Inversion of electromagnetic induction data using a novel wavelet-based and scale-dependent regularization term

Wouter Deleersnyder^{1,2}, Benjamin Maveau¹, Thomas Hermans² and David Dudal^{1,3}

¹Department of Physics, KU Leuven Campus Kortrijk - KULAK, Etienne Sabbelaan 53, 8500 Kortrijk, Belgium. E-mail: wouter.deleersnyder@kuleuven.be

²Department of Geology, Ghent University, Krijgslaan 281 - S8, 9000 Gent, Belgium

³Department of Physics and Astronomy, Ghent University, Ghent, Krijgslaan 281 - S9, 9000 Gent, Belgium

Accepted 2021 May 4. Received 2021 April 2; in original form 2020 August 20

SUMMARY

The inversion of electromagnetic induction data to a conductivity profile is an ill-posed problem. Regularization improves the stability of the inversion and a smoothing constraint is typically used. However, the conductivity profiles are not always expected to be smooth. Here, we develop a new inversion scheme in which we transform the model to the wavelet space and impose a sparsity constraint. This sparsity constrained inversion scheme will minimize an objective function with a least-squares data misfit and a sparsity measure of the model in the wavelet domain. A model transform to the wavelet domain allows to investigate the temporal resolution (periodicities at different frequencies) and spatial resolution (location of the peaks) characteristics of the model, and penalizing small-scale coefficients effectively reduces the complexity of the model. The novel scale-dependent regularization term can be used to favour either blocky or smooth structures, as well as high-amplitude models in globally smooth structures in the inversion. Depending on the expected conductivity profile, a suitable wavelet basis function can be chosen. The scheme supports multiple types of regularization with the same algorithm and is thus flexible. Finally, we apply this new scheme on a frequency domain electromagnetic sounding data set, but the scheme could equally apply to any other 1-D geophysical method.

Key words: Controlled source electromagnetics (CSEM); Electromagnetic theory; Inverse theory; Wavelet transform.

1 INTRODUCTION

Electromagnetic induction (EMI) surveys aim to image the electrical properties of the subsurface non-invasively. Via a petrophysical law, the electrical properties are then related to subsurface characteristics such as the porosity and tortuosity, the electrical conductivity of the pore fluid, the saturation and the clay content (Archie *et al.* 1942; Revil *et al.* 2017; Duy Thanh *et al.* 2019). EMI methods are used in a broad range of applications such as archaeological prospection (Saey *et al.* 2012), soil contamination (Pettersson & Nobes 2003), detection of electric anomalies [such as unexploded ordnance (Fernández *et al.* 2010)], agriculture (Jadoon *et al.* 2015) and saltwater intrusion (Paepen *et al.* 2020).

EMI instrumentation can often be used with different intercoil spacings or heights, resulting in data with a sensitivity varying with depth, allowing to invert the data and obtain a conductivity profile of the subsurface. Geophysical inversion is an ill-posed problem, because the solution is usually non-unique, which is generally cured with regularization techniques. Additionally, a slightly different data set will generally yield an entirely different result. Deterministic methods impose one or more constraints, such as the

smoothness constraint imposed by the traditional Tikhonov regularization (Tikhonov 1943). The smoothness constraint is not always adapted to the subsurface structure (Linde *et al.* 2015). In earlier work on EMI inversion, Tikhonov regularization, truncated singular value decomposition (Deidda *et al.* 2014) or various sparsity based inversion schemes (Tantum *et al.* 2012) have been commonly applied. While many alternative minimum-structure schemes exist, it is not always easy to find the most adapted scheme. One flexible inversion scheme that supports multiple types of minimum-structure constraints could simplify the generation of an ensemble of various inversion models.

The literature about full-waveform inversion is richer and faces similar problems. Full-waveform inversion is concerned with fitting the model parameters, velocities and impedances, to the recorded seismic or ground penetrating radar data. The ill-posedness of the problem similarly requires a suitable regularization scheme in order to obtain a realistic model. The non-uniqueness is tackled typically with Tikhonov regularization (Virieux & Operto 2009). More recently, total variation regularization was applied to this field, including blocky regularization schemes with several focusing functions (Guitton 2012), such as the ℓ_1 -norm.

The high non-linearity of full-waveform inversion results typically in a multimodal objective function (i.e. multiple local minima), for which multiscale approaches tend to be successful (Bunks *et al.* 1995). In EM data inversion, the objective function may also be multimodal, depending on the choice of the forward model operator. In such cases, Bunks *et al.* (1995) suggested the use of the wavelet decomposition as a direction of future work, because it can decompose the problem by scale in an efficient fashion.

Wavelets are commonly used in signal processing (e.g. denoising), image compressing and compressed sensing. In compressed sensing, where band-limited signals can be recovered from samples below the Nyquist frequency limit (Candes & Tao 2005; Candes & Wakin 2008), sparsity promoting inversion is essential. The particular attraction to this relatively new field resulted in a considerable development of sparsity inversion and accompanying optimization algorithms, which was advantageous for geophysical minimum-structure sparsity inversion. Note that minimum structure in a compressed sensing framework would assume that the earth possesses minimum structure in some basis and that we know that basis. However, we do not state to know the basis in which the properties of the subsurface should possess minimum structure in a compressed sensing framework. In this text, with minimum-structure inversion, we aim for the schemes which are in essence extremal problems, provided with a bounding property of the Earth. This means that we filter unrealistic electrical conductivities and that we aim for the model with the least number of features that explains the data equally well than other, more exciting models with more (potentially attractive) features. This is the Occam's inversion approach taken by Constable *et al.* (1987), which was also the basis for smoothness-based regularization.

Seismologists previously applied wavelet theory to inverse problems, for example for reducing the dimensionality for large scale inverse problems (Voronin *et al.* 2015) or as a minimum-structure regularization technique in seismic tomography (Loris *et al.* 2010; Fang & Zhang 2014). Nittinger & Becken (2016) present a 2-D, wavelet-based sparsity inversion scheme for magnetotelluric data with Daubechies *et al.* (2004)'s work on the iterative soft thresholding algorithm for solving objective functions with both an ℓ_1 and ℓ_2 -norm. While this work was confined to smooth wavelets and thus smooth inversion results, in Nittinger & Becken (2018), the work was expanded to an inversion scheme that combines two or more wavelet families. Specifically for frequency domain EMI, the literature about wavelet-based sparsity inversion is limited. Liu *et al.* (2017) present an inversion scheme for FDEM data, based on the discrete wavelet transform in combination with an ℓ_1 -norm, where the resolution is as good as the space domain. In their methodology, a finite difference method was used to solve the forward problem. Due to the slow forward modelling, only about ten iterations could be considered in the inverse problem.

Here, a wavelet-based minimum-structure regularization scheme is developed with a novel regularization term. Our scheme uses a fast, quasi-linear (approximate) forward modelling operator (see Section 2.1). This allows us to compute much more iterations than with finite difference methods (in fact, computational power is not an issue anymore), while using a more accurate forward model than the commonly used linear LIN approximation. There remains a small modelling error due to the approximations in the forward operator. In Section 3.3, we demonstrate that this modelling error and non-linearity pose not much of an issue. Inspired by the technicalities of the wavelet transform (*cf.* Theorem 2.1 in Section 2.3) and the pragmatic idea that small scale details should be penalized more heavily than the coarser main structure of the conductivity

profile, each length scale of the conductivity profile is penalized with a different weight, we refer to this as scale-dependent weights. The wavelet basis determines the type of constraint (blocky, smooth or in between) that is imposed and hence this is a flexible inversion procedure. The wavelets considered in this text are limited to Daubechies wavelets, we review their properties and motivate this choice in Section 2.3. Our results show mainly two advantages. First, the scheme is more flexible as it can recover both blocky and smooth conductivity profiles. Secondly, it can recover high amplitude anomalies in a globally smooth conductivity profile while using a simple inversion scheme with one Daubechies wavelet. In Sections 3.1–3.3, we apply our regularization scheme on synthetic data and compare it with the commonly used Tikhonov regularization as a benchmark and traditional non-scale-dependent wavelet-based regularization schemes. Finally, we apply our scheme to a real data set of the Gontrode forest, Belgium, a context at which a 1-D inversion scheme is realistic (Bobe *et al.* 2020).

