standard machine learning or hybrid models (in which case a machine learning training problem must be solved), or by performing PROSAIL inversion (which involves solving a model inversion problem). When inverting PROSAIL, the objective is to obtain an optimal configuration $\mathbf{c}^*$ for which its simulated spectrum $M(\mathbf{c}^*)$ matches the observed spectrum S as well as possible, minimising the spectral loss function value $\mathcal{L}\big(M(\mathbf{c}^*), S\big)$. The implicit assumption here is that the $\mathbf{c}^*$ found through PROSAIL inversion corresponds to the $\mathbf{c}^*$ of the parameter retrieval problem.

Ill-posedness on the PROSAIL inversion problem indicates that a unique solution cannot be reliably found (e.g., due to parameter non-identifiability through compensation effects among parameters or spectral ambiguity, particularly in the infrared bands of Sentinel-2), or that the loss landscape is non-continuous. In contrast, ill-posedness on the parameter retrieval problem can indicate that the retrieval problem may be underdetermined or ill-defined, due to the information contained in the spectral data S being insufficient to uniquely determine $\mathbf{c}^*$, or because there is a mismatch between real-world measured data and the input data expected by PROSAIL. If PROSAIL inversion is ill-posed, the parameter retrieval using it is also ill-posed, but PROSAIL inversion is not necessarily ill-posed if parameter retrieval is. This is a highly relevant distinction because, if the parameter retrieval is ill-posed, but not the PROSAIL inversion, this implies that tweaks to the inversion techniques (e.g., LUT-based hybrid

models) or fully data-driven approaches, would still suffer from ill-posedness. In this vein, we hypothesise three possible main causes of the ill-posedness experienced by practitioners when performing parameter retrieval through PROSAIL inversion.

First, as model inversion is often assumed to be an ill-posed problem (see, e.g., Darvishzadeh et al. (2008); Verrelst et al. (2014, 2019a)), the PROSAIL inversion problem may indeed be ill-posed, which we already test for in E1 and E2. However, even if the PROSAIL inversion problem itself is not ill-posed, the overarching parameter retrieval problem could still be. It is possible that noise and uncertainty in the observed data are causing the uncertainty of solutions for different instances to overlap, resulting in ill-posedness for the retrieval task. Moreover, it is possible that there are observable spectra for which there are no realistic solutions, for example, due to limitations of the scope of the simulation model, or due to the effects of spectral mixing (Cavalli, 2023).

Therefore, in addition to PROSAIL inversion itself, we consider two possible alternative causes through which parameter retrieval through PROSAIL could be ill-posed: *noise combined with ill-conditionedness*, where small amounts of spectral noise can overpower the signal of matching simulated and observed spectra, and *spectral mixing* (mixed-pixels), where the observed spectra originate from multiple heterogeneous source spectra, whose combination may not correspond to a meaningful configuration. These characteristics prominently differentiate real-world settings from idealised simulation settings, making them appealing candidates to evaluate. We performed duplicate versions of Experiments 1 and 2 for each of these conditions, to verify that the well-posedness of the PROSAIL inversion still holds, even in these changed conditions. Furthermore, since there is no guarantee that our list of possible causes of ill-posedness was exhaustive, we included experiments on real-world Sentinel-2 data (details can be found in Appendix C). If the patterns hold even for real-world data, this means that PROSAIL inversion is not the source of the ill-posedness.

We will explain our experiments to test these possible causes in the following.

2.3.1. Noise and conditioning

Noise on the spectral observations, for example, through interference on or sensor limitations of the spectrometer used, has previously been found to have a strong impact on retrieval performance (de Sa et al., 2021). When there is noise on the observations S, the spectrum is perturbed into a new position $S^{'}$ in the spectral space. When this happens, the optimal solution $\mathbf{c}^{*}$ will shift away from the true configuration $\mathbf{c}^{+}$ (we later show examples of this in Figure 3), because the loss function $\mathcal{L}$ to minimise is now considering the perturbed spectrum $S^{'}$ instead of the noise-free version S. Since the noise per instance is unknown *a priori*, two instances i_1 and i_2, with highly similar spectra $S^{'}_{i_1} \approx S^{'}_{i_2}$, could have entirely dissimilar true solutions $\mathbf{c}^{+}_{i_1} \not\approx \mathbf{c}^{+}_{i_2}$, because their original, dissimilar noise-free spectra $S_{i_1} \neq S_{i_2}$ were pushed together through unpredictable noise. This problem can be exacerbated depending on the level of *conditioning* of the problem, which can be interpreted as the sensitivity of the loss landscape to perturbations to the spectral observations (Belsley and Oldford, 1986; Varah, 1973).

Experiment 3 (E3). We tested for this property in Experiment 3 by perturbing dataset D by adding various levels β of randomly sampled Gaussian noise (in contrast to the deterministic perturbations of E2) at 1%, 2%, 5%, 10% and 20% of the mean band value $\overline{S_b}$ in D to the spectral data per band b: $S'^{b} = S^{b} + \mathcal{N}(0, \beta * \overline{S^{b}})$. We then computed the average shift of the optima between the noise-free version of the

data, and the noisy versions thereof. If this shift is high, especially for relatively low amounts of noise (indicating ill-conditionedness), any point in the parameter space that an optimum could have shifted from, can be considered a potential true solution to the problem. In this case, noise on the observed spectra can be considered a source of ill-posedness for the parameter retrieval problem.

2.3.2. Spectral mixing

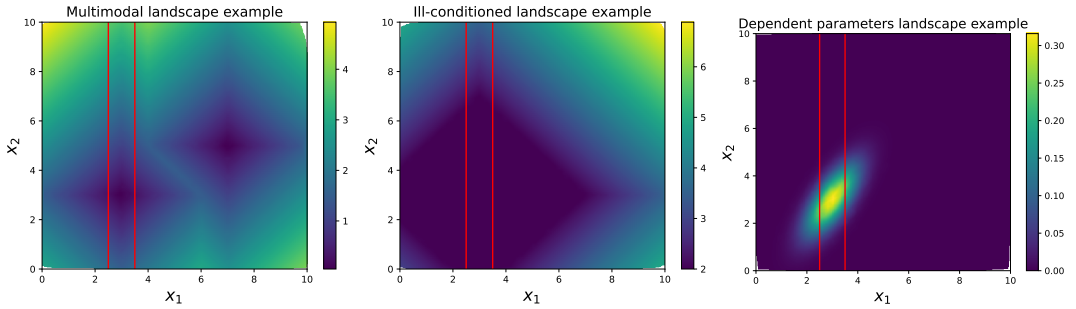
In real-world applications, spectral mixing is effectively inevitable. PROSAIL can only model a single set of parameters (though some of the structural parameters are already aggregates), thereby assuming a homogeneous vegetation cover for the entire area covered by a pixel (which can be interpreted as, for example, a mean of the vegetation types in the area). Spaceborne observations by the Sentinel-2 satellite cover, at best, an area of $100m^2$, which may contain a combination of several highly diverse vegetation types, or even land cover not related to vegetation (such as buildings or geological features). There is no guarantee that, if the spectra from different vegetation sources are mixed at a certain proportion, their optimal parameter values would consist of, e.g., a weighted average with the same proportional representation. Therefore, the optimum for an observed spectrum, when mixed, may not correspond to a meaningful parameter configuration.

Experiment 4 (E4). In Experiment 4, we tested whether a clear solution can still be found for observations where spectral mixing has occurred. We modified the LUT generation procedure to generate new instances as a weighted combination of three randomly sampled configurations and their simulated spectra, parameterised by the randomly sampled weight parameters α_1, α_2 and α_3 (where $\alpha_1 + \alpha_2 + \alpha_3 = 1$). This results in three distinct, independent spectra S_1, S_2 and S_3, simulated from their true parameter configurations $\mathbf{c}_1^+$, $\mathbf{c}_2^+$ and $\mathbf{c}_3^+$, that are combined into a single mixed spectral observation $S' = \alpha_1 S_1 + \alpha_2 S_2 + \alpha_3 S_3$. We then ran our optimisation setup for S' to find the predicted configuration $\mathbf{c}^*$, and compared this to the individual true configurations $\mathbf{c}_1^+$, $\mathbf{c}_2^+$ and $\mathbf{c}_3^+$, as well as their weighted mean $\alpha_1\mathbf{c}_1^+ + \alpha_2\mathbf{c}_2^+ + \alpha_3\mathbf{c}_3^+$. In retrieval settings, the weighted mean of configurations is a likely target value to predict for mixed spectra.

If the retrieval performance for all these cases (particularly the mean) is worse than the performance in cases without spectral mixing, spectral mixing can be considered a cause of ill-posedness in parameter retrieval. The retrieval problem would then become ill-defined, since the spectrum no longer corresponds to any single configuration, thereby violating characteristic 1 (a solution exists) of well-posed problems, despite being trivially satisfied for PROSAIL inversion itself. It is also possible that the problem would violate characteristic 2 (the solution is unique) of well-posed problems, because the underlying mechanisms of the spectral mixing can be considered a type of random noise. In these cases, it may be advisible for future work to further explore the impact of spectral unmixing techniques (see, e.g., (Cavalli, 2023; Kowalski et al., 2023; Xu and Somers, 2021)) on parameter retrieval through multispectral data.

