




Article

Tsallis Entropy, Likelihood, and the Robust Seismic Inversion

Igo Pedro de Lima ^{1,†}, Sérgio Luiz E. F. da Silva ^{2,*,†} , Gilberto Corso ^{1,3,†}  and João M. de Araújo ^{1,2,†} 

¹ Programa de Pós-Graduação em Ciência e Engenharia de Petróleo - Universidade Federal do Rio Grande do Norte, Natal RN 59078-970, Brazil; igoivcds@hotmail.com (I.P.d.L.); gfcorso@gmail.com (G.C.); joamedeiros@fisica.ufrn.br (J.M.d.A.)

² Departamento de Física Teórica e Experimental, Universidade Federal do Rio Grande do Norte, Natal RN 59078-970, Brazil

³ Departamento de Biofísica e Farmacologia, Universidade Federal do Rio Grande do Norte, Natal RN 59078-970, Brazil

* Correspondence: sergioluizufnr@ufrn.edu.br

† These authors contributed equally to this work.

Received: 2 April 2020; Accepted: 17 April 2020; Published: 19 April 2020



Abstract: The nonextensive statistical mechanics proposed by Tsallis have been successfully used to model and analyze many complex phenomena. Here, we study the role of the generalized Tsallis statistics on the inverse problem theory. Most inverse problems are formulated as an optimisation problem that aims to estimate the physical parameters of a system from indirect and partial observations. In the conventional approach, the misfit function that is to be minimized is based on the least-squares distance between the observed data and the modelled data (residuals or errors), in which the residuals are assumed to follow a Gaussian distribution. However, in many real situations, the error is typically non-Gaussian, and therefore this technique tends to fail. This problem has motivated us to study misfit functions based on non-Gaussian statistics. In this work, we derive a misfit function based on the q -Gaussian distribution associated with the maximum entropy principle in the Tsallis formalism. We tested our method in a typical geophysical data inverse problem, called post-stack inversion (PSI), in which the physical parameters to be estimated are the Earth's reflectivity. Our results show that the PSI based on Tsallis statistics outperforms the conventional PSI, especially in the non-Gaussian noisy-data case.

Keywords: Tsallis entropy; maximum likelihood; q -Gaussian; inverse problems; seismic imaging

1. Introduction

Recently, Tsallis nonextensive entropy [1] was applied to describe and analyze many complex phenomena that the standard statistical mechanics seem inadequate to address [2–4]. The reason for the success of the Tsallis theory comes from the fact that the observational data are better described within the Tsallis framework [5]. In this scenario, the non-Gaussian distributions that arise from the Maximum Entropy Principle (MEP) [6,7] associated with Tsallis entropy [1,8] help to explain many characteristics of complex systems, such as long-range correlations [9] and the scale-free phenomena [10,11].

The Tsallis statistics are an extension of the standard statistical mechanics and they are based on the Boltzmann–Gibbs–Shannon (BGS) entropy measure. The successful applications of Tsallis entropy [12] motivates us to explore the mathematical tool behind the Tsallis framework. In this study, we included

the Tsallis statistics in the inverse problems theory, which is an important field in applied physics [13], engineering [14], seismology [15], biomedical imaging [16], machine learning [17], and geophysics [18], among many others.

The essence of an inverse problem consists of obtaining information about physical parameters from indirect observations [18,19]. To achieve this task, the inverse problem is formulated as an optimisation problem that aims to minimize the difference between the modelled data and the observed data (residuals or errors), in which the usual misfit function is based on the least-squares distance of the residuals [19]. In addition, inverse problems are inherently ill-posed in the sense of Hadamard [20], meaning that the solution is not unique and a little perturbation in the observed data can severely impact the search for the global minimum of the misfit function [21].