2 METHODS

2.1 Forward model

The forward model describes the soil's response $\mathbf{d} \in \mathbb{C}^{n_d \times 1}$ to the magnetic dipole placed at height h_0 above the surface and a soil with specific parameter distribution $\mathbf{m} \in \mathbb{R}^{n_m \times 1}$ with n_m entries, measured at n_d intercoil spacings s . The parameter distribution $\mathbf{m}$ is referred to as model or model parameters and contains the electrical conductivities in Siemens per metre for a specific (usually equidistant) parametrization of the subsurface. The complex part or quadrature-phase of the FDEM subsurface's response $\mathbf{d}$ is the secondary magnetic field in Ampere per metre, resulting from the induced eddy currents. Finding the operator $\mathcal{F}$, such that

$$\mathbf{d} = \mathcal{F}(\mathbf{m}) \quad (1)$$

holds, is called the *forward problem*. Wait (1951) derived such a model, taking into account the loop-loop couplings between eddy currents and the dampening of the electromagnetic fields, for the 1-D approximation (horizontally stratified earth) in the quasi-stationary Maxwell regime. The model involves recursive relations with integration over Bessel functions, which have to be executed numerically. This computationally demanding task slows down an inversion procedure in which the forward problem needs to be calculated at every iteration. Wait later derived a linear model (Wait 1962), which is extensively used in the work of McNeill (1980) and many others (e.g. Simpson *et al.* 2009; De Smedt *et al.* 2011), known as the low induction number (LIN) approximation. The LIN approximation is linear and neglects the couplings and electromagnetic dampening. It is only valid at low induction numbers B which is defined as the intercoil distance s divided by the skin depth δ , that is $B = \frac{\omega \mu_0 \sigma s^2}{2}$, where ω is the angular frequency of the dipole, μ_0 is the magnetic permeability of the vacuum and σ the electric conductivity of the medium. The meaning of sufficiently low induction numbers is that if the skin depth is much larger than the path the electromagnetic field has to traverse, the dampening can be neglected. Indeed, the path length is restricted by the fall-off of the magnetic dipole, which has the same order as the intercoil distance s . A more recently proposed model by Maveau *et al.* (2020), based on the LIN approach, takes into account the electromagnetic dampening and provides accurate results under more lenient conditions, especially in more conductive (e.g. high salinity) backgrounds. We refer to this model as the damped model. Using an approximate forward

operator $\mathcal{F}_{\text{approx}}$ introduces a biased modelling error κ , we write

$$\mathcal{F}(\mathbf{m}) = \mathcal{F}_{\text{approx}}(\mathbf{m}) + \kappa. \quad (2)$$

Another method to solve the forward problem is to use finite elements methods (FEM) or finite volume methods (FVM), such as provided for example in SimPEG (Cockett *et al.* 2015). These methods can yield accurate results for sufficiently fine meshes. For cylindrically symmetric set-ups (i.e. horizontally stratified conductivity profile and a vertical dipole), the computation is relatively fast. In other cases, the computational burden of a detailed mesh is large. By consequence, only a few iterations in the inversion procedure can be considered.

The forward operator plays an important role in the inversion procedure (it will, for example, affect the number of local minima in the objective function). In Appendix A, we describe the cost related to the choice of the forward operator. We provide a summary of the calculations that are involved for each forward model and a comparison of execution times. In this text, we focus on the inversion with wavelets and do not investigate in details the potential multimodality of the objective function. At first, we generate synthetic data and invert that data via the LIN approximation for its computational benefits. Furthermore, this method is sufficient to illustrate the working principle and the proof of concept of the inversion scheme with wavelets. In Section 3.3, however, we do elaborate on the role of the modelling error κ and the forward operator $\mathcal{F}$.

2.2 Inversion approach

Due to the non-linearity of many forward operators, the inverse problem is typically solved as an optimization problem. A suitable parameter distribution of the subsurface is determined by a minimum of an objective function ϕ . The objective function takes the parameters distribution or model parameters as input and expresses how well the parameters fit the data. In minimum structure inversion procedures, an additional measure of model complexity, a regularization or model misfit term, is added to the objective function. The minimum is usually obtained via gradient descent-like methods.

The objective function in minimum structure inversion is a combination of a data misfit functional ϕ_d , a model misfit functional ϕ_m and a regularization parameter λ . The latter balances the relative importance of both misfit terms. The data misfit functional ϕ_d measures how well the model parameters $\mathbf{m}$ fit the data $\mathbf{d}$ and is defined as a least-squares term

$$\phi_d = \frac{1}{2} \|\mathbf{D}_d(\mathbf{d} - \mathcal{F}(\mathbf{m}))\|_2^2, \quad (3)$$

where $\mathbf{D}_d$ is a diagonal data weighting matrix whose elements are usually the reciprocals of the estimated noise standard deviation (Strutz 2010). The data weighting matrix penalizes the data misfit for data points with minimal measurement error less than data points with larger measurement error. In FDEM inversion, the inverse problem is formulated in terms of the magnetic field ratio, that is the secondary magnetic field divided by the primary magnetic field at the location of the receiver. The output of the forward problem (1) is then a dimensionless, real vector $\mathbf{d}$. The model misfit functional is a regularization term that is introduced to handle the ill-posedness of the problem. Indeed, geophysical inversion problems are ill-conditioned and the solution is typically not unique. A traditional method for minimum structure inversion is to apply a smoothing constraint, such as with Tikhonov regularization (Tikhonov 1943). Tikhonov regularization improves the stability of

the inversion, but promotes smooth solutions and hence cannot recover blocky structures. Other regularization terms exist, such as the ℓ_1 -norm inversion in original model space (Farquharson 2007), the minimum gradient support functional (Thibaut *et al.* 2021) or a sequential inversion (Guillemoteau *et al.* 2016) to name a few. There are many alternatives such as covariance-based inversion (Hermans *et al.* 2012).

In this paper, we develop a different approach. Suppose that there exists a basis in which the true model parameters $\mathbf{m}$, known to possess minimum structure, are represented in a sparse fashion in $\mathbf{x} \in \mathbb{R}^{n_x \times 1}$. Then,

$$\mathbf{x} = \mathbf{W}\mathbf{m}, \quad (4)$$

where $\mathbf{W}$ is the basis transformation. Such a basis transformation, in combination with a sparsity promoting measure, yields an appropriate model misfit, because a complex structure in the model parameters $\mathbf{m}$ will have many non-zero entries in the $\mathbf{x}$ -representation and will therefore be heavily penalized by the sparsity promoting measure. Since (Donoho 2006) has shown that the ℓ_1 -norm is a good approximation of the ℓ_0 -norm, the ℓ_1 -norm is a well-known sparsity promoting measure. Finding the solution of such a ℓ_2 - ℓ_1 -minimization problem can be accomplished by reformulating the optimization problem as a linear program (Nocedal & Wright 2006). For linear programs, multiple efficient solution methods are available (Candes *et al.* 2008). In our optimization problem, measurement errors induce noise and therefore basis pursuit denoising (BPDN) is preferred (Chen & Donoho 1994). This will solve the optimization problem faster than an exact linear program. BPDN is a quadratic convex problem and can be solved by many general solvers, such as described in Nocedal & Wright (2006). We will rely on gradient decent-like methods and do not state that our optimization algorithm is the optimal choice, but it performs fine for our purposes. The details are described in Section 2.5. Since the ℓ_1 -norm is not differentiable at zero, we will use the perturbed ℓ_1 -norm measure of Eklblom (Eklblom 1987)

$$\mu_{\text{Eklblom}}(x) = \sqrt{x^2 + \epsilon}, \quad (5)$$

which is very similar to the ℓ_1 -norm for small ϵ and which is also convex. The latter is a welcome property for the optimization problem. The model misfit in terms of $\mathbf{x}$ in our inversion scheme is thus

$$\phi_m(\mathbf{x}) = \sum_j \sqrt{x_j^2 + \epsilon}. \quad (6)$$

In summary, we have

$$\begin{aligned} & \min_{\mathbf{x}} \phi \\ &= \min_{\mathbf{x}} (\phi_d + \lambda \phi_m) \\ &= \min_{\mathbf{x}} \left(\frac{1}{2} \|\mathbf{D}_d(\mathbf{d} - \mathcal{F}_{\text{approx}}(\mathbf{W}^{-1}\mathbf{x}))\|_2^2 + \lambda \sum_{j=1}^{n_x} \sqrt{x_j^2 + \epsilon} \right), \end{aligned} \quad (7)$$

where the regularization parameter λ is still undetermined. A regularization parameter can be too small, that is the case when the data misfit ϕ_d is smaller than the actual noise on the data, that is when we have overfitting. On the other hand, with a too large regularization parameter, the parameter distribution will exhibit too little structure. Hansen (2010) discusses the common automatic methods to estimate an optimal regularization parameter. More recent and sophisticated techniques exist (Valentine & Sambridge 2018), (Lima *et al.* 2019), but we will rely on the L-curve criterion, which

is still commonly used today. This method requires us to solve the inverse problem for multiple regularization parameters and plots the model misfit in terms of the data misfit. It is expected that the outcomes of each inversion will appear in an L-shape: there will be branches in which a change in the regularization parameter will yield a significant change in either the data or model misfit. There will be a corner in which the change in the regularization parameter will yield a roughly equal significant change in both misfits. The corner is usually the point in the (ϕ_d, ϕ_m) -plane closest to the origin and thus with both misfits minimized. The corner is usually more pronounced and easier to detect in log-log plane. In general, there is no guarantee that the curve will exhibit an L-shape. Local corners pose difficulties for the automatic corner detection. We rely on the adaptive pruning algorithm (Hansen *et al.* 2007), because it is a robust method for the appropriate corner selection.