2.4. Impact of range constraint priors (RQ3)

A commonly used method for reducing the ill-posedness of PROSAIL inversion is the addition of priors in the form of range constraints, which was previously found to improve performance (Combal et al., 2003). We consider three possible (not mutually exclusive) mechanisms through which the ill-posedness of PROSAIL inversion could be



(a) Multimodality (mechanism i))

(b) Ill-conditioning (mechanism ii))

(c) Dependency (mechanism iii))

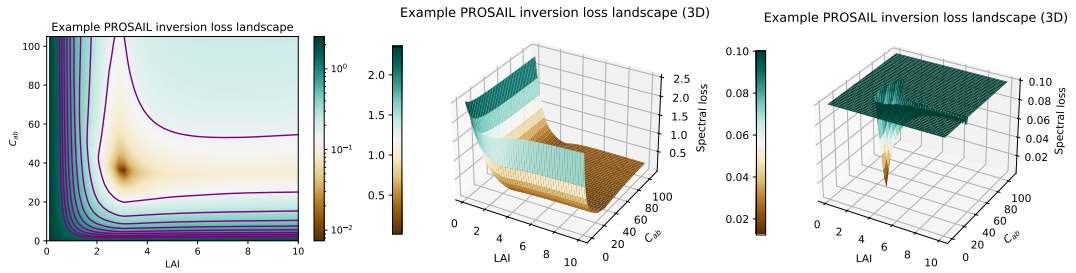
Figure 2.: Examples of landscapes and the possible mechanisms of reduced ill-posedness through range constraint priors on the x_1 variable (red vertical lines). In Figure 2a, the constraint on the x_1 variable excludes the second optimum at $(7, 5)$, thereby reducing ill-posedness through mechanism i). In Figure 2b, there are many points with a value very close to 2 (plotted simply as 2 for illustrative purposes), but the range constraint prevents errors larger than 1 (the width of the range) through mechanism ii); without the constraint, the model may have predicted points like $(4, 6)$. Finally, in Figure 2c (where, for plotting convenience, we are maximising instead of minimising), due to the dependency between x_1 and x_2, the range constraint prior on x_1 reduces the viable range for x_2 from about $[1, 5]$ to about $[1.5, 4.5]$ through mechanism iii).

reduced by adding range constraint priors: i) excluding competing optima, ii) reducing the maximal magnitude of errors, and iii) parameter dependency. We show abstract examples of these possible mechanisms in Figure 2.

If the loss landscape is multimodal (see Section 2.2 for details), the mechanism through which the ill-posedness is reduced would be intuitive: restricting the ranges of parameters to known feasible parts of the parameter space would rule out optima in different parts of the parameter space, making it more likely to converge to the correct optimum. This mechanism is illustrated in Figure 2a.

Meanwhile, if the landscape is unimodal, but it is ill-conditioned (see Section 2.3.1 for details) with a wide range of values evaluating to similar losses, adding range constraints could shrink the range of potential points with similar loss values that an optimiser could converge to for noisy spectral observations. In this case, performance would be improved because the degree to which predictions can be wrong would be limited; this mechanism is illustrated in Figure 2b. However, in this case it is possible that retrieval performance does not significantly differ from random sampling within the specified prior ranges; if this is the case, performing any retrieval at all would not provide any additional posterior information over the already available prior knowledge, which may reduce the appeal of using priors to reduce ill-posedness for parameter retrieval problems.

Finally, if there are dependencies between parameters (for example, more possible C_{ab} combinations for low LAI), constraining the ranges of some parameters may result in reduced ranges for other parameters as well, as illustrated in Figure 2c. In the example, the x_2 parameter has a range of about 1 to 5. However, its lowest values can only be reached if the value for x_1 is also low, and its highest values are only reachable if the value for x_1 is high. By adding a range constraint to x_1 as prior knowledge, the



(a) Loss landscape contour plot (b) Loss landscape 3D plot (c) 3D plot with clipped losses

Figure 3.: Example of a loss landscape for PROSAIL inversion, as a 2D plot with contour lines at a log scale emphasising small differences for low values (3a), a 3D plot (3b), and a 3D plot with clipped loss values (3c). The landscape is unimodal, meaning only a single point in the landscape locally minimises the spectral loss function. However, for most of the loss landscape, the differences in the spectral loss are very small, and only get larger as the parameters near a value of 0 (as indicated by the concentration of contour lines in Figure 3a). In Figure 3b, there appear to be no differences in loss values within a large plateau of configurations with similar spectral losses. Only when artificially clipping the loss values to set all losses > 0.1 to 0.1 can the global minimum be seen in the 3D loss landscape (shown in Figure 3c).

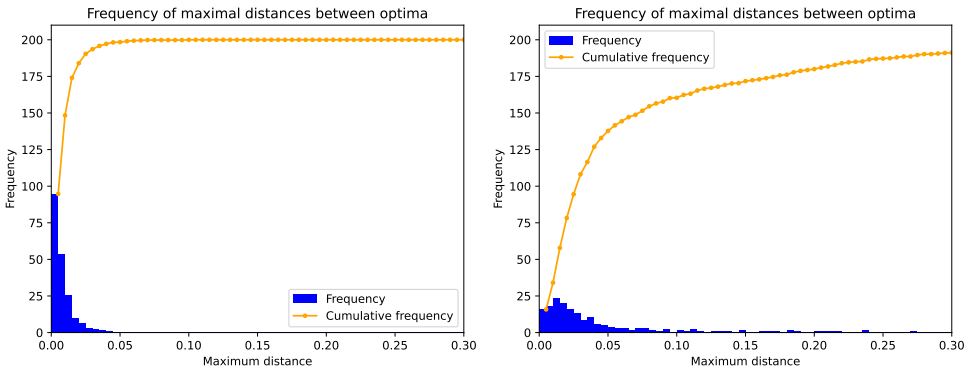
viable range for x_2 is also reduced to a range from about 1.5 to 4.5, thereby reducing the ill-posedness.

Experiment 5 (E5). We experimentally gauged in E5 whether the impact of adding priors comes from mechanism i), ii) or iii). If our results for the multimodality experiment described in Section 2.2 show a multimodal landscape, we can assume that range constraint priors reduce ill-posedness through mechanism i) (excluding competing optima). To test for mechanisms ii) and iii), we performed optimisation on the instances in D (with 10% Gaussian noise on spectral observations as described in Section 2.3.1, to ensure the problem is ill-posed), setting priors on LAI with a range interval around the true value at 0%, 10%, 30%, 50% and 100% of the total LAI range. We then repeated this procedure by using random sampling for LAI within its range interval, as a baseline measurement.

If the retrieval performance, measured only on LAI, is better for tighter range intervals than for larger intervals, it is likely that the ill-posedness was reduced through mechanism ii). However, if there is no difference with the random sampling-based baseline, though the prior knowledge would improve performance, it is not synergistic with the retrieval method, and rather replaces it entirely, as the posterior equals the prior. Finally, if the retrieval performance for *other* parameters is significantly better with tighter range constraints for LAI than with a looser or absent constraint, it is likely that reducing the range of LAI also reduced the viable ranges of other parameters, making mechanism iii) more likely.

3. Results

In the following we will, unless otherwise stated, analyse the results for the PMAE loss function. Due to their patterns being largely the same as the results for PMAE, the results for SAM can be found in Appendix D.2.



(a) PMAE (b) SAM

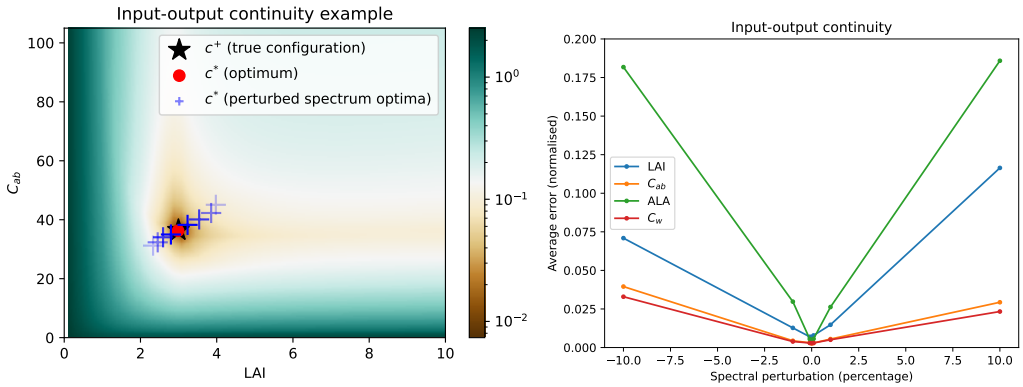
Figure 4.: Histogram and cumulative frequency of the maximal distances between 5 optima (initialised uniformly randomly) obtained using iterated local search, per instance. The bins on the x-axis (restricted to 0–0.3 instead of the full 0–1 range for better visibility) represent distance in the parameter space as a proportion of the total parameter range. For example, for the LAI parameter with a range of 0 to 10, a proportional distance of 0.1 represents an LAI difference of 1. The results for the PMAE loss function can be found in in Figure 4a, and the results for SAM can be found in Figure 4b. Since almost all 1000 maximum distances are contained within the first bins with the smallest distances, the restarted, randomly initialised optimisation procedure appears to converge to the same local optimum every time, indicating unimodality.