In this paper, we explore a novel methodology in the geophysical data inverse problem using Tsallis entropy and a probabilistic maximum-likelihood approach. This methodology allows us to perform seismic data inversion to estimate the physical parameters of subsurface structures, known as post-stack inversion (PSI) [22]. The main objective of the PSI methodology is to estimate the reflectivity of subsurface structures by minimizing the difference between the observed seismic data and the theoretically modelled data using the misfit function [22,23]. We construct the misfit function by applying the probabilistic maximum-likelihood [19,24] method for the q-Gaussian probability distribution, which is obtained through the MEP for Tsallis entropy [8,25,26]. From a statistical viewpoint, this is equivalent to assuming that the residuals obey the q-generalized Gauss' error law [27]. The results show that the PSI based on Tsallis statistics is a powerful tool for the reconstruction of seismic images, especially in situations where the data is noisy. Our proposal outperforms the conventional PSI approach because the noise present in the data is seldom Gaussian [28].

In the next section, we briefly review the conventional PSI formulation. In Section 3, we review the Tsallis framework and derive the misfit function associated with the nonextensive entropy, as proposed in the Tsallis work [1]. This section is the core of this work and, therefore, it is dedicated to PSI formulation on the Tsallis framework. In Section 4, we describe the parameters that we used in the numerical simulations to test our methodology to a geophysical problem from oil reservoir imaging. In addition, we compare the conventional Gaussian-oriented PSI with the PSI based on q-statistics. Finally, in Section 5, we give our final remarks and we place this manuscript in a broader context.

2. Conventional PSI Formulation

The PSI is used to quantitatively infer rock-properties. In particular, it describes the subsurface structures from the reflection data (or post-stack data) recorded in a seismic survey [22]. The typical seismic experiment consists of an explosive source and a set of mechanical sensors that capture the waves reflected from the geological structures. Using a conventional approach, the PSI is formulated as a least-squares optimisation problem [22,23], whose main goal is to minimize the difference between the observed data (d^{obs}) and the modelled data (d^{mod}) using the misfit function:

$$\min_{\mathbf{r}} \phi_G(\mathbf{r}) := \frac{1}{2} \sum_{i=1}^n \left(d_i^{mod}(t, r) - d_i^{obs}(t) \right)^T \left(d_i^{mod}(t, r) - d_i^{obs}(t) \right) \quad (1)$$

In this equation, \mathbf{r} represents the reflectivity of the medium under study (model parameters) and t denotes the time. In addition, the superscript T refers to the transpose operator and n represents the size of the total reflectivity series recorded in the seismic survey.

The modelled data relies on the convolution between the seismic source, $s(t)$, and the medium reflectivity series, $r(t)$, given by [29]:

$$d^{mod}(t, r) = s(t) * r(t) = Gr(t) \tag{2}$$

where $*$ represents the convolution operator and G represents a kernel matrix computed from the seismic source, in which its action on any vector $r(t)$ will result in the convolution between $s(t)$ and $r(t)$.

The least-squares PSI (hereafter conventional PSI) determines the *maximum a posteriori* solution that can be used to estimate the model parameters. From a probabilistic maximum-likelihood viewpoint, the errors are assumed to be independent and identically distributed by a Gaussian probability distribution [18,19,24]. Consequently, the minimization of $\phi_G(\mathbf{r})$ is equivalent to the maximization of the standard Gaussian likelihood, given by:

$$\mathcal{L}_G(\mathbf{r}) \propto \exp \left[-\frac{1}{2} \sum_{i=1}^n \left(d_i^{mod}(t, r) - d_i^{obs}(t) \right)^T \left(d_i^{mod}(t, r) - d_i^{obs}(t) \right) \right] \tag{3}$$

in which the *maximum a posteriori* of \mathbf{r} is obtained by minimizing the negative log of $\mathcal{L}_G(\mathbf{r})$.

It is well-known that the MEP [6,7] for the BGS entropy

$$S(p(x)) = - \int p(x) \ln(p(x)) dx \tag{4}$$

under the constraints

$$\int p(x) dx = 1 \tag{5}$$

and

$$\int x^2 p(x) dx = 1. \tag{6}$$

straightforwardly yields a standard Gaussian distribution, where $p(x)$ is a probability density function. The constraint (5) is the normalization condition, while the constraint (6) restricts the second moment to unity. In summary, the conventional PSI maximizes the BGS entropy under the constraints (5) and (6).