2.3 Wavelet basis

The basis transformation $\mathbf{W}$ transforms a minimum structure model $\mathbf{m}$ to a sparse representation $\mathbf{x}$ in the wavelet domain. The discrete wavelet transform allows to represent the conductivity profile in a sparse fashion, because it has both spatial and temporal resolution (whereas the Fourier transform only has temporal resolution). The nomenclature stems from time-dependent signals, but it is also more generally used. In space, the temporal resolution refers to the spatial frequency or wavenumber. By studying the wavelet transform, we can analyse both the location of the peaks (spatial resolution) and periodicities at different frequencies (temporal resolution). The sparsity of the transform is guaranteed by a theorem presented later in this text. Note that this text is based on the discrete wavelet transform, which should not be confused with the continuous wavelet transform, often used for frequency analysis. Most of the material in this section can be found in Strang & Nguyen (1996) and Mallat (1989).

While the wavelet transform has a complete and independent theory based on the definition of multiresolution analysis (Daubechies 1992), we sketch a more intuitive summary with analogies to Fourier series. In Fourier series, a 2π periodic signal in the continuous time domain is represented in terms of basis functions $\{e^{ikt}\}_{k \in \mathbb{Z}}$ with Fourier coefficients $f_k = \langle e^{ikt}, f \rangle$ (Mallat 1999). Indeed, we write

$$f(t) = \sum_{k \in \mathbb{Z}} f_k e^{ikt}. \quad (9)$$

Wavelet theory constructs a set of basis functions, consisting of *scaling functions* φ and *wavelet functions* ψ . Analogously to Fourier series, the signal is expanded in a set of basis functions. Unlike Fourier series, these basis functions have compact support (i.e. they only exist on a finite interval) and f does not need to be periodic. Finding such bases and determining their properties are the major matter of interest of wavelet theory. Instead of considering sines and cosines at different frequencies, the following translations and dilations of the wavelet and scaling functions span the function space:

$$\varphi_{n,k}(t) = 2^{n/2} \varphi(2^n t - k) \quad \text{or} \quad \psi_{n,k}(t) = 2^{n/2} \psi(2^n t - k), \quad (10)$$

where $n \in \mathbb{Z}$ is the dilation parameter that makes the basis function's compact support larger or smaller (temporal resolution). The k -parameter describes translations along the t -axis (spatial resolution). We refer to ψ_n or φ_n as a set of basis functions at resolution

level n and these sets span the spaces W_n and V_n , respectively. Assume that the bases are orthonormal (this is not required, biorthogonal wavelets can also be used). The scaling function is more fundamental than the wavelet function, the wavelet function at resolution level n is a linear combination of scaling functions at resolution level $n+1$.

As with Fourier series, we can decompose an arbitrary function $f_n(t) \in (W_n \oplus V_n)$ as a series in both scaling and wavelet functions:

$$f_n(t) = \sum_{k \in \mathbb{Z}} v_{nk} \varphi_{nk}(t), \quad f_n(t) = \sum_{k \in \mathbb{Z}} w_{nk} \psi_{nk}(t), \quad (11)$$

where $v_{nk} = \langle \varphi_{nk}, f \rangle$ and $w_{nk} = \langle \psi_{nk}, f \rangle$ are scaling and wavelet coefficients, respectively. For example, a piecewise constant function $f_n(t)$ at regular intervals can be viewed as a sum of exclusively db1 scaling basis functions at a resolution level n and their translations (see Fig. A1 in Appendix B for the shape of the db1 scaling function). In the discrete wavelet transform, the function $f(t)$ and wavelets are discretely sampled. We drop the time-dependence in this notation and we refer to it as a signal instead of a function.

A one level (discrete) *wavelet transform* is the decomposition of the signal $f_n \in V_n$ in its components $f_{n-1} \in V_{n-1}$ and $g_{n-1} \in W_{n-1}$.

$$f_n = \sum_k v_{nk} \varphi_{nk} = \underbrace{\sum_k v_{n-1,k} \varphi_{n-1,k}}_{f_{n-1}} + \underbrace{\sum_k w_{n-1,k} \psi_{n-1,k}}_{g_{n-1}}, \quad (12)$$

where $v_{n-1,k}$ are the scaling coefficients and $w_{n-1,k}$ the wavelet coefficients at resolution level $n-1$. The signal g_{n-1} describes the details that were present in the signal on scale n , but have disappeared from the coarser $(n-1)$ -scale f_{n-1} . A higher level wavelet transform is obtained by repeating the one level wavelet transform on f_{n-1} . This process can be interpreted in the framework of filter banks, where the wavelet coefficients are obtained after passing through a high pass filter (contains details of signal) and a subsequent downsampling procedure. A low pass filter and the subsequent downsampling generate the scaling coefficients. These scaling functions are then again passed through the high and low pass filter, etc. The high and low pass filter coefficients are a result from specific conditions, for example that they allow for perfect reconstruction to the model space, that is such that there exists an inverse procedure to go from the coefficients in wavelet space to the original signal. For every wavelet basis function, there exists a unique set of low pass and high pass filters. Their relation is not obvious, for which we refer to Mallat (1999). However, the filter bank interpretation allows for a fast and reliable computation of the wavelet transform with complexity $\mathcal{O}(N)$, in which the explicit form of the basis function is not required. This computation, known as the fast wavelet transform (FWT), is analogous to the fast Fourier transform.

In our inversion approach, we require an explicit matrix $\mathbf{W} \in \mathbb{R}^{n_x \times n_m}$ for the basis transformation. This is obtained by computing n_m wavelet transforms of a Dirac train (a Dirac train δ_i is a vector with zeros, and with a spike (a one) at the i th index)

$$\mathbf{W} = \mathbf{W}\mathbb{I} = \mathbf{W}[\delta_1 \delta_2 \dots \delta_{n_m}] = [\mathbf{W}\delta_1 \mathbf{W}\delta_2 \dots \mathbf{W}\delta_{n_m}], \quad (13)$$

where $\mathbf{W}\delta_i$ is computed with an available implementation of the FWT, such as the PyWavelets package (Lee *et al.* 2006) in Python. We distinguish the lengths of the signal in the different domains, more specifically $n_x \geq n_m$. This results from signal extension, which is an analogous situation as with convolutions. (In the filter bank interpretation, the relation between v_{nk} and $v_{n+1,k}$ is a convolution with a downsampled and time reversed impulse response.) Wavelets

with larger support widths yield wavelet representations with more coefficients. With finite signals, one must think about how to handle the boundaries. It is impossible to correlate a basis function with a part of a signal when that basis function has wider support than the length of that part of the signal. The signal is therefore by default symmetrically or periodically extended [e.g. in Matlab's implementation (Misiti *et al.* 2009)]. In a geological context, we generally do not expect the model parameters to be periodic, so we opt for a smooth signal extension. Another strategy to limit the boundary distortions, is to compute a maximum *useful* level of decomposition. The decomposition stops when the signal becomes shorter than the support width of the wavelet. The maximum level $N_{\max}$ corresponds with

$$N_{\max} = \left\lfloor \log_2 \left(\frac{n_m}{F-1} \right) \right\rfloor, \quad (14)$$

where $\lfloor \cdot \rfloor$ is the floor-operator and F is the number of filter coefficients, which scales with the support of the wavelet.

Finally, some properties about wavelets are introduced to better understand the approximating abilities of a specific wavelet. The number of vanishing moments is the most decisive property of a wavelet. A wavelet has p vanishing moments when

$$\int t^k \psi dt = 0 \quad \text{for } k = 0, 1, \dots, p-1. \quad (15)$$

The number of vanishing moments is related with the compact support of the wavelet: an orthonormal wavelet with p vanishing moments has at least a support of size $2p-1$ (Daubechies 1988). Wavelets with minimal compact support for a given p are called Daubechies wavelets (dbp). Further, if $f(t)$ is p times differentiable, its wavelet coefficients decay (Strang & Nguyen 1996):

Theorem 2.1. Decay of the Wavelet: If $f(t)$ is p times differentiable, its wavelet coefficients decay like 2^{-jp} :

$$|w_{j,k}| \leq C 2^{-jp} \|f^{(p)}(t)\|, \quad (16)$$

with C some positive constant.