3.1. RQ1: Ill-posedness characteristics

Experiment 1. We visualised an example loss landscape for PROSAIL inversion in Figure 3. In this example loss landscape, the landscape appears to be unimodal, as can be seen in Figure 3a, with a large plateau of nearly identical spectral loss values surrounding that optimum that can be seen in Figure 3b. The optimum itself can only be seen when limiting the range of spectral losses by a substantial margin in Figure 3c (2.5 to 0.1, a reduction of 96%). Therefore, based on this example, PROSAIL inversion appears to be unimodal, but ill-conditioned: given the small margins of loss function values separating the optimum from the plateau, a small perturbation to the input spectrum would likely have a large impact on the output prediction.

To generalise this observation to a general pattern over all instances, we aggregated the results of Experiment 1 over instances by plotting a histogram of the maximum distances in Figure 4. As the figure shows, the optimisation algorithm almost always converged to the same optimum after being randomly restarted, indicated by the heavy skew toward the lowest value bins. For this experiment, we included the results for SAM (Figure 4b) because, while they were similar to those of PMAE (Figure 4a), the skew was slightly less extreme, though most maximal differences were still within a 5% distance of the parameter ranges. Overall, these results provide empirical evidence that there exists only a single local and global optimum in the loss landscape of PROSAIL inversion, leading to a unimodal problem. Therefore, PROSAIL inversion meets requirement ii) of well-posedness.

Experiment 2. The continuity between inputs (spectra) and outputs (optimal solution) for PROSAIL inversion tested in Experiment 2 can be found in Figure 5. In



(a) Example visualisation

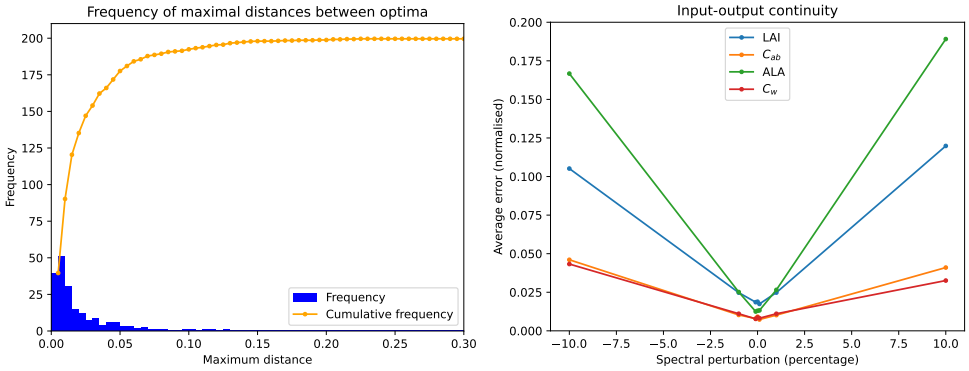
(b) Shift per perturbation

Figure 5.: The continuity of the output for PROSAIL inversion (predicted configuration c^*) with respect to perturbations to the input (spectrum). Figure 5a shows an example of a single instance (with only the LAI and C_{ab} parameters free), with colours representing its loss values, forming a loss landscape. The true configuration c^+ is visualised as a star, which the noise-free optimum c^* (the red dot) matches nearly perfectly. The blue +-signs represent the shifted versions of the optimum c^*, where points with lower opacity represent a larger perturbation to the spectral input. Figure 5b aggregates this phenomenon (mean) over all 1000 instances, normalised to a 0-1 range based on the bounds of the parameter range, showing that it is a consistent pattern.

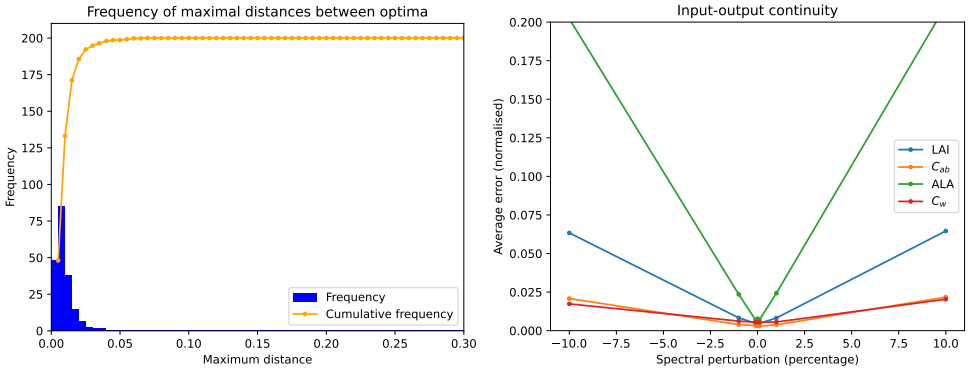
the example from Figure 5a, the blue points marked by a +-sign formed a clear line through the parameter space, where the points with a lower perturbation (high opacity) were closer to the optimum for the unperturbed spectrum, while the points with a higher perturbation (low opacity) were further removed from this original optimum. This indicates that, in this example, the location of the optimum (output) shifted smoothly and continuously with changes to the input (observed spectrum), thereby meeting requirement iii) of well-posedness. As Figure 5b shows, this pattern continued to hold when aggregated over all 1000 instances: for all parameters, a convex shape can be observed, indicating that optima converged to the optimum of a specific spectrum as their spectra approached this spectrum. If the property of input-output continuity had not been met, there would have been instances where small perturbations would have resulted in sudden, large jumps across the parameter space, which would not have resulted in the parabola-like shapes in Figure 5b that we observe. Therefore, we conclude that PROSAIL inversion meets requirement iii) of well-posedness.

In conclusion for RQ1, the results from Experiments 1 and 2 imply that PROSAIL inversion is not an ill-posed problem: there is always a solution, this solution is unique, and the output moves continuously with respect to the input.

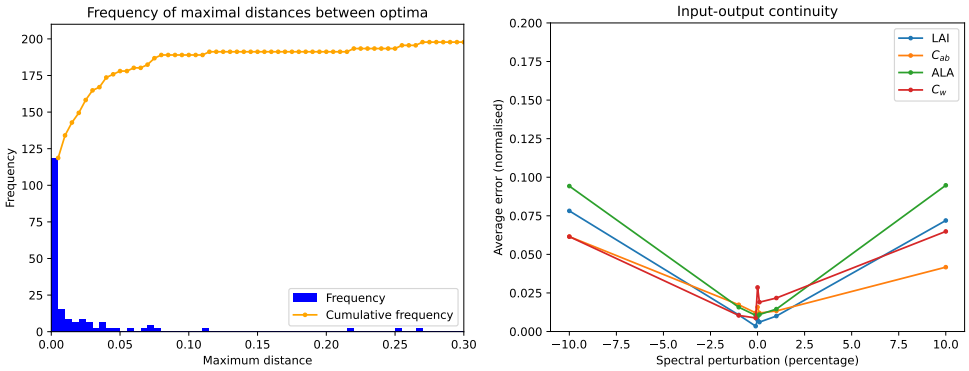
Our results for the additional runs of E1 and E2 for noisy, spectrally mixed and real-world data can be found in Figure 6. As the figure shows, the PROSAIL inversion loss landscape continued to show unimodal patterns, with a large skew toward convergence to the same optimum every time. The input-output continuity likewise remained intact, with the exception of real-world data in Figure 6f. This shows that some PROSAIL inversion loss landscapes may sometimes violate requirement 3 of well-posedness for real-world spectral data, although there would still only be a single



(a) Noisy, E1 (b) Noisy, E2



(c) Spectrally mixed, E1 (d) Spectrally mixed, E2



(e) Real-world, E1 (f) Real-world, E2

Figure 6.: Repeat of Experiments 1 and 2 for noisy (6a and 6b), spectrally mixed (6c and 6d), and real-world (6e and 6f) data.

optimum in this case (though it would be harder to find). Provided the optimisation algorithm can still reliably converge to the optimum (as in our results in Figure 6e), this would not have a large impact on retrieval performance. It is also worth noting is that the normalised error rates for no perturbations (0%) were often larger than 0. This suggests that the best fitting simulated spectrum did not perfectly match the observed real-world spectrum, thereby potentially resulting in a violation of property 1 of well-posedness (a solution exists) for the parameter retrieval problem (but not PROSAIL inversion).