3. Tsallis Framework and Seismic Inversion

Constantino Tsallis postulated a generalization of the BGS entropy [1] that describes non-equilibrium states and also describes the equilibrium thermostatics [30]. Formally, the Tsallis framework (or q -framework) is based on the q -generalized logarithm function [1,8]

$$\ln_q(x) = \frac{x^{1-q} - 1}{1 - q} \tag{7}$$

and its inverse function, the q -generalized exponential:

$$\exp_q(x) = \left[1 + (1 - q)x \right]_+^{\frac{1}{1-q}} \tag{8}$$

with

$$\ln_q(\exp_q(x)) = \exp_q(\ln_q(x)) = x \tag{9}$$

where $q \in \mathbb{R}$ is the nonextensive parameter of the Tsallis theoretical framework and $[a]_+ = \max\{0, a\}$. In the $q \rightarrow 1$ case, the expressions in (7) and (8) reduce to the usual exponential and logarithmic functions, respectively. For the continuous case, the q -entropy associated with the q -framework is given by:

$$S_q = -k_B \int p^q(x) \ln_q(p(x)) dx \tag{10}$$

where k_B is the Boltzmann constant. In the limit $q \rightarrow 1$, the q -entropy in (10) recovers BGS entropy [1]

$$S_{q \rightarrow 1} = \lim_{q \rightarrow 1} -k_B \int p^q(x) \ln_q(p(x)) dx = -k_B \int p(x) \ln(p(x)) dx \tag{11}$$

Furthermore, Tsallis entropy (10) is usually written as [1]:

$$S_q(p(x)) \equiv \frac{k_B}{q-1} \left(1 - \int p^q(x) dx \right) \tag{12}$$

which is the mathematical definition that we will use in this paper.

3.1. Maximum Tsallis Entropy and the q -Gaussian Distribution

The MEP [6,7] for Tsallis entropy (12) under the constrains of normalization (5) and the q -normalized expectation

$$\sigma_q = \frac{\int x^2 p^q(x) dx}{\int p^q(x) dx} \tag{13}$$

is associated with the following functional entropy to be maximized:

$$\mathcal{S}(p(x), \alpha_1, \alpha_2) = S_q(p(x)) - \alpha_1 \left(\int p(x) dx - 1 \right) - \alpha_2 \left(\int x^2 p^q(x) dx - \sigma_q \int p^q(x) dx \right) \tag{14}$$

where α_1 and α_2 are Lagrange multipliers.

The optimisation of Equation (14) consists of calculating the functional entropy saddle point (stationary point), $\frac{\delta \mathcal{S}}{\delta p(x)} = 0$:

$$\int \delta p(x) \left[\frac{q k_B}{1-q} p^{q-1}(x) - \alpha_1 - q \alpha_2 \left(x^2 p^{q-1}(x) - \sigma_q p^{q-1}(x) \right) \right] dx = 0 \tag{15}$$

To satisfy Equation (15) the following equation should be satisfied:

$$\frac{q k_B}{1-q} p^{q-1}(x) - \alpha_1 - q \alpha_2 \left(x^2 p^{q-1}(x) - \sigma_q p^{q-1}(x) \right) = 0 \tag{16}$$

which straightforwardly yields the q -probability function:

$$p(x) = \left[\left(\frac{q}{\alpha_1} \right) \left(k_B - (1-q)\alpha_2(x^2 - \sigma_q) \right) \right]^{\frac{1}{1-q}} \tag{17}$$