Hence, wavelet coefficients decay with increasing resolution level. Inversely, discontinuities or singularities in the signal f yield large wavelet coefficients corresponding to the same spatial location on all resolution levels. This theorem, therefore, reveals an important trade-off: higher vanishing moments improve the approximating abilities of a wavelet (which is interesting in applications as it yields sparser representations) but involves wavelets with larger compact support. The larger the compact support, the more difficult to meet the conditions of the theorem: the signal f must then be piecewise smooth on a larger interval. Daubechies wavelets are by definition the wavelets with best approximating abilities: they are for a given number of vanishing moments, the orthogonal wavelet with the smallest compact support.

The Daubechies wavelet with two vanishing moments (db2) is shown in Fig. 1. Their scaling and wavelet functions in 1(a) and 1(b) have a support width of $2p-1$ or 3. Figs 1(d)–(f) show, respectively the one, two and three level discrete wavelet transform of signal f in Fig. 1(c). Signal f is a discretely sampled conductivity profile obtained from (Hermans & Irving 2017). It exhibits minimum structure in the sense that neighbouring samples have similar electrical conductivities. The samples can be viewed as scaling coefficients v_n at resolution level n . The one level DWT (with by default symmetric signal extension) transforms these scaling coefficients to scaling coefficients v_{n-1} and wavelet coefficients w_{n-1} at a coarser resolution level $n-1$. The former describes the signal at a coarser

scale, while the latter coefficients contain the details. In Fig. 1(d), the scaling coefficients, the 20 coefficients x_i in the first half of the figure, describe the original signal with less detail [and are also rescaled with a factor $\sqrt{2}$, see eq. (10)]. One can immediately notice that for such minimum-structure signals the majority of the wavelet coefficients, in the second half of the figure, are small. When the small wavelet coefficients, for example smaller than a threshold or a percentage of the maximum scaling coefficient, are ignored (set to zero), one gets a sparser representation of the original signal. The back transformation to model space (sparse signal f or dashed line in Fig. 1(c)) is still very similar to the original signal f . This is obtained by the common practice of setting all wavelet coefficients equal to zero if it is smaller than 5 per cent of the largest scaling coefficient. When this type of compression is applied on signal f , it results in a representation with all scaling coefficients and only one wavelet coefficient. To go from the one level to the two level DWT, the same procedure is applied on the scaling coefficients v_{n-1} (the wavelet coefficients w_{n-1} remain unaltered). The scaling coefficients v_{n-2} again describe the original signal at an even coarser scale and w_{n-2} describe the details that are lost since v_{n-1} . Fig. 1(f) shows the three level DWT. Note that the number of coefficients in wavelet space is slightly larger than in model space, due to the signal extension.

A wavelet with p vanishing moments is orthogonal to polynomials of degree $p-1$ (see eq. 15). Hence, the db2 wavelet is orthogonal to linear functions. This guarantees that wavelet coefficients will be zero for linear pieces in the signal f . The greater the number of vanishing moments, the more complex structures can be represented in a sparse fashion. This clearly shows that wavelets with higher vanishing moments have better approximating abilities. In Figs 1(j)–(i) we consider a discretely sampled linear function g and its maximum four level wavelet transform. One immediately notices that all wavelet coefficients (right-hand side) are zero. In the scaling coefficients, one can still recognize the original (rescaled) structure. The minimum-structure linear function can consequently be represented *exactly* with only three coefficients in the wavelet representation. To increase intuition, we return to eq. (9) in which we expand a signal into a basis. It is not obvious by eye that a linear combination of the irregular shaped db2 scaling function yields the original function. In this example, there are only three terms: three scaling functions with support width 3. Fig. 1(i) shows the superposition of the three scaling functions with non-zero coefficients. If we add the signals, we get a linear function in the original interval $[0, 1]$. Note that we used smooth signal extension here, as a result of which we effectively avoided boundary distortions.

2.4 Scale-dependent regularization

Finally, we introduce what we call scale-dependent regularization. The problem with the regularization term in eq. (7) is that it gives each coefficients in x the same weight. However, as we *know* that some components are more likely to be zero than others, we account for this in the objective function. The first coefficient will never be zero, because it would impose that the average value over the conductivity profile would be zero. Wavelet coefficients at smaller scales are expected to be zero, since these correspond with neighbouring model parameters having equal conductivities. Our wavelet basis, however, exhibits better minimum structure when the high resolution wavelet coefficients $x_i \in W_{-1}$ are sparse. Theorem 2.1 makes this more formal for any type of wavelet with p vanishing moments, recall that if f is locally smooth and p times differentiable, the theorem states that at scale $j+1$ the wavelet

Figure 1. Example of the wavelet transform with Daubechies wavelets with two vanishing moments (db2) on two discrete signals f and g . Vertical lines in figures (d)–(f) show the boundary between scaling coefficients v and wavelet coefficients w at different resolution levels.

coefficients, localized where f is smooth and p times differentiable, are approximately smaller than those on scale j by a factor of 2^p . This result can be used to define a new regularization term (where we assume no signal extension) on the wavelet representation $\mathbf{x} = [v_{0,0}, w_{0,0}, w_{1,0}, w_{1,1}, w_{2,1}, w_{2,1}, \dots, w_{n,k}, \dots]$:

$$\begin{aligned} \phi_m(\mathbf{x}) &= \frac{1}{E} \left(\sum_k \mu(w_{0,k}) + 2^1 \sum_k \mu(w_{1,k}) \right. \\ &\quad \left. + 2^2 \sum_k \mu(w_{2,k}) + 2^3 \sum_k \mu(w_{3,k}) + \dots \right) \\ &= \frac{1}{E} \sum_n 2^n \sum_k \mu(w_n, k), \end{aligned} \quad (17)$$

where we have made the sparsity constraint for high resolution coefficients more stringent and μ is the Eklblom measure. Note that we have dropped the number of vanishing moments p in the exponential as in Theorem 2.1, because this would be an excessive penalization of the small-scale wavelet coefficients for a large number of vanishing moments p (During the development phase of this inversion scheme, the factor 2^{jp} effectively suppressed all the small-scale coefficients to zero). Furthermore, we normalize the model misfit functional with the Euclidean norm (E) of coefficients to be able to better compare between different parametrizations.

2.5 Overview of the inversion scheme

Solving the inverse problem is basically minimizing an objective function. Constructing a suitable function, however, requires some thought and multiple assumptions must be made. Our inversion scheme is illustrated in Fig. 2. An iterative optimization method starts with an initial model or guess. In principle, this model can be based on prior knowledge of the geological context (e.g. groundtruth knowledge obtained in the vicinity of the surveying site) or by relying on the apparent conductivity, although the initial model can likewise be generated randomly. The latter option is less suitable for multimodal objective functions.

Following the generation of the initial model, it is transformed into the logarithmic domain. This is a common action in geophysical inversion schemes to ensure positive parameters, such as for the electrical conductivity. Of course, we will conduct a back-transform to the original domain later in the inversion scheme, where all values from the logarithmic domain are transformed to the $\mathbb{R}^+$ domain.

The model in the logarithmic domain is subsequently transformed into wavelet domain by the discrete wavelet transform. This involves a few preliminary choices that must be made before the transform can be applied. Each decision affects the objective function and hence potentially affects its minimum. First of all, one must decide which wavelet to use. Since in this paper only one dimension is considered, we choose wavelets from the Daubechies family for their approximating abilities, given the trade-off between compact

Figure 2. The inversion scheme in $\log_{10}$ -space domain.

support width and the number of vanishing moments. A second decision is the level of the wavelet transform. In principle, we require a sparse representation of a geologically realistic model, because we work with a sparsity promoting measure. We have already argued in the previous section that for a signal of length 2^N , it is not necessarily better to apply an N level DWT, due to boundary distortions. Nevertheless, we choose $N = \log_2(n_m)$ which turns out to work equally well. Finally note that the shape of the wavelet will affect the result of the inversion, especially if we opt for a strong degree of regularization. Accordingly, if a blocky model is expected, it is better to opt for blocky wavelets which have sharper edges, such as the db1 wavelet. If soft boundaries are expected between the layers, then smoother wavelets are more appropriate. These wavelets usually have a larger number of vanishing moments and have a stabilizing effect due to their larger compact support.

Using the perturbed ℓ_1 -norm measure of Eklblom (5) introduces an extra parameter ϵ . For values $x \gg \epsilon$, the Eklblom-measure behaves as the ℓ_1 -norm, while for values $x \ll \epsilon$ it behaves like a scaled sum-of-squares measure. While the latter property is theoretically less desirable, it does not pose any problem as this will only occur for small values for x which only have a negligible effect on the back-transformation to the space-domain. Geometrically speaking, the diamond representing the ‘unit circle’ of the ℓ_1 -norm will have rounded corners which allow more solutions than the ℓ_1 -norm, but with nearly zero components. In our scheme, components of the model in wavelet space x_i smaller than 10^{-4} have little effect on the actual conductivity profile, thus for $\epsilon = 10^{-6} \ll 10^{-4}$ we have a sparsity measure on the relevant domain of $\mathbb{R}$. Tests on the sensitivity of ϵ in our scheme reveal that there is only a small effect for this specific parameter. It is therefore not further studied.