3.2. RQ2: Ill-posedness causes

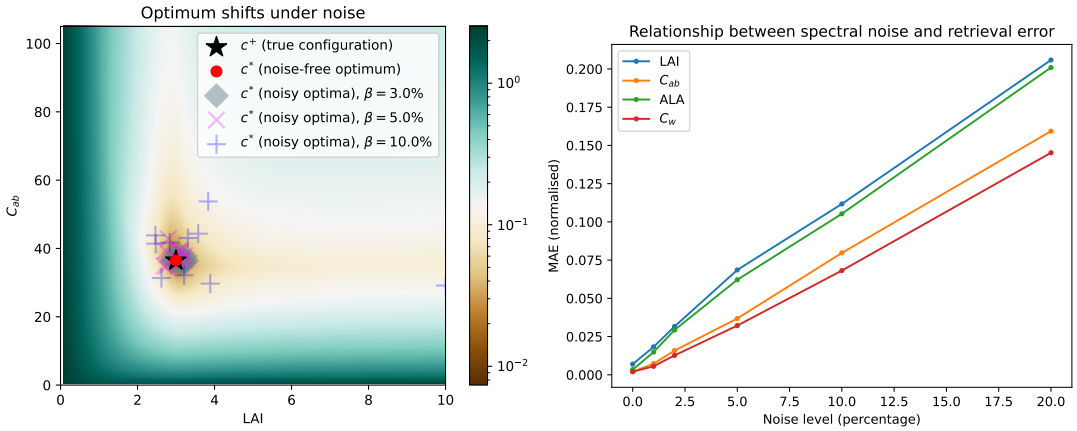
Given our findings that the PROSAIL well-posedness appeared to hold for noisy, spectrally mixed and real-world data, we will test for the impact of our hypothesised causes of parameter retrieval ill-posedness in the following.

Experiment 3. For Experiment 3, we visualised the impact of random noise in the spectral inputs on retrieval performance in Figure 7, with an example loss landscape visualised in Figure 7a. A prediction error on the optimum due to noise on the observations can be interpreted as a 'shift' of the optimum $\mathbf{c}^*$ in the loss landscape, away from the unknown true configuration $\mathbf{c}^+$, represented by the grey, magenta and blue markers in Figure 7a. As illustrated in the figure, each time a Gaussian noise was re-applied to an originally 'clean' spectrum (as described in Section 2.3), a new point in the parameter space globally minimised the spectral loss, for which we visualised 10 examples per noise level in the figure. The shifts were fairly small for 3% spectral noise, noticeable for 5%, and highly disruptive with large outliers (e.g., reaching the maximum LAI value of 10) for 10% noise – a realistic setting, as the Sen2Cor atmospheric correction algorithm alone can introduce substantial noise to a spectrum (de Sa et al., 2021; Sola et al., 2018; Verrelst et al., 2019b). For example, the RMSE for Sen2Cor per band reported by Sola et al. (2018) represents around 14% to 20% of the average band values in our real-world dataset, with one outlier at nearly 40%. Because, for any given result, we cannot know the degree to which this optimum has been noise-shifted away from $\mathbf{c}^+$, nor the direction in which it was shifted, any point on the loss landscape that the optimum $\mathbf{c}^*$ could have shifted from, could be considered a potential solution to the retrieval problem, thereby violating requirement ii) of well-posedness.

In Figure 7b, we aggregated the numerical results and plotted the mean absolute error for retrieval (normalised as a proportion of the total possible range of the parameter, e.g., 0-10 for LAI) against the intensity of Gaussian noise added to the spectral observations. As might be expected, the mean absolute error increased with the noise added to the spectra; this is consistent with results reported by de Sa et al. (2021). In the results for PMAE shown in Figure 7b, the normalised loss for all parameters showed a linear relationship with the noise level. In the SAM results, which can be found in Appendix D.2, the patterns for ALA and LAI parameters were not linear, but instead sharply increased for low levels of noise, while increasing only marginally for higher noise levels. However, in all cases higher levels of noise resulted in higher parameter retrieval error rates.

We conclude that spectral noise seems to be a contributing factor to the ill-posedness of parameter retrieval using spectral information, which can be the case even when PROSAIL inversion itself is well-posed. This effect is likely exacerbated by the ill-conditionedness found in Figure 3: relatively small perturbations on the spectral observations caused by noise could result in large jumps across the loss landscape, given the large plateau-like region of the loss landscape where noise could relatively easily overpower the signal of the loss function.

Experiment 4. In this experiment, we tested whether spectral mixing could be an additional explanation for the ill-posedness experienced when performing parameter retrieval. The results for this experiment can be found in Table 3. The cells in this table contain normalised mean absolute error rates for different parameters (rows), when the predictions in $\mathbf{c}^*$ are compared to 4 different types of 'true' values (columns): the weighted mean of the configurations that correspond to the 3 mixed spectra that formed the observations, and the parameter values of these 3 configurations them-



(a) Example visualisation

(b) Performance per noise level

Figure 7.: The impact of spectral noise on retrieval performance. Figure 7a shows an example of a single instance (with only the LAI and C_{ab} parameters free), with colours representing its loss values, forming a loss landscape. The true configuration c^+ is visualised as a star, which the noise-free optimum c^* (the red dot) matches nearly perfectly. The grey, magenta and blue markers represent the shifted versions of the optimum c^*, when Gaussian spectral noise at an intensity β of 3%, 5% and 10% is applied 10 times to the same instance (every repeat sampling a new, unpredictable noise term). Figure 7b aggregates this 'shifted optimum' phenomenon over all 1000 instances, showing that it is a consistent pattern, and the intensity of the shifts increases as the noise level increases.

selves.

As Table 3 shows, there was a notable drop in retrieval performance, especially for LAI, compared to the performance expected for noise-free data that was not mixed. For example, the loss values for the weighted mean of the mixed configurations in Table 3 appear comparable to the loss values for spectra with 2.5% to 10% Gaussian noise applied to them in Figure 7b, while the expected loss for non-mixed noise-free spectra based on this figure is close to 0. This decreased performance suggests that the optimum for a mixed spectrum, while consistently converging to a target value most closely related to the weighted mean of the optima of the constituent spectra, and thereby retaining some fidelity, does not perfectly align with such a target. The behaviour of the optimal outcomes c^* in response to a linear mixture of input spectra appears to be governed by complex, non-linear and unpredictable mechanisms. If a user is interested in retrieving, e.g., the weighted mean of the true parameter configurations, there are many points around the optimum that could correspond to this desired value. Therefore, the ill-posedness caused by spectral mixing appears similar to that of random Gaussian noise in E3, as long as the underlying non-linear mechanics remain unpredictable.

Although the magnitude of this effect was relatively small for an extreme type of spectral mixing (fully independently generated configurations), these results suggest that spectral mixing is a contributing factor to ill-posedness for parameter retrieval by violating at least one of characteristics 1 and 2 of well-posedness.

In conclusion for RQ2, we consider the impact of spectral noise to be the most likely cause of the ill-posedness experienced in parameter retrieval from multispectral

Table 3.: Results for E4 on the impact of spectral mixing. Every cell represents the (normalised) MAE between the optima found for the mixed spectrum S' and the quantities listed in the columns. The first column represents the weighted mean of the true configurations of the constituent spectra S_1, S_2 and S_3, while the other columns represent the MAE compared to these individual constituent configurations. This suggests that the solution for mixed spectra matches the weighted mean of the constituent configurations more closely than the configuration of any individual constituent spectrum.

	Normalised MAE target			
Parameter	$\alpha_1\mathbf{c_1^+} + \alpha_2\mathbf{c_2^+} + \alpha_3\mathbf{c_3^+}$	$\mathbf{c_1^+}$	$\mathbf{c_2^+}$	$\mathbf{c_3^+}$
LAI	0.117 ± 0.089	0.164 ± 0.151	0.158 ± 0.149	0.161 ± 0.151
C_{ab}	0.048 ± 0.048	0.125 ± 0.106	0.12 ± 0.101	0.121 ± 0.104
ALA	0.106 ± 0.134	0.236 ± 0.193	0.237 ± 0.196	0.238 ± 0.187
C_w	0.033 ± 0.045	0.077 ± 0.064	0.079 ± 0.065	0.083 ± 0.067

Table 4.: Mean absolute error rates for parameter retrieval performance for the four different parameters (rows), with columns representing the interval size of a range constraint prior on LAI (with 100% covering the full original parameter range). The 'LAI (uniform)' row represents the performance of estimating LAI through uniform random sampling, while in other columns, performance is acquired through optimisation. In each row, the prior range size in a column marked with a lower number (e.g., [1]) retrieves a parameter significantly better (significance level $\alpha = 0.05$) than one with a higher number (e.g., [2]). Adding range constraint priors on LAI greatly improved LAI retrieval performance, while also improving ALA (but not C_{ab} and C_w) retrieval performance.