To rewrite the probability function in (17) using q -generalized functions, consider the expression

$$p(x) = [1 + (1-q)A_q]^{\frac{1}{1-q}} [1 + (1-q)(-B_q)x^2]^{\frac{1}{1-q}} \tag{18}$$

where A_q and B_q depend on q and the Lagrange multipliers α_1 and α_2 . After some algebra, we derive the following expression for the q -probability distribution:

$$p(x) = \exp_q(A_q) \exp_q(-B_q x^2) \tag{19}$$

with

$$A_q = \frac{q}{\alpha_1(1-q)} \left[\frac{k_B}{1-q} + \alpha_2 \sigma_q - \frac{\alpha_1}{q} \right] \quad \text{and} \quad B_q = \frac{\alpha_2}{k_B + (1-q)\alpha_2 \sigma_q} \tag{20}$$

In this context, Equation (19) corresponds to the q -Gaussian probability distribution, which is used in the geophysical inversion in this manuscript. We notice that no q -Gaussian exists in the $q \geq 3$ case because it does not satisfy the normalization condition, Equation (5).

3.2. The q -misfit Function

Assuming that all residual data ($\mathbf{x} = x_1, x_2, \dots, x_n$) are independent and follow a q -Gaussian distribution (20), the q -misfit function is obtained by the log-likelihood (we take the logarithm for convenience):

$$-\ln(\mathcal{L}_q(\mathbf{x})) = -\ln\left(\prod_{i=1}^n p_q(x_i)\right) = -\ln\left(\prod_{i=1}^n \exp_q(A_q) \exp_q(-B_q x_i^2)\right) \tag{21}$$

This formula can be written as:

$$-\ln(\mathcal{L}_q(\mathbf{x})) = -\ln[\exp_q(A_q)] - \frac{1}{1-q} \sum_{i=1}^n \ln[1 + (1-q)(-B_q) x_i^2]_+ \tag{22}$$

We notice that minimize Equation (22) is equivalent to minimizing the function:

$$\phi_q(\mathbf{x}) = \frac{1}{q-1} \sum_{i=1}^n \ln[1 + (q-1)B_q x_i^2]_+ \tag{23}$$

which in our geophysical inversion problem is formulated as the following optimisation problem:

$$\min_{\mathbf{r}} \phi_q(\mathbf{r}) := \frac{1}{q-1} \sum_{i=1}^n \ln[1 + (q-1)B_q (Gr_i(t) - d_i^{obs}(t))^T (Gr_i(t) - d_i^{obs}(t))]_+ \tag{24}$$

In addition, we consider that $B_q = 1/(3-q)$, as proposed in [25] for $1 < q < 3$. We notice that the conventional misfit function is recovered from Equation (24) for the limit $q \rightarrow 1$. Using the L'Hôpital's rule:

$$\begin{aligned} \lim_{q \rightarrow 1} \phi_q(\mathbf{r}) &= \lim_{q \rightarrow 1} \sum_{i=1}^n \frac{\frac{\partial}{\partial q} \left\{ \ln \left[1 + \left(\frac{q-1}{3-q} \right) (Gr_i(t) - d_i^{obs}(t))^T (Gr_i(t) - d_i^{obs}(t)) \right]_+ \right\}}{\frac{\partial(q-1)}{\partial q}} \\ \lim_{q \rightarrow 1} \phi_q(\mathbf{r}) &= \lim_{q \rightarrow 1} \sum_{i=1}^n \frac{\left(\frac{1}{3-q} \right)^2 (Gr_i(t) - d_i^{obs}(t))^T (Gr_i(t) - d_i^{obs}(t))}{\left[1 + (q-1)B_q (Gr_i(t) - d_i^{obs}(t))^T (Gr_i(t) - d_i^{obs}(t)) \right]_+} \\ \lim_{q \rightarrow 1} \phi_q(\mathbf{r}) &= \frac{1}{2} \sum_{i=1}^n (Gr_i(t) - d_i^{obs}(t))^T (Gr_i(t) - d_i^{obs}(t)) = \phi_G(\mathbf{r}) \end{aligned} \tag{25}$$

quod erat demonstrandum.