There are various optimization methods available, classified into trust-region, line-search and non-gradient methods. In this paper, we exclusively use the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm. This algorithm is implemented in Python (SciPy, Jones *et al.* 2001) with the line search algorithm that meets the Wolfe conditions (Moré & Thuente 1994). This method is a Newton’s method and approximates the Hessian at every iteration. In practice, the choice of tolerances (stop criteria) will affect the actual outcome. In general, the tolerance on the objective function is set to 10^{-11} and 10^{-10} for the gradient which is sufficient (for further details, see Jones *et al.* 2001). For the forward model, we choose the

damped model, because of its appealing balance between accuracy and computational cost.

3 RESULTS

3.1 Scale-dependent regularization

In the first examples, we demonstrate the advantage of the scale-dependence of our novel regularization scheme. Whereas other schemes (such as Liu *et al.* 2017; Nittinger & Becken 2016) penalize each wavelet coefficient with equal weight, our scheme penalizes each wavelet coefficient proportionally with their resolution level n .

To make a valid comparison between the two wavelet-based regularization schemes, we first eliminate all other factors in the inversion process that affects the outcome. Such factors are, for example, the presence of numerous local minima in the objective function, the presence of a modelling error κ in the forward operator or a limited maximal number of iterations due to limited computational resources. We eliminate the modelling error by using the same forward operator for the synthetic data generation and the inversion. The objective function will have good properties with the LIN approximation because the model is linear and consequently the data misfit term will be quadratic, while the Eklblom measure is convex. Furthermore, choosing this forward model overcomes the computational burden. We elaborate on the effect of the forward model in Section 3.3.

The first example in Fig. 3 shows a realistic synthetic, three-layer conductivity profile, where a highly conductive layer of 0.8 Sm^{-1} from a depth of 2.5–5 m is embedded in a less conductive half-space of 0.1 Sm^{-1} . This is a case where smooth inversion typically yields a very smeared out profile, while a blocky outcome is better suited. We show the result with db1 wavelets (see Fig. A1 in Appendix B for the shape of the wavelets, in the next section with other wavelets and smooth inversion) with both scale-dependent regularization ($2^j \lambda$) and the traditional non-scale-dependent sparsity inversion with equal regularization (const. λ). The L-curve criterion is used for the estimation of the optimal regularization parameter. While the convexity of the L-curve is not guaranteed, the L-curves displayed the typical L-shape (not shown). The data misfits ϕ_d corresponding to the inversion models are presented in Table 1.

Figure 3. Comparison of traditional non-scale-dependent versus scale-dependent wavelet-based regularization schemes with the LIN approximation as forward operator.

Table 1. Data-misfits ϕ_d for the inversion models in Fig. 3.

	db1	db6
const. λ	1.18×10^{-4}	1.75×10^{-6}
$2^j \lambda$ (scale-dependent)	1.37×10^{-4}	1.47×10^{-6}

The synthetic data is generated for a vertical dipole in both components (horizontal coplanar HCP and perpendicular PERP) for 20 equidistant intercoil distances from 1 to 20 m at a height of 0.1 m above the soil and where we have added 5 per cent Gaussian multiplicative noise. We choose to fit 32 model parameters. The result of the scale-dependent sparsity inversion outperforms the traditional wavelet-based inversion. The scale-dependent scheme follows the true model almost exactly, while the latter has too much structure at the small scales. This example illustrates the advantage of scale-dependent regularization over traditional wavelet-based inversion with db1 wavelets.

The second example shows a conductivity profile from a borehole logging from an alluvial aquifer (Hermans & Irving 2017). It has 10 times smaller conductivity values and therefore the variation in the magnetic field data will be more subtle. In this case, we opt for a more realistic undetermined problem in which we fit 64 model parameters with the db6 wavelet to data from a vertical dipole for both components and five equidistant intercoil distances from 1 to 10 m. The results in Fig. 3 show for both schemes a wrong location of the main peak at 2.5 m depth. However, we observe that the scale-dependent regularization has a much narrower peak and whose amplitude is closer to the exact values indicating that the scale-dependent regularization seems also better suited for other wavelet bases. The fact that wavelets are capable of representing smooth profiles with sharp anomalies in a sparse fashion is truly a result of its multiresolution character. While this behaviour was already observed for a combination of two or more wavelet families (Nittinger & Becken 2018), it is now observed with only one wavelet type in combination with a scale-dependent regularization term. In Loris *et al.* (2010), db4 was suggested as the wavelet that was able to retain sharp features, while not obtaining too blocky results as with db1. In our scheme, wavelets with larger compact support can also lead to inversion models with sharp features. The misidentification of the peak is related to the combination of the noise level with the challenging nature of the inverse problem. The inversion of the noise-free data set (shown in the Appendix C in Fig. A2) yields a correct location of the peak (for both regularization methods).

3.2 Model space versus wavelet space

In this section, we illustrate our proposed regularization scheme and take the traditional scheme with smoothing Tikhonov regularization as benchmark. Our flexible scheme can recover blocky structures and sharp boundaries, which is lacking in minimum-structure inversions with smoothing constraints. To illustrate the advantage of wavelet space over standard model space, we perform a numerical test on the three-layered model described in the previous section, with the only difference that we fit 64 model parameters instead of 32. The results after inversion are shown in Fig. 4 the corresponding data misfits ϕ_d are shown in Table 2. From the shapes of the wavelet, it is expected that the db1-wavelet yields the best (blocky) inversion. We therefore compare the results of the db1-wavelet with smoothing inversion in Fig. 4(a). The wavelet-based inversion follows the model closely, while the smoothing inversion locates the peak correctly, but with an underestimated value for the electrical conductivity as expected. The smooth inversion also creates a high conductivity artefact close to the surface. In addition, the interface between the two different layers is not identified. Figs 4(b)–(d) show inversion results for wavelets with increasing vanishing moment (and increasing regularity). To facilitate result comparison, the regularization parameter was fixed to 10^{-5} for all inversions (the L-curve criterion selected very similar regularization parameters for each wavelet). As expected from the shape of the wavelet, the db2-wavelet is capable of both showing sharp and relatively smooth transitions. The peak is slightly dislocated, this is possibly an effect of the asymmetric shape of the wavelet function. For other wavelets, the inversion results are relatively identical and smoother, but with a larger peak than commonly found with Tikhonov regularization. So far, we could not determine a method to select an optimal wavelet from the data, as this remains strongly dependent on the expected geological variations. On the other hand, one can impose the characteristics of the inversion scheme as desired. For blocky inversion, one chooses a low number of vanishing moments, while for smooth inversion, a larger vanishing moment can be opted for. Note that in general the maximal electrical conductivity value is recovered more exactly with wavelet inversion than for smooth Tikhonov inversion (i.e. large number of vanishing moments) and that the results are more symmetrical near the actual conductive layer. The only disadvantage is the extra structure at larger depths. Those are oscillations which are a result of the shape of the wavelet (the more vanishing moments, the more oscillations

Figure 4. Models after inversion with no modelling error ($\kappa = 0$) and the wavelet-based regularization with various Daubechies wavelets versus Tikhonov regularization.

Table 2. Data-misfits ϕ_d for the inversion models in Fig. 4.

ℓ_2	db1 ($\lambda = 3.2 \times 10^{-4}$)	db1 ($\lambda = 10^{-5}$)	db2	db3	db4	db5	db6	db7	db8	db9
1.116	1.363	1.120	1.135	1.121	1.110	1.103	1.119	1.119	1.117	1.132 $\times 10^{-4}$

as can be seen from Fig. A1 in Appendix B). We do not conclude that a specific wavelet basis is particularly better than other bases, what is for example sometimes concluded in other literature about wavelet inversion (Loris *et al.* 2010; Fang & Zhang 2014).

The scale-dependent wavelet based inversion is advantageous for recovering smoother models as well. As in the previous section, we compare the inversion with the conductivity profile from the alluvial aquifer. For each inversion, the L-curve criterion determines the optimal regularization parameter which is different for each wavelet. The results for Tikhonov, db1 and db6 wavelet inversion are shown in Fig. 5 and data misfits ϕ_d are presented in Table 3. The results for db1 and Tikhonov are similar. Both schemes recover peaks at the same locations with more or less the same values for electrical conductivity. As with the previous test with this profile, the wavelets with more vanishing moments can recover the anomalies better. The case with the correct location of the peak in the absence of experimental noise is shown in Fig. A3 in Appendix C.