Range interval	**LAI prior range constraint size**				
	0%	**10%**	**30%**	**50%**	**100%**
LAI (uniform)	[1]0.0 ± 0.0	[2]0.473 ± 0.289	[3]1.413 ± 0.863	[4]2.214 ± 1.37	[5]2.978 ± 2.124
LAI	[1]0.0 ± 0.0	[2]0.479 ± 0.341	[3]0.808 ± 0.835	[4]0.955 ± 1.133	[4]0.972 ± 1.167
C_{ab}	[1]7.737 ± 11.865	[2]8.143 ± 11.575	[3]8.262 ± 11.559	[2]8.235 ± 11.358	[3]8.23 ± 11.401
ALA	[1]5.978 ± 5.95	[2]8.151 ± 9.201	[3]8.795 ± 9.488	[4]8.943 ± 9.685	[3]8.868 ± 9.373
C_w	[1]0.003 ± 0.006	[2]0.004 ± 0.006	[3]0.004 ± 0.005	[4]0.004 ± 0.005	[3]0.004 ± 0.005

data, with a possible additional contribution by spectral mixing.

3.3. RQ3: Impact of range constraint priors

Experiment 5. In this experiment, we were interested whether prior information in the form of range constraint priors can indeed reduce ill-posedness, and if so, through what mechanisms it is effective. Given our results for Experiment 1 in Figure 4, showing PROSAIL inversion to be a unimodal problem, mechanism i) (excluding competing optima) is unlikely to be a big factor. The results for mechanisms ii) (reducing the maximal magnitude of errors) and iii) (parameter dependency) are shown in Table 4.

As the table shows, the retrieval error for LAI increased as the size of the prior range size increased. For prior range sizes up to 10% the performance was highly similar to that of uniform sampling within the range, indicating that the posterior after performing retrieval was equal to the prior knowledge. The range of ill-posed solutions likely extended beyond this prior range; therefore, introducing a range con-

straint reduced the ill-posedness. In this case, performing retrieval would not add any additional information. However, for larger range constraints, performing retrieval resulted in better performance than uniform sampling in the range interval, while the performance also deteriorated with larger intervals. These results indicate that mechanism ii) is a likely factor in the efficacy of range constraint priors for improving ill-posed retrieval performance, usually without the prior knowledge replacing the retrieval method outright.

For mechanism iii), as Table 4 shows, the retrieval performance of C_{ab} and C_w did not appear to be strongly affected by the LAI prior range interval size, while the retrieval performance of ALA appears to be correlated due to its errors increasing with the increased prior range interval, though this was mainly the case for highly precise ($< 10\%$) intervals. Therefore, while it is likely that mechanism iii) plays some part in improving retrieval performance, its impact may be limited.

We conclude for RQ3 that range constraint priors help improve retrieval performance primarily through mechanism ii), but mechanism iii) may also contribute in some cases.

4. Discussion

In this work, we set out to analyse the ill-posedness of PROSAIL inversion (RQ1), to establish possible causes for the ill-posedness experienced by domain practitioners (RQ2), and to confirm that a commonly used strategy for reducing ill-posedness, adding range constraint priors, indeed reduces the ill-posedness of the problem (RQ3).

Our results for **Experiment 1** (Figure 4) and **Experiment 2** (Figure 5) show that PROSAIL inversion itself is unlikely to be ill-posed, since it meets the criteria of a well-posed problem. Our analysis focused on the retrieval of LAI, leaf angle, chlorophyll content and water content; however, it is possible that the retrieval of other parameters, notably parameters with a limited impact on the spectral output of PROSAIL, could still be ill-posed.

Our results for **Experiment 3** (Figure 7) indicate that spectral noise can cause the predicted configuration for a problem instance to shift away from the true configuration. Since this noise is unknown *a priori*, any configuration in the parameter space that a prediction for an observed spectrum could have shifted from is a potential true solution to the problem instance, thereby making the retrieval problem ill-posed. Spectral mixing also appeared to contribute toward ill-posedness (**Experiment 4**; Table 3).

Given these results, it appears that the ill-posedness experienced by domain practitioners performing parameter retrieval does not stem from the PROSAIL radiative transfer model, but rather from the inherent limitations of the overarching retrieval task and the information contained in the spectral data. This would carry strong implications for future work in this field, as multispectral data would not contain sufficient information to reliably retrieve the parameters of interest.

Finally, our results for **Experiment 5** indicate that the use of range constraints improved retrieval performance more for smaller range intervals, and less for larger intervals, while performing better than random sampling in the range interval after interval sizes of 10%. This indicates that mechanism ii) (reduced magnitudes of errors) is likely to be a factor.

In contrast, mechanism iii) (inter-parameter dependencies) likely has a modest contribution to the efficacy of range constraint priors to reduce ill-posedness. This is a

surprising result, given that biophysical parameters are highly likely to affect one another in nature. A possible explanation for this result may be that models like PROSAIL can perform a simulation for any configuration, regardless of its biological plausibility, and the relationship between parameters (like the example in Figure 2c) depends solely on whether they amplify one another's effects on the spectrum. This is not necessarily related to the degree of co-occurrence of certain parameter settings in nature. In fact, such natural relationships may be promising candidates to further constrain the search space.

4.1. Future work

There are several remaining challenges in understanding the ill-posedness of the parameter retrieval problem. First, performing analyses on labelled real-world data could evaluate whether fully data-driven approaches, not reliant on PROSAIL, would indeed be susceptible to the same ill-posedness, as our experimental results suggest. The data collection involved would be a major obstacle to testing this hypothesis in practice, as it would require extremely accurate ground-truth data for all parameters simultaneously, as well as effectively noise-free spectral observations to estimate the impact of noise. Future work in this direction, if the data requirements were sufficiently met for such a study to become feasible, would be valuable.

Though we were able to test two hypothesised causes of ill-posedness and mechanisms for the impact of range constraint priors, we cannot be certain that our list of hypotheses is exhaustive. It is possible that there are other factors involved that play even larger roles. Therefore, beyond the results of this study itself, we stress in particular the perspective we adopted to test our hypotheses through a careful consideration of the loss landscapes underlying the inversion problem. We invite other scholarly work, should others come up with a set of additional hypotheses, to similarly test these through systematic computational experiments on the loss landscape where appropriate.

Our experimental results suggest that ill-posedness could not be overcome solely with algorithmic contributions that improve the identification of the parameter configuration that best fits the observations (e.g., through better optimisation or by performing more efficient learning). Even if the optimum were predicted perfectly accurately every time, our findings indicate that this optimum itself is not a perfect target. Statistical uncertainty quantification techniques may help capture these uncertainties, although these should be selected with care. Since the ill-posedness appears to originate in the parameter retrieval problem itself, rather than any particular property of the PROSAIL model and its inversion, it is likely that purely data-driven methods would be similarly affected by the ill-posedness.

Instead, we would encourage explorations into novel contributions for biophysical parameter retrieval from a data-centric perspective, focusing on increasing the information content in the observed data (for example, through incorporating additional data sources, higher resolution data, hyperspectral data, or temporal autocorrelation), regardless of whether the method used for mapping the observations to estimated parameters includes PROSAIL inversion or not. For example, the use of specialist hybrid models, that have been trained on a subset of data that is relevant to a study area (either through manual selection of training data, or through the use of active learning heuristics combined with a study area-specific validation set) has already seen successful applications of PROSAIL inversion (Candiani et al., 2022; Sahoo et al., 2023).

Perhaps an automatic generation or selection of appropriate specialist models might be a fruitful next step to explore. Alternatively, advances in deep learning techniques may be able to contribute toward 'denoising' the observed noisy spectra through the application of, e.g., denoising autoencoders (Vincent et al., 2008). If noisy spectral data could be reduced to the equivalent of the noise-free simulated data, this would greatly alleviate the ill-posedness caused by noise.

5. Conclusions

Our systematic analysis of PROSAIL inversion and its ill-posedness indicates that pure PROSAIL inversion for the retrieval of the often studied parameters leaf area index, chlorophyll content, average leaf angle and water content, meets all the requirements of a well-posed problem: there is a solution, the solution is unique, and the output moves continuously with respect to the input. Given the ill-posedness experienced by domain practitioners, these results suggest that PROSAIL inversion itself is not an ill-posed problem, but rather the associated parameter estimation problem when relying on multispectral data. This seems mainly caused by spectral noise, while spectral mixing also appears to play a role.

Finally, we found that range constraint priors can alleviate the ill-posedness of the retrieval problem through a reduction of the magnitude of possible errors, in addition to possible indirect effects through inter-parameter dependencies.

A problem can only be effectively addressed if it is well understood. Currently, much focus in parameter retrieval work using PROSAIL inversion has focused on method-centric approaches, for example, by efficiently sampling points using active learning, or on training specialised hybrid models with reduced ranges for specific study sites. We hope that, building on the results reported here, future work can also more efficiently explore novel solutions that might improve the viability of parameter retrieval when fewer assumptions (e.g., about the study site) can be made. In particular, we believe that the exploration of data-centric improvements, such as the automatic training or selection of specialist models for a given application area, spectral denoising or data fusion approaches, may be a fruitful endeavour.

Acknowledgements

This work was supported by the Dutch Research Council (NWO) under grant OCENW.KLEIN.425; the European Space Agency (ESA) under No. 4000136204/21/NL/GLC/my; TAILOR (EU Horizon 2020 research and innovation programme) under GA No 952215; and by an Alexander von Humboldt Professorship in AI held by Holger Hoos.

Disclosure of interest

The authors report there are no competing interests to declare.