The shape of the wavelet (and scaling function) affects the inversion result (this was already mentioned in Nittinger & Becken 2016), and choosing the wavelet should be based on prior knowledge. This can be understood from the theoretical framework, namely by penalizing large coefficients of the wavelet representation of a model, the

Figure 5. Models after inversion with no modelling error ($\kappa = 0$) and the wavelet-based regularization with the db1 and db6-wavelet versus Tikhonov regularization.

Table 3. Data-misfits ϕ_d for the inversion models in Fig. 5.

ℓ_2	db1	db6
1.70	1.44	1.59 $\times 10^{-6}$

Table 4. Data-misfits ϕ_d for the inversion models in Fig. 6.

$\ell_2, n_m = 32$	$\ell_2, n_m = 64$	$\ell_2, n_m = 128$	db1	db2	db3	db6
2.70	2.10	2.27	2.46	2.31	2.21	1.96

$\times 10^{-6}$

Figure 6. Models after inversion with no modelling error ($\kappa = 0$) and the wavelet-based regularization with various Daubechies versus Tikhonov regularization. **Figure 7.** Magnetic field ratio generated via the different forward operators from the conductivity profile in Fig. 8.

outcome is altered with exactly the shape of the wavelet or scaling function. Therefore, the shape is often recognized in the inversion result. Consider another example with a softer interface between two different layers. In Fig. 6, such a conductivity profile from (Hermans *et al.* 2012) is shown from De Panne, Flanders, where a saltwater lens is lying on a clay layer. We show the inversion results for four

different wavelets: db1, db2, db3 and db6. In general, better results (in terms of similarity between the true model and inversion model) are obtained when more model parameters are considered for more vanishing moments. The data misfits ϕ_d are shown in Table 4. For each case, the Tikhonov smoothing inversion is also shown in the corresponding number of model parameters. For all those cases, the Tikhonov regularization cannot identify the actual peak, while the wavelet inversion can, independently from the chosen wavelet. We again observe a decreasing blockyness for larger vanishing moments and an artificial second peak for db3 and db6, due to the shape of the chosen wavelet and the data not providing a resolution of such secondary peaks.

Regularized deterministic inversion is always limited in terms of uncertainty estimation and bears the risk to introduce inconsistent prior information through the choice of the regularization functional. The use of our algorithm, which is flexible in the sense that several wavelet bases can be used, offer a straightforward way to assess the influence of the choice of the wavelet, and identify zones that are more likely to display undesired features related to the regularization (e.g. secondary peaks). For example, one might compare the results from db1 to db6 to identify such occurrences.

This touches upon an issue addressed in Loris *et al.* (2010) where the selective nature of the wavelet coefficients and the ℓ_1 criterion was able to reconstitute a large scale anomaly in the absence of

Figure 8. Effect of forward operator (and modelling error κ) on the inversion result.

Figure 9. FDEM field data obtained with the DUALEM 421S, from Bobe *et al.* (2020).

sufficient data, what turned out to be a lucky exception than a rule. This is a relevant remark, as FDEM data has a limited resolution. Moreover, Loris *et al.* (2010) conclude that the success of the ℓ_1 method heavily depends on the choice of a suitable basis, which in the absence of prior knowledge is unknown. Interpreting the results as an ensemble might resolve these issues.

3.3 The role of the forward model operator

In the previous section in the absence of unmodelled errors ($\|\kappa\| = 0$), our proposed inversion scheme successfully recovered blocky structures. Next, we test the effectiveness of our scheme with an approximated forward operator $\mathcal{F}_{\text{approx}}$ instead of the exact computation $\mathcal{F}$, for the purpose of computational efficiency. The synthetic data is generated from the operator $\mathcal{F}$ to mimic a true survey (including 5 per cent multiplicative Gaussian noise), hence we examine our scheme in the presence of biased error κ . Using the scheme for both the LIN approximation and the damped model allows comparing the effect of the degree of unmodelled errors on the result of the inversion, especially in a conductive setting where the LIN assumption breaks down.

The magnetic field data generated from the exact model $\mathcal{F}$, the LIN approximation $\mathcal{F}_{\text{LIN}}$ and the damped model $\mathcal{F}_{\text{damped}}$ are shown in Fig. 7. From this data, it is clear that the noise is biased and larger for the LIN approximation than for the damped model, that is their error

Table 5. Data-misfits ϕ_d for the inversion models in Fig. 8.

	db1	ℓ_2
$\mathcal{F}_{\text{LIN}}$	2.17×10^{-4}	2.11×10^{-4}
$\mathcal{F}_{\text{damped}}$	1.43×10^{-4}	1.27×10^{-4}

is $\|\kappa_{\text{LIN}}\| = 0.23$ and $\|\kappa_{\text{damped}}\| = 0.026$, respectively. As expected, we get $\|\kappa_{\text{damped}}\| < \|\kappa_{\text{LIN}}\|$. In terms of data-misfit (3), the unbiased experimental noise equals 0.000149, the biased unmodelled error in damped and experimental noise equals 0.0002664, while for the LIN approximation this equals 0.00996. Note that the error and its norm will differ for different conductivity profiles.

The L-curve criterion can easily locate vertical branch due to the experimental random noise, but we can not use this principle to estimate a regularization parameter such that the data misfit meets the value of the sum of the experimental and biased noise. Furthermore, we cannot rely on the discrepancy principle, since the biased noise κ is unknown. The optimal estimated regularization parameters following the L-curve criterion yield the estimates in the range of the experimental noise. Their data misfits ϕ_d for the LIN and damped model are 0.000217 and 0.000143, respectively. The damped model has a misfit approximately equal to the experimental noise level, while for LIN this is slightly larger. The results are shown in Fig. 8 and the corresponding data misfits ϕ_d are presented in Table 5. The conductivity profile obtained via the damped model has only a slightly deviant electrical conductivity, while for the LIN approximation slightly too much structure was introduced. However, both models recover the profile relatively well. Surprisingly, the smooth inversion with the LIN approximation yields better results (in terms of similarity between the true model and inversion model). This example illustrates the importance of a good approximation for wavelet-based inversion, yet, our method seems relatively robust.

3.4 Field data case

In this section, we demonstrate the wavelet-based inversion scheme on real FDEM field data, obtained from Bobe *et al.* (2020). The measurement location is in Gontrode forest, Belgium, known for its horizontal stratigraphy of different soil layers (i.e. no 2-D or 3-D effects are expected). For the sounding, the DUALEM 421S was used with six receiver coils (three HCP coils at 1, 2 and 4 m and

Figure 10. Inversion results from Tikhonov regularization (ℓ_2) and wavelet-based regularization with Daubechies (db) wavelets. Estimation of regularization parameter via L-curve criterion. The blue reference model is the FDEM inversion obtained from Bobe *et al.* (2020) within the 2σ -uncertainty bands in log-space.

Table 6. Data-misfits ϕ_d for the inversion models in Fig. 10.

	db1	db2	db3	db6	ℓ_2
Data misfit $\phi_d (\times 10^{-2})$	2,470	2,435	2,355	2,437	2,519

three PRP coils at 1.1, 2.1 and 4.1 m) with an operating frequency of 9.0 kHz. The sounding was performed at zero, 30 and 60 cm height. The magnetic field ratios are shown in Fig. 9.

Bobe *et al.* (2020) conducted a visual inspection of the subsurface for the validation of the results, by drilling a hole with an Edelman auger. The top 0.15 m consists of organic rich forest material with many air-filled voids. Between 0.15 and 0.8 m, a clay-rich sand was observed. The clay content decreased with further depth. At 1.5 m, the groundwater table was reached.

The result of the inversion procedure with the damped model as forward operator is shown for db1, db2, db3 and db6 wavelets and for smooth inversion in Fig. 10. The data misfits ϕ_d are presented in Table 6. The results of the inversion with other wavelets exhibited similar characteristics. We have used an equidistant parameterization in log10-space of 64 model parameters up to 30 m depth and only the outcome of the first 3 m is shown (because of the little sensitivity in the lower depth region). The results can be compared to the outcome of Bobe *et al.* (2020), shown as the blue reference model within the 2σ -uncertainty bands originating from a Kalman-ensemble generator approximating the stochastic posterior solution of the problem, in which, the clay-rich sand was not recovered. It shows a steady increase in electrical conductivity from 0.01 S/m at the surface to almost 0.8 S/m at 2 m depth, followed by a constant plateau.

We expect the first 0.15 m to be highly resistive (air voids are an insulator). Clay is known to have a relatively high bulk electrical conductivity value, which explains the peak in the conductivity profiles. Sand, on the other hand, has a low bulk electrical conductivity. Groundwater contains more free electrical charges and therefore, the electrical conductivity usually increases in the saturated zone. This behaviour is recovered in almost all wavelet-based inversion results, except for the db1 wavelet, which approximately corresponds to the expected solution of a 3-layered blocky model.