Data availability

The code required to reproduce our results can be found at the following repository: `https://github.com/ADA-research/PROSAIL-ill-posedness`. The SEN2-MSI-T dataset specification and download scripts are available at `https://github.com/ADA-research/VPint2`. All other data was generated using the PROSAIL radiative transfer model.

References

L. Arp, H. Hoos, P. van Bodegom, A. Francis, J. Wheeler, D. van Laar, and M. Baratchi. Training-free thick cloud removal for Sentinel-2 imagery using value propagation interpolation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 216:168–184, 2024. ISSN 0924-2716.

T. Bäck. *Evolutionary computation 1: Basic algorithms and operators*. CRC press, 2018.

C. Bacour, S. Jacquemoud, M. Leroy, O. Hautecœur, M. Weiss, L. Prévot, N. Bruguier, and H. Chauki. Reliability of the estimation of vegetation characteristics by inversion of three canopy reflectance models on airborne POLDER data. *Agronomie*, 22(6):555–565, 2002.

C. Bacour, F. Baret, D. Béal, M. Weiss, and K. Pavageau. Neural network estimation of LAI, fAPAR, fCover and LAI× Cab, from top of canopy MERIS reflectance data: Principles and validation. *Remote Sensing of Environment*, 105(4):313–325, 2006.

G. Bai, J. Dash, L. Brown, C. Meier, C. Lerebourg, E. Ronco, N. Lamquin, V. Bruniquel, M. Clerici, and N. Gobron. Gbov (ground-based observation for validation): a Copernicus service for validation of vegetation land products. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 4592–4594. IEEE, 2019.

D. A. Belsley and R. Oldford. The general problem of ill conditioning and its role in statistical analysis. *Computational Statistics & Data Analysis*, 4(2):103–120, 1986.

K. Berger, J. P. Rivera Caicedo, L. Martino, M. Wocher, T. Hank, and J. Verrelst. A survey of active learning for quantifying vegetation traits from terrestrial Earth observation data. *Remote Sensing*, 13(2):287, 2021.

N. A. Binh, L. T. Hauser, P. Viet Hoa, G. Thi Phuong Thao, N. N. An, H. S. Nhut, T. A. Phuong, and J. Verrelst. Quantifying mangrove leaf area index from Sentinel-2 imagery using hybrid models and active learning. *International Journal of Remote Sensing*, 43(15-16):5636–5657, 2022.

L. A. Brown, C. Meier, H. Morris, J. Pastor-Guzman, G. Bai, C. Lerebourg, N. Gobron, C. Lanconelli, M. Clerici, and J. Dash. Evaluation of global leaf area index and fraction of absorbed photosynthetically active radiation products over north america using copernicus ground based observations for validation data. *Remote Sensing of Environment*, 247:111935, 2020.

G. Caballero, A. Pezzola, C. Winschel, A. Casella, P. Sanchez Angonova, J. P. Rivera-Caicedo, K. Berger, J. Verrelst, and J. Delegido. Seasonal mapping of irrigated winter wheat traits in Argentina with a hybrid retrieval workflow using Sentinel-2 imagery. *Remote Sensing*, 14 (18):4531, 2022.

G. Candiani, G. Tagliabue, C. Panigada, J. Verrelst, V. Picchi, J. P. Rivera Caicedo, and M. Boschetti. Evaluation of hybrid models to estimate chlorophyll and nitrogen content of maize crops in the framework of the future CHIME mission. *Remote Sensing*, 14(8):1792, 2022.

R. M. Cavalli. Spatial validation of spectral unmixing results: A systematic review. *Remote Sensing*, 15(11):2822, 2023.

A. Chakraborty and A. K. Kar. Swarm intelligence: A review of algorithms. *Nature-inspired computing and optimization: Theory and applications*, pages 475–494, 2017.

B. Combal, F. Baret, M. Weiss, A. Trubuil, D. Macé, A. Pragnere, R. Myneni, Y. Knyazikhin,

and L. Wang. Retrieval of canopy biophysical variables from bidirectional reflectance: Using prior information to solve the ill-posed inverse problem. *Remote Sensing of Environment*, 84(1):1–15, 2003.

R. Darvishzadeh, A. Skidmore, M. Schlerf, and C. Atzberger. Inversion of a radiative transfer model for estimating vegetation LAI and chlorophyll in a heterogeneous grassland. *Remote Sensing of Environment*, 112(5):2592–2604, May 2008. ISSN 0034-4257. . URL `https://www.sciencedirect.com/science/article/pii/S0034425707004968`.

N. C. de Sa, M. Baratchi, L. T. Hauser, and P. van Bodegom. Exploring the impact of noise on hybrid inversion of PROSAIL RTM on Sentinel-2 data. *Remote Sensing*, 13(4):648, 2021.

A. Guo, W. Huang, B. Qian, H. Ye, Q. Jiao, X. Cheng, and C. Ruan. A hybrid model coupling PROSAIL and continuous wavelet transform based on multi-angle hyperspectral data improves maize chlorophyll retrieval. *International Journal of Applied Earth Observation and Geoinformation*, 132:104076, 2024.

H. H. Hoos and T. Stützle. Stochastic local search. In *Handbook of Approximation Algorithms and Metaheuristics*, pages 297–307. Chapman and Hall/CRC, 2018.

F. Hutter, H. Hoos, and K. Leyton-Brown. An efficient approach for assessing hyperparameter importance. In *International conference on machine learning*, pages 754–762. PMLR, 2014.

S. Jacquemoud and F. Baret. PROSPECT: A model of leaf optical properties spectra. *Remote Sensing of Environment*, 34(2):75–91, 1990.

S. Jacquemoud, S. Flasse, J. Verdebout, and G. Schmuck. Comparison of several optimization methods to extract canopy biophysical parameters-application to CAESAR data. In *Proc. 6th Int. Symp. Physical Measurements and Signatures in Remote Sensing*, pages 291–298. CNES Paris, 1994.

S. Jacquemoud, F. Baret, B. Andrieu, F. Danson, and K. Jaggard. Extraction of vegetation biophysical parameters by inversion of the PROSPECT+ SAIL models on sugar beet canopy reflectance data. Application to TM and AVIRIS sensors. *Remote Sensing of Environment*, 52(3):163–172, 1995.

S. Jacquemoud, C. Bacour, H. Poilve, and J.-P. Frangi. Comparison of four radiative transfer models to simulate plant canopies reflectance: Direct and inverse mode. *Remote Sensing of Environment*, 74(3):471–481, 2000.

S. Jacquemoud, W. Verhoef, F. Baret, C. Bacour, P. J. Zarco-Tejada, G. P. Asner, C. François, and S. L. Ustin. PROSPECT+ SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment*, 113:S56–S66, 2009.

S. I. Kabanikhin. Definitions and examples of inverse and ill-posed problems. *Journal of Inverse and Ill-posed Problems*, 2008.

K. Kowalski, A. Okujeni, and P. Hostert. A generalized framework for drought monitoring across central European grassland gradients with Sentinel-2 time series. *Remote Sensing of Environment*, 286:113449, 2023.

Liangyun, Liu and Bowen, Song. ValLAI_Crop: Validation dataset for coarse-resolution satellite LAI product over Chinese cropland, 2021. URL `https://data.niaid.nih.gov/resources?id=zenodo_4080910`. Last accessed: 20-03-2025.

M. Meroni, R. Colombo, and C. Panigada. Inversion of a radiative transfer model with hyperspectral observations for LAI mapping in poplar plantations. *Remote Sensing of Environment*, 92(2):195–206, 2004. ISSN 0034-4257.

NASA. SMAPVEX08 In Situ Vegetation Data, 2008. URL `https://cmr.earthdata.nasa.gov/search/concepts/C1000001420-NSIDC_ECS.html`. Last accessed: 20-03-2025.

National Ecological Observatory Network (NEON). NEON). URL `https://data.neonscience.org`. Last accessed: 20-03-2025.

X. Quan, B. He, and X. Li. A Bayesian network-based method to alleviate the ill-posed inverse problem: A case study on leaf area index and canopy water content retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 53(12):6507–6517, 2015.

M. Ranghetti, M. Boschetti, L. Ranghetti, G. Tagliabue, C. Panigada, M. Gianinetto, J. Verrelst, and G. Candiani. Assessment of maize nitrogen uptake from PRISMA hyperspectral data through hybrid modelling. *European Journal of Remote Sensing*, 56(1):2117650, 2023.

K. Richter, C. Atzberger, F. Vuolo, P. Weihs, and G. d'Urso. Experimental assessment of the Sentinel-2 band setting for RTM-based LAI retrieval of sugar beet and maize. *Canadian Journal of Remote Sensing*, 35(3):230–247, 2009.

M. Rossi, G. Candiani, F. Nutini, M. Gianinetto, and M. Boschetti. Sentinel-2 estimation of CNC and LAI in rice cropping system through hybrid approach modelling. *European Journal of Remote Sensing*, pages 1–20, 2022.

R. N. Sahoo, S. Gakhar, R. G. Rejith, J. Verrelst, R. Ranjan, T. Kondraju, M. C. Meena, J. Mukherjee, A. Daas, S. Kumar, et al. Optimizing the retrieval of wheat crop traits from uav-borne hyperspectral image with radiative transfer modelling using Gaussian process regression. *Remote Sensing*, 15(23):5496, 2023.

I. M. Sobol. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, 55(1-3):271–280, 2001.

I. Sola, A. García-Martín, L. Sandonís-Pozo, J. Álvarez-Mozos, F. Pérez-Cabello, M. González-Audícana, and R. M. Llovería. Assessment of atmospheric correction methods for Sentinel-2 images in Mediterranean landscapes. *International journal of applied earth observation and geoinformation*, 73:63–76, 2018.

J. Sun, L. Wang, S. Shi, Z. Li, J. Yang, W. Gong, S. Wang, and T. Tagesson. Leaf pigment retrieval using the PROSAIL model: Influence of uncertainty in prior canopy-structure information. *The Crop Journal*, 10(5):1251–1263, 2022.

M. Varah, James. On the numerical solution of ill-conditioned linear systems with applications to ill-posed problems. *SIAM Journal on Numerical Analysis*, 10(2):257–267, 1973.

W. Verhoef. Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. *Remote Sensing of Environment*, 16(2):125–141, 1984.

J. Verrelst, J. P. Rivera, G. Leonenko, L. Alonso, and J. Moreno. Optimizing LUT-Based RTM Inversion for Semiautomatic Mapping of Crop Biophysical Parameters from Sentinel-2 and -3 Data: Role of Cost Functions. *IEEE Transactions on Geoscience and Remote Sensing*, 52(1):257–269, Jan. 2014. ISSN 1558-0644. . Conference Name: IEEE Transactions on Geoscience and Remote Sensing.

J. Verrelst, Z. Malenovskỳ, C. Van der Tol, G. Camps-Valls, J.-P. Gastellu-Etchegorry, P. Lewis, P. North, and J. Moreno. Quantifying vegetation biophysical variables from imaging spectroscopy data: A review on retrieval methods. *Surveys in Geophysics*, 40:589–629, 2019a.

J. Verrelst, J. Vicent, J. P. Rivera-Caicedo, M. Lumbierres, P. Morcillo-Pallarés, and J. Moreno. Global sensitivity analysis of leaf-canopy-atmosphere RTMs: Implications for biophysical variables retrieval from top-of-atmosphere radiance data. *Remote sensing*, 11(16):1923, 2019b.

P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th International Conference on Machine Learning*, pages 1096–1103, 2008.

J. Wang, R. Lopez-Lozano, M. Weiss, S. Buis, W. Li, S. Liu, F. Baret, and J. Zhang. Crop specific inversion of PROSAIL to retrieve green area index (GAI) from several decametric satellites using a Bayesian framework. *Remote Sensing of Environment*, 278:113085, 2022.

F. Xu and B. Somers. Unmixing-based Sentinel-2 downscaling for urban land cover mapping. *ISPRS Journal of Photogrammetry and Remote Sensing*, 171:133–154, 2021.

R. H. Yuhas, A. F. Goetz, and J. W. Boardman. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. In *JPL, Summaries of the Third Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop*, 1992.

Appendix A. Implementation details

A.1. Loss functions

When performing analyses on a loss landscape, the selection of an appropriate loss function (similarity metric d) can be of great importance. In the case of multispectral data, the simplest approach would be to consider all bands independently, and apply the mean absolute error (MAE) (or L_1 loss) function to compare the simulated spectrum $\hat{S}$ and the observed spectrum S as:

$$MAE(\hat{S}, S) = \frac{1}{B} \sum_{b=0}^{B} |\hat{s}_b - s_b| \tag{A1}$$

Here B is the number of bands in the spectrum. However, the values of spectral bands can greatly differ in magnitude, causing the MAE to be biased towards the bands with the greatest expected reflectance values (such as infrared). A relatively simple way to alleviate this problem is to use the proportional mean absolute error (PMAE), which is equivalent to the mean absolute percentage error (MAPE) used in other work (de Sa et al., 2021), but does not convert the representation to percentages:

$$PMAE(\hat{S}, S) = \frac{1}{B} \sum_{b=0}^{B} \frac{|\hat{s}_b - s_b|}{s_b} \tag{A2}$$

MAE and PMAE are intuitive loss functions for the reconstruction of spectral data. However, in some applications, the brightness or albedo of a spectrum overall is less important than the ratio of band values compared to other band values, forming the hue of the spectrum. The spectral angle mapper (SAM) loss function (Yuhas et al., 1992) can be used to try to capture this aspect of a reconstruction, and can be computed as:

$$SAM(\hat{S}, S) = \arccos \frac{\hat{S} \cdot S}{(\hat{S} \cdot \hat{S}) * (S \cdot S)} \cdot \frac{180}{\pi} \tag{A3}$$

A.2. Optimisation procedure

We treat the numerical optimisation procedure as a black-box optimisation problem. Within black-box optimisation algorithms, the tradeoff between *exploration* (identifying promising new parts of the search space) and *exploitation* (improving already known promising solutions until convergence) is often a central concept. In unimodal landscapes, only a single optimum exists, allowing greedy algorithms such as hill climbing or greedy local search (for an overview of stochastic local search methods, see (Hoos and Stützle, 2018)) to quickly converge to a local optimum, committing fully to exploitation. If the landscape is multimodal, there are multiple different local optima, causing greedy algorithms to get stuck in local optima. In this case, algorithms that add more exploration to their optimisation heuristics would be necessary.

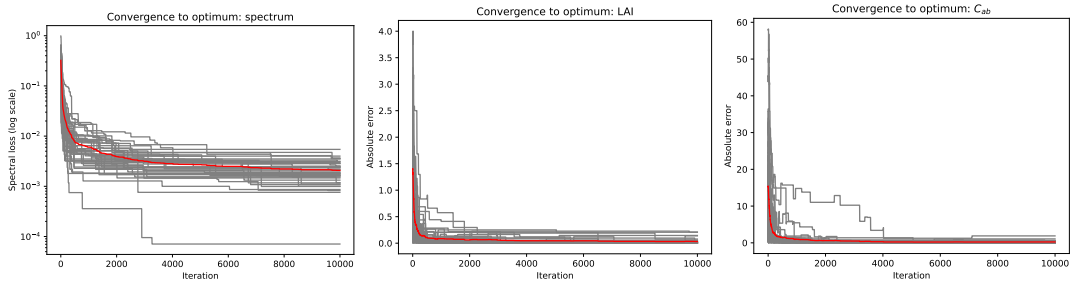
Our experiments used greedy local search as an optimisation algorithm. For every instance, we initialised the current solution c to the mean of the prior distributions of the free parameters. At every optimisation step, we generated a candidate new solution c' by perturbing the elements of c: $c'_p = c_p + \mathcal{N}(0, \sigma_p)$, where p is a parameter

in c and $\sigma_p = 0.05 \cdot \frac{max(p)-min(p)}{2}$. Here σ_p represents the intensity of the perturbation for a parameter p; we set it to 5% of the middle-way point of the parameter range (e.g., if a parameter ranges from 0 to 10, its perturbation intensity would be $0.05 \cdot \frac{10-0}{2} = 0.25$), though other mutation strategies could also be considered. If $\mathcal{L}(M(c'), S) < \mathcal{L}(M(c), S)$, c' becomes the new c. This procedure is iterated until the function evaluation budget has been exhausted.

Using greedy local search resulted in two main advantages. First, convergence will be fast, reducing the computational load of our experiments. Second, using a greedy algorithm allows us to test for multimodality (since the algorithm could converge to different optima when repeating an optimisation procedure), which is important to Section 2.2.

We note that, if the results for our experiments described in Section 2.2 indicate that PROSAIL inversion is a multimodal problem, a global optimisation method would need to be used to enable reliable convergence to a global optimum.

Appendix B. Optimisation convergence



(a) Spectral loss function (b) LAI retrieval loss (c) C_{ab} retrieval loss

Figure B1.: Visualisation of convergence to a stable solution based on the spectral loss function (B1a), the retrieval loss for LAI (B1b), and the retrieval loss for C_{ab} (B1c). In these plots, every grey line represents a randomly sampled instance from D, and the red line represents the mean of these runs. On average (the red lines), the parameters converged to stable values within 1000 iterations, with later iterations only slightly improving the spectral loss further.

We performed this additional experiment to verify that the function evaluation budget used in our experiments is sufficient to converge to a stable optimum, both in terms of the spectral loss value and the parameters of the configuration c^*. We performed our optimisation approach on 100 random instances, and plotted the loss values for the spectrum (which the optimisation algorithm uses to perform optimisation) and for the retrieved parameters in the configuration (which the algorithm does not have access to, and is plotted for evaluation purposes). A very low rate of improvement for later iterations would indicate that the budget is sufficient to converge to a stable optimum.