Based on the lithological profile, db3 seems to provide the most realistic results. It also corresponds to the lowest data misfit ϕ_d in Table 6. The main difference with smooth reference inversion

from Bobe *et al.* (2020) is the first peak in electrical conductivity due to the clay content. The smooth inversion (ℓ_2) with identical parameterization as the wavelet-based inversion and the damped model as forward operator also seems to provide consistent results. Differences between the inversion models indicate zones where the resolution might be lower, what can be nicely shown using different bases with outcomes that have similar data misfits.

The wavelet-based regularization scheme yields results with similar characteristics, yet the results are different. This is the faith of deterministic inversion, however the wavelet-based inversion scheme would also suit in stochastic inversion methods.

4 CONCLUSION

We have introduced an alternative inversion scheme for EMI surveys that can be extended to any other 1-D geophysical method. It involves a new regularization term based on the wavelet transform and scale-dependent weighting which can easily be combined with the existing framework of deterministic inversion (gradient-based optimization methods, L-curve criterion for optimal regularization parameter).

Our results show that the scale-dependent wavelet-based inversion scheme with sparsity constraint is more flexible than Tikhonov regularization since it can recover both blocky and smooth conductivity profiles, by adequately choosing the number of vanishing moments of the wavelet. Furthermore, our scheme can recover high amplitude anomalies in combination with globally smooth profiles what is generally not possible with ℓ_2 -regularized inversion schemes. The scale-dependency of our scheme allows the use of wavelets with few vanishing moments and is, therefore, an improvement with respect to existing wavelet-based regularization schemes. The nature of the inversion method allows for high flexibility because the shape of the wavelet can be exploited to generate multiple representations of the inverse model. Depending on available prior knowledge, the final result can be chosen or the results can be interpreted together. Although the choice of the wavelet is case-specific, we have observed few variations in the recovered model above db6. Together with the computational speed-up w.r.t. schemes using computationally more expensive forward models, our method thus offers a viable alternative to the existing methodologies.

We have shown that the inversion scheme was robust even for approximate forward models. However, this should be confirmed in

future work for more challenging conditions such as saline grounds and where a potential multimodal character of the objective function, due to an increased non-linear behaviour of the forward model, poses new challenges.

ACKNOWLEDGEMENTS

The authors thank Christin Bobe for making available the data from the field work in Gontrode, Belgium and Marieke Paepen for the conversations on the inversion of those data. The research leading to these results has received funding from FWO (Fund for Scientific Research, Flanders, grant 1113020N). We also acknowledge three reviewers, Michael Becken, Maximilian Weigand and an anonymous reviewer, for their thoughtful and valuable comments on this paper.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, Wouter Deleersnyder, upon reasonable request.

REFERENCES

- Anderson, W.L., 1979. Numerical integration of related Hankel transforms of orders 0 and 1 by adaptive digital filtering, *Geophysics*, **44**(7), 1287–1305.
- Archie, G.E. *et al.*, 1942. The electrical resistivity log as an aid in determining some reservoir characteristics, *Trans. AIME*, **146**(01), 54–62.
- Bobe, C., Hanssens, D., Hermans, T. & Van De Vijver, E., 2020. Efficient probabilistic joint inversion of direct current resistivity and small-loop electromagnetic data, *Algorithms*, **13**(6), 144, doi:10.3390/a13060144.
- Bunks, C., Saleck, F.M., Zaleski, S. & Chavent, G., 1995. Multiscale seismic waveform inversion, *Geophysics*, **60**(5), 1457–1473.
- Candes, E.J. & Tao, T., 2005. Decoding by linear programming, *IEEE Trans. Inform. Theory*, **51**(12), 4203–4215.
- Candès, E.J. & Wakin, M.B., 2008. An introduction to compressive sampling, *IEEE Signal Process. Mag.*, **25**(2), 21–30.
- Candes, E.J., Wakin, M.B. & Boyd, S.P., 2008. Enhancing sparsity by reweighted l1 minimization, *J. Fourier Anal. Appl.*, **14**(5–6), 877–905.
- Chen, S. & Donoho, D., 1994. Basis pursuit, in *Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers*, Vol. **1**, pp. 41, IEEE.
- Cockett, R., Kang, S., Heagy, L.J., Pidlisecky, A. & Oldenburg, D.W., 2015. SimPEG: An open source framework for simulation and gradient based parameter estimation in geophysical applications, *Comput. Geosci.*, **85**, 142–154.
- Constable, S.C., Parker, R.L. & Constable, C.G., 1987. Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, **52**(3), 289–300.
- Daubechies, I., 1988. Orthonormal bases of compactly supported wavelets, *Commun. Pure Appl. Math.*, **41**(7), 909–996.
- Daubechies, I., 1992. *Ten Lectures on Wavelets*, Vol. **61**, SIAM.
- Daubechies, I., Defrise, M. & De Mol, C., 2004. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Commun. Pure Appl. Math.*, **57**(11), 1413–1457.
- De Smedt, P. *et al.*, 2011. Reconstructing palaeochannel morphology with a mobile multicoil electromagnetic induction sensor, *Geomorphology*, **130**(3–4), 136–141.
- Deidda, G.P., Fenu, C. & Rodriguez, G., 2014. Regularized solution of a non-linear problem in electromagnetic sounding, *Inverse Problems*, **30**(12), 125014, doi:10.1088/0266-5611/30/12/125014.
- Deidda, G.P., de Alba, P.D., Fenu, C., Lovicu, G. & Rodriguez, G., 2020. FDEMtools: a MATLAB package for FDEM data inversion, *Numer. Algorithms*, **84**(4), 1313–1327.
- Donoho, D.L., 2006. For most large underdetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution, *Commun. Pure Appl. Math.*, **59**(6), 797–829.
- Duy Thanh, L., Jougnot, D., Van Do, P. & Van Nghia, A. N., 2019. A physically based model for the electrical conductivity of water-saturated porous media, *Geophys. J. Int.*, **219**(2), 866–876.
- Eklblom, H., 1987. The l1-estimate as limiting case of an lp-or huber-estimate, in *Statistical Data Analysis Based on the L1-Norm and Related Methods: 31/08/1987-04/09/1987*, pp. 109–116, Elsevier.
- Fang, H. & Zhang, H., 2014. Wavelet-based double-difference seismic tomography with sparsity regularization, *Geophys. J. Int.*, **199**(2), 944–955.
- Farquharson, C.G., 2007. Constructing piecewise-constant models in multi-dimensional minimum-structure inversions, *Geophysics*, **73**(1), K1–K9.
- Fernández, J.P., Shubitidze, F., Shamatava, I., Barrowes, B.E. & O'Neill, K., 2010. Realistic subsurface anomaly discrimination using electromagnetic induction and an SVM classifier, *EURASIP J. Adv. Signal Process.*, **2010**(1), 305890, doi:10.1155/2010/305890.
- Guillemoteau, J., Simon, F.-X., Lück, E. & Tronicke, J., 2016. 1d sequential inversion of portable multi-configuration electromagnetic induction data, *Near Surf. Geophys.*, **14**(5), 423–432.
- Guittou, A., 2012. Blocky regularization schemes for full-waveform inversion, *Geophys. Prospect.*, **60**(5), 870–884.
- Hansen, P.C., 2010. *Discrete Inverse Problems: Insight And Algorithms*, Vol. **7**, SIAM.
- Hansen, P.C., Jensen, T.K. & Rodriguez, G., 2007. An adaptive pruning algorithm for the discrete l-curve criterion, *J. Comput. Appl. Math.*, **198**(2), 483–492.
- Hermans, T. & Irving, J., 2017. Facies discrimination with electrical resistivity tomography using a probabilistic methodology: effect of sensitivity and regularisation, *Near Surf. Geophys.*, **15**(1), 13–25.
- Hermans, T., Vandenbohede, A., Lebbe, L., Martin, R., Kemna, A., Beaujean, J. & Nguyen, F., 2012. Imaging artificial salt water infiltration using electrical resistivity tomography constrained by geostatistical data, *J. Hydrology*, **438**, 168–180.
- Jadoon, K.Z., McCabe, M.F. & Moghadas, D., 2015. Application of electromagnetic induction to monitor changes in soil electrical conductivity profiles in arid agriculture, in *First Conference on Proximal Sensing Supporting Precision Agriculture*, Vol. **2015**, pp. 1–5, European Association of Geoscientists & Engineers.
- Jones, E., Oliphant, T., Peterson, P., *et al.*, 2001. *SciPy: Open Source Scientific Tools for Python*. [Online; accessed 12-01-2019].
- Lee, G., Wasilewski, F., Gommers, R., Wohlfahrt, K., O'Leary, A. & Nahrstaedt, H., 2006. *Pywavelets-Wavelet Transforms in Python*, <https://github.com/PyWavelets/pywt> [Online; accessed 2018-MM-DD].
- Lima, W.A., Silva, J.B., Santos, D.F. & Costa, J.C., 2019. A robust interactive estimation of the regularization parameter, *Geophysics*, **84**(3), IM19–IM33.
- Linde, N., Renard, P., Mukerji, T. & Caers, J., 2015. Geological realism in hydrogeological and geophysical inverse modeling: a review, *Adv. Water Resour.*, **86**, 86–101.
- Liu, Y., Farquharson, C.G., Yin, C. & Baranwal, V.C., 2017. Wavelet-based 3-D inversion for frequency-domain airborne EM data, *Geophys. J. Int.*, **213**(1), 1–15.
- Loris, I., Douma, H., Nolet, G., Daubechies, I. & Regone, C., 2010. Nonlinear regularization techniques for seismic tomography, *J. Comput. Phys.*, **229**(3), 890–905.
- Mallat, S., 1999. *A Wavelet Tour of Signal Processing*, Elsevier.
- Mallat, S.G., 1989. A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. Pattern Anal. Mach. Intell.*, **(7)**, 674–693.
- Maveau, B., Delrue, S. & Dudal, D., 2020. A damped forward EMI model for a horizontally stratified earth, *Explor. Geophys.*, **51**(7), 1–12.
- McNeill, J., 1980. Electromagnetic terrain conductivity measurement at low induction numbers, Technical Note TN-6, Geonics Limited, Ontario, Canada.
- Misiti, M., Misiti, Y., Oppenheim, G. & Poggi, J., 2009. *MATLAB Wavelet Toolbox™ 4 User'S Guide*, pp. 153, The MathWorks, Inc. Natick.

- Moré, J.J. & Thuente, D.J., 1994. Line search algorithms with guaranteed sufficient decrease, *ACM Transactions on Mathematical Software (TOMS)*, **20**(3), 286–307.
- Nittinger, C. & Becken, M., 2018. Compressive sensing approach for two-dimensional magnetotelluric inversion using wavelet dictionaries, *Geophys. Prospect.*, **66**(4), 664–672.
- Nittinger, C.G. & Becken, M., 2016. Inversion of magnetotelluric data in a sparse model domain, *Geophys. J. Int.*, **206**(2), 1398–1409.
- Nocedal, J. & Wright, S., 2006. *Numerical Optimization*, Springer Science & Business Media.
- Paepen, M., Hanssens, D., Smedt, P.D., Walraevens, K. & Hermans, T., 2020. Combining resistivity and frequency domain electromagnetic methods to investigate submarine groundwater discharge in the littoral zone, *Hydrol. Earth Syst. Sci.*, **24**(7), 3539–3555.
- Pettersson, J.K. & Nobes, D.C., 2003. Environmental geophysics at Scott base: ground penetrating radar and electromagnetic induction as tools for mapping contaminated ground at antarctic research bases, *Cold Regions Sci. Technol.*, **37**(2), 187–195.
- Revil, A. et al., 2017. Complex conductivity of soils, *Water Resour. Res.*, **53**(8), 7121–7147.
- Saey, T., De Smedt, P., Meerschman, E., Islam, M.M., Meeuws, F., Van De Vijver, E., Lehouck, A. & Van Meirvenne, M., 2012. Electrical conductivity depth modelling with a multireceiver emi sensor for prospecting archaeological features, *Archaeol. Prospect.*, **19**(1), 21–30.
- Simpson, D., Van Meirvenne, M., Saey, T., Vermeersch, H., Bourgeois, J., Lehouck, A., Cockx, L. & Vitharana, U.W., 2009. Evaluating the multiple coil configurations of the em38dd and dual-em-21s sensors to detect archaeological anomalies, *Archaeol. Prospect.*, **16**(2), 91–102.
- Strang, G. & Nguyen, T., 1996. *Wavelets and Filter Banks*, SIAM.
- Strutz, T., 2010. *Data Fitting and Uncertainty: A Practical Introduction to Weighted Least Squares and Beyond*, Vieweg and Teubner.
- Tantum, S.L., Scott, W.R., Morton, K.D., Collins, L.M. & Torrione, P.A., 2012. Target classification and identification using sparse model representations of frequency-domain electromagnetic induction sensor data, *IEEE Trans. Geosci. Remote Sens.*, **51**(5), 2689–2706.
- Thibaut, R., Kremer, T., Royen, A., Ngun, B.K., Nguyen, F. & Hermans, T., 2021. A new workflow to incorporate prior information in minimum gradient support (mgs) inversion of electrical resistivity and induced polarization data, *J. appl. Geophysics*, **187**, 104286, doi:10.1016/j.jappgeo.2021.104286.
- Tikhonov, A.N., 1943. On the stability of inverse problems, *Dokl. Akad. Nauk SSSR*, **39**, 195–198.
- Valentine, A.P. & Sambridge, M., 2018. Optimal regularization for a class of linear inverse problem, *Geophys. J. Int.*, **215**(2), 1003–1021.
- Virieux, J. & Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics, *Geophysics*, **74**(6), WCC1–WCC26.
- Voronin, S., Mikesell, D. & Nolet, G., 2015. Compression approaches for the regularized solutions of linear systems from large-scale inverse problems, *GEM-Int. J. Geomath.*, **6**(2), 251–294.
- Wait, J.R., 1951. The magnetic dipole over the horizontally stratified earth, *Can. J. Phys.*, **29**(6), 577–592.
- Wait, J.R., 1962. A note on the electromagnetic response of a stratified earth, *Geophysics*, **27**(3), 382–385.

APPENDIX A: COMPUTATIONAL COST RELATED TO THE CHOICE OF THE FORWARD OPERATOR

The bottleneck in gradient based optimization is typically not the calculation of the field response via the forward model, but is the calculation of the derivative or sensitivity matrix. For (semi-)analytical models, analytical expressions of the derivatives are available. However, evaluating such expressions are not always computationally cheap. In the LIN-approximation, the magnetic field ratio has a linear relation with the electrical conductivity. For linear operators, the

Table A1. Execution time comparison for the calculation of the fields and Jacobians for 100 iterations.

	Wait (non-linear)	Damped (quasi-linear)	LIN (linear)
Fields	4.386 s	0.851 s	0.00497 s
Jacobian	20.6 s ^a	3.80 s	0.005 s

^aThe calculation of the Jacobian for Wait's model was performed with the FDEMtools package (Deidda et al. 2020) on a Matlab engine, which may affect the validity of the comparison. However, execution times of the same order of magnitude are expected in similar Python routines.

general expression of forward operator in eq. (1) reduces to a matrix F

$$d_i = F_{ij} m_j. \quad (\text{A1})$$

The linearity of the LIN approximation allows to write the matrix F , independent of the conductivities m_j and therefore the computation of the derivative is straightforward.

The damped model is a quasi-linear, closed form relation between the conductivities and the fields. The general expression of the forward model eq. (1) reduces now to

$$d_i = F_{ij}(\mathbf{m}) m_j, \quad (\text{A2})$$

more specifically and as discussed in Section 2.1, the matrix F is now dependent on the background conductivity. When that conductivity is chosen to be constant, the model is linear. The optimal background conductivity, however, is based on the knowledge of the profile. Maveau et al. (2020) suggest to use the weighted average (with the thickness of each layer as weight) of the conductivities above the layer for which the contribution is calculated. In an inversion process, the conductivities change and thus the matrix has to be calculated at each iteration step. This involves multiple evaluations of modified Bessel functions, but no integrations over Bessel functions, as in the forward model of Wait.

The Jacobian for the non-linear forward model by Wait (1951) is derived in multiple texts. For example in Deidda et al. (2014), the complexity of the scheme is $\mathcal{O}((3n^2 + 8mn)q)$ floating point operations, where n is the number of layers, m the number of data values, and q the nodes in the quadrature formula used to approximate the Hankel transform (see Anderson 1979). And as with the damped model, the Jacobian matrix has to be computed at every iteration. However, Deidda et al. (2014) state that the evaluation of the analytical expression is still up to seven times faster than the finite differences approximation. Therefore, those methods were not considered in our analysis.

To make this discussion more tangible, the execution time of the computations of the fields and Jacobian are presented in Table A1. The routines were executed on a standard scientific laptop (with an Intel i7-8850H processor). The comparison confirms an increasing computational burden for the linear LIN, quasi-linear damped and non-linear full solution of Wait respectively.

APPENDIX B: DAUBECHIES WAVELETS

APPENDIX C: RESULTS WITHOUT EXPERIMENTAL NOISE

Figure A1. Scaling functions of the wavelets used in this paper.

Figure A2. Identical inverse problem as in Section 3.1 but without any multiplicative noise.

Figure A3. Identical inverse problem as in Section 3.1 but without any multiplicative noise.