The results of this experiment can be found in Figure B1. As the figure shows, the function evaluation budget allotted to the optimisation algorithm is sufficient to converge to stable values on average, and a convergence of the spectral loss (Figure B1a) corresponds to a convergence of the retrieval losses of individual parameters

(Figures B1b and B1c). As a result, our experimental setup appears to be well suited to answer our research questions.

Appendix C. Real-world Sentinel-2 data

Although our main experiments focussed on simulated data, since this allowed us access to both noise-free and noisy data, it is possible that there are other factors, beyond the Gaussian noise and spectral mixing we considered in Section 2.3, that result in ill-posedness for parameter retrieval using real-world data. To ensure that the patterns we found (well-posedness of PROSAIL inversion) hold for real-world data as well, we performed additional analyses on real-world data, as shown in Figure 6.

The Sentinel-2 dataset in question that we used was the SEN2-MSI-T cloud removal dataset (Arp et al., 2024). While any Sentinel-2 Level-2A-based dataset would work, this dataset offered a convenient mix of scale and diversity. The dataset consists of 5 land cover classes, each split into a geographically diverse set of 4 scenes, resulting in 20 total locations. Every scene contained a cloud-free observation at 5 time steps within a period of 6 months, resulting in a total of 100 geospatially- and temporally diverse images. From each image, we took the centre pixel as the representative spectrum for the image (using multiple pixels from the same image could have biased the evaluation to the specific images of the dataset). Therefore, this dataset enabled us to evaluate on a curated, diverse set of 100 real-world instances and check if the PROSAIL inversion loss landscapes were still well-posed.

Appendix D. Additional results

D.1. Parameter importance per band

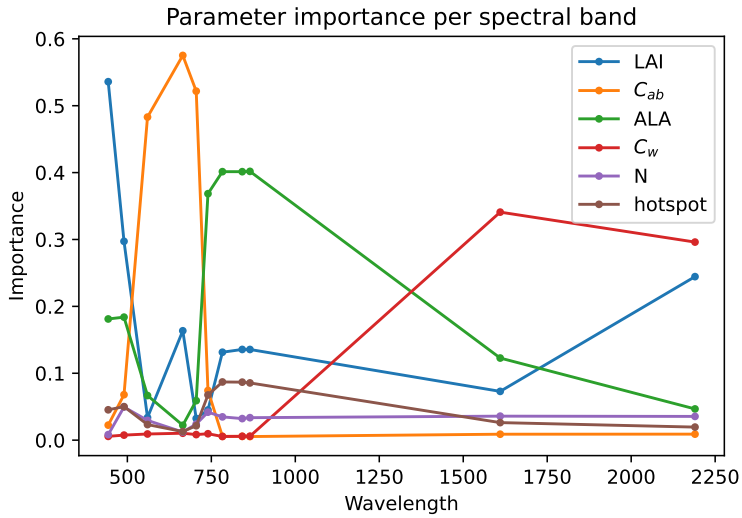
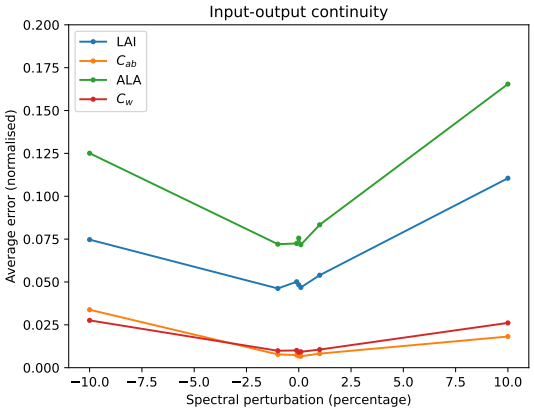


Figure D1.: Relative importance of different parameters for different wavelengths. Every point on a line represents the average importance of the parameter of that line in determining the MAE for a particular spectral band's wavelength.

In addition to our main parameter importance experiment, we computed the importance of 6 of the globally most important parameters at every Sentinel-2 band, and plotted this is Figure D1. As can be seen in the figure, the relative importance of the parameters can vary greatly between spectral bands. Therefore, parameters with a relatively low global importance may still be relatively easily retrievable, due to the sensitivity of the loss landscape to these parameters in local, specialised parts of the parameter space. All parameters from our selection have a part of the spectrum where they have the highest impact (LAI for ultraviolet and blue, chlorophyll for green, red and near-infrared, leaf angle for short-wave infrared, and leaf water content for infrared).

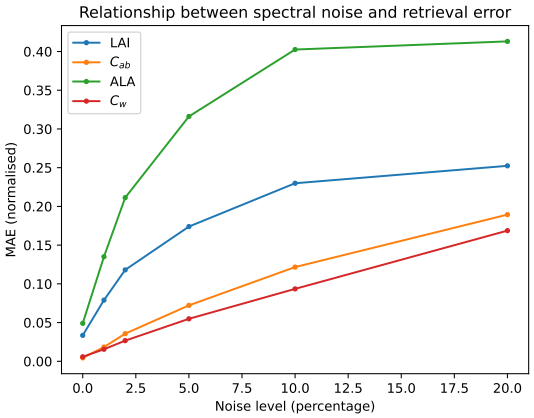
D.2. SAM main experiment results

These supplementary figures contain the results for our experiments for the SAM loss function, corresponding to the same analyses we provide in Section 3 for the PMAE loss function. We have moved these figures to the supplementary material because their patterns largely conformed to those for PMAE. Figure D2 contains results for Experiment 2, Figure D3 contains results for Experiment 3, Table D1 contains results for Experiment 4, and Table D2 contains results for Experiment 5.



(a) Shift per perturbation

Figure D2.: The continuity of the output for PROSAIL inversion (predicted configuration c^*) with respect to perturbations to the input (spectrum), aggregated over all 1000 instances, normalised to a 0-1 range based on the bounds of the parameter range, showing that it is a consistent pattern. Unlike in the PMAE results, the best solution for no perturbation (0 on the x-axis) did not have a near-zero error rate for LAI and ALA; this suggests that SAM may not be an appropriate choice as an optimisation loss function.



(a) Performance per noise level

Figure D3.: The impact of spectral noise on retrieval performance, aggregating the 'shifted optimum' phenomenon over all 1000 instances, showing that it is a consistent pattern, and the intensity of the shifts increases as the noise level increases.

Table D1.: Results for E4 on the impact of spectral mixing. Every cell represents the (normalised) MAE between the optima found for the mixed spectrum S' and the quantities listed in the columns. The first column represents the weighted mean of the true configurations of the constituent spectra S_1, S_2 and S_3, while the other columns represent the MAE compared to these individual constituent configurations. This suggests that the solution for mixed spectra matches the weighted mean of the constituent configurations more closely than the configuration of any individual constituent spectrum.

	Normalised MAE target			
Parameter	$\alpha_1\mathbf{c_1^+} + \alpha_2\mathbf{c_2^+} + \alpha_3\mathbf{c_3^+}$	$\mathbf{c_1^+}$	$\mathbf{c_2^+}$	$\mathbf{c_3^+}$
LAI	0.136 ± 0.167	0.199 ± 0.182	0.203 ± 0.187	0.206 ± 0.19
C_{ab}	0.057 ± 0.058	0.128 ± 0.113	0.121 ± 0.109	0.123 ± 0.11
ALA	0.24 ± 0.172	0.296 ± 0.233	0.299 ± 0.241	0.307 ± 0.233
C_w	0.043 ± 0.053	0.082 ± 0.069	0.084 ± 0.072	0.087 ± 0.073

Table D2.: Mean absolute error rates for parameter retrieval performance for the four different parameters (rows), with columns representing the interval size of a range constraint prior on LAI (with 100% covering the full original parameter range). The 'LAI (uniform)' row represents the performance of estimating LAI through uniform random sampling, while in other columns, performance is acquired through optimisation. In each row, the prior range size in a column marked with a lower number (e.g., [1]) retrieves a parameter significantly better (significance level $\alpha = 0.05$) than one with a higher number (e.g., [2]). Adding range constraint priors on LAI greatly improved LAI retrieval performance, while also improving ALA (but not C_{ab} and C_w) retrieval performance.

Range interval	**LAI prior range constraint size**				
	0%	**10%**	**30%**	**50%**	**100%**
LAI (uniform)	[1]0.0 ± 0.0	[2]0.473 ± 0.289	[3]1.413 ± 0.863	[4]2.214 ± 1.37	[5]2.978 ± 2.124
LAI	[1]0.0 ± 0.0	[2]0.755 ± 0.33	[3]1.606 ± 1.054	[4]1.995 ± 1.572	[5]2.151 ± 1.851
C_{ab}	[2]10.86 ± 15.61	[1]10.35 ± 15.497	[3]11.455 ± 15.626	[4]12.084 ± 15.965	[5]12.23 ± 16.092
ALA	[1]24.908 ± 22.832	[2]30.773 ± 22.476	[3]34.7 ± 23.514	[4]35.712 ± 24.424	[4]35.725 ± 24.357
C_w	[1]0.004 ± 0.007	[1]0.005 ± 0.007	[3]0.005 ± 0.008	[4]0.005 ± 0.008	[5]0.005 ± 0.008