Therefore, the optimisation problem formulated in (24) becomes:

$$\min_{\mathbf{r}} \phi_q(\mathbf{r}) := \frac{1}{q-1} \sum_{i=1}^n \ln \left[1 + \left(\frac{q-1}{3-q} \right) (Gr_i(t) - d_i^{obs}(t))^T (Gr_i(t) - d_i^{obs}(t)) \right]_+ \tag{26}$$

We call the function $\phi_q(\mathbf{r})$ a q -misfit function and we call q -PSI the optimisation problem employing the q -misfit function. It is worth noting that the term in brackets is always positive in the $1 < q < 3$ case. However, for $q < 1$, $\phi_q(\mathbf{r})$ is given by Equation (26) if $|Gr_i(t) - d_i^{obs}(t)| < [(3 - q)/(1 - q)]^{1/2}$ [25] and zero otherwise. To simplify the notation, henceforth operation $[a]_+$ will be implicit in the equations.

3.3. PSI as a Local Optimisation Problem

As mentioned earlier, the PSI is usually solved using local optimisation methods. Starting from a reflectivity series \mathbf{r}_0 (initial model), the optimisation problem is solved iteratively by updating the model according to:

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{H}^{-1} \nabla_{\mathbf{r}} \phi(\mathbf{r}_k) \tag{27}$$

where α_k is the steplength [31] of the k -th iteration, \mathbf{H} is the Hessian matrix, and $\nabla_{\mathbf{r}} \phi(\mathbf{r}_k) = \partial \phi(\mathbf{r}_k) / \partial \mathbf{r}$ is the gradient of the misfit function $\phi(\mathbf{r}_k)$. Therefore, in addition to the misfit function, we need to determine their corresponding gradient and Hessian.

In the q -PSI case, the gradient is given by:

$$\nabla_{\mathbf{r}} \phi_q(\mathbf{r}) = \sum_{i=1}^n G^T \left[\frac{2 (Gr_i(t) - d_i^{obs}(t))}{[3 - q + (q - 1) (Gr_i(t) - d_i^{obs}(t))^T (Gr_i(t) - d_i^{obs}(t))]^{1/2}} \right] \tag{28}$$

For comparison, the gradient of the conventional misfit function is given by:

$$\nabla_{\mathbf{r}} \phi_G(\mathbf{r}) = \sum_{i=1}^n G^T (Gr_i(t) - d_i^{obs}(t)) \tag{29}$$

wherein it is easy to see that at the limit $q \rightarrow 1$, the q -misfit function gradient—Equation (28)—becomes the conventional misfit gradient function—Equation (29).

By comparing Equations (28) and (29), we can see that the gradient of the q -misfit function is weighted by the factor $2/[3 - q + (q - 1)(Gr_i(t) - d_i^{obs}(t))^T (Gr_i(t) - d_i^{obs}(t))]$. Consequently, the large residual data are dampened in the model update process, Equation (27), which makes the q -PSI less sensitive to large errors than the conventional PSI approach.

In our numerical study, because of the fast convergence, we perform the optimisation with the quasi-Newton method *limited-memory* Broyden-Fletcher-Goldfarb-Shanno (*l*-BFGS) [31], in which the inverse of the Hessian is obtained through an approximation computed from previous gradients [32]. Thus, we do not explicitly calculate the second derivative of the misfit function, which saves a lot of memory and reduces the computational cost. This trick is valuable because geophysics problems typically have a large number of variables (number of model parameters).

4. Numerical Results

We illustrate our method using a portion of the Marmousi2 model, which is based on the geology of the Kwanza Basin region (Angola) [33,34]. The acoustic impedance distribution of this model is shown in Figure 1a. The impedance model consists of 2000 and 400 grid cells in the vertical and horizontal directions, respectively (800,470 total grid points). From this impedance model, we obtained the reflectivity model that will be used as the true model in this study, which is depicted in Figure 1b. The mathematical relationship between acoustic impedance (Z) and reflectivity (r) is given by [29,35]:

$$r(t) = \frac{1}{2} \frac{d}{dt} \left[\ln (Z(t)) \right] \tag{30}$$

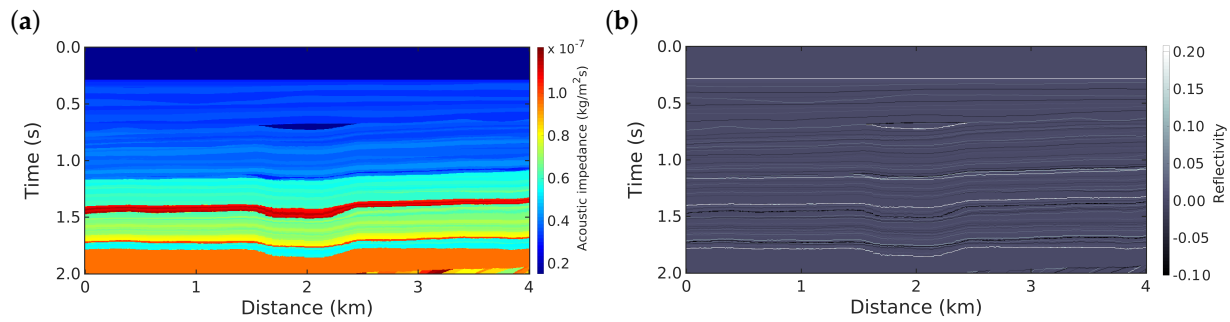


Figure 1. (a) The portion of the Marmousi2 acoustic impedance model in the depth-time domain. (b) Reflectivity model (true model), which is extracted from the impedance model in Figure 1a using Equation (30).

A Ricker wavelet [36] is considered to be the source signature, which is defined by:

$$s(t) = \left(1 - 2\pi^2\nu_p^2t^2\right) \exp\left(-\pi^2\nu_p^2t^2\right) \tag{31}$$

where ν_p is the most energetic frequency. In this study, we considered $\nu_p = 55$ Hz.

For the optimisation solver, we use the *l*-BFGS algorithm [32] to minimize the misfit function for both methods, in which the stop criteria is a tolerance ϵ in the gradient norm or if the step length does not satisfy the Wolfe conditions [31]. The gradient norm is $\|\nabla_r\phi(\mathbf{r})\| < \epsilon = 10^{-12}$. If one of these criteria is satisfied, then the inversion process stops. Figure 2b shows the initial model that was used for all of the numerical simulations.

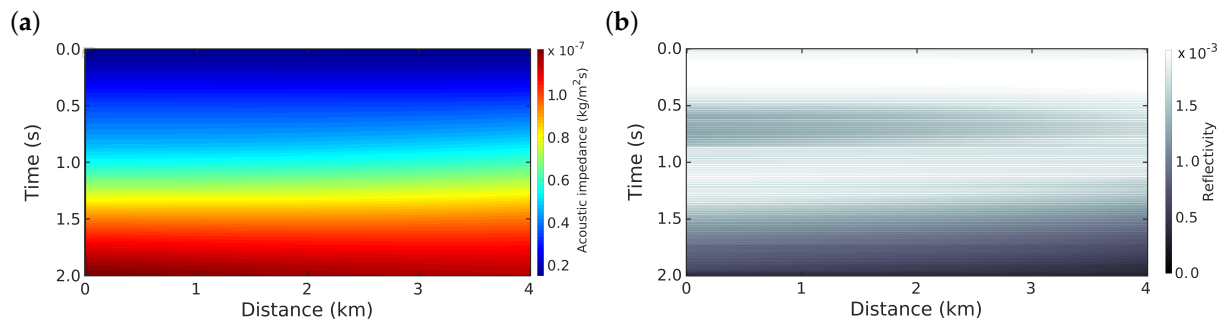


Figure 2. (a) Initial acoustic impedance model in the depth-time domain, and its (b) reflectivity model (Initial model).

We considered two scenarios: in the first scenario, an ideal case with noiseless data is considered to demonstrate the good working order of the algorithms. Then, in the second scenario, we considered a dataset with spikes, to simulate a non-Gaussian noise. The spikes were added over 1% of the samples (chosen randomly using a uniform distribution) by rescaling the signal amplitudes by a factor of 15β , in which β follows a standard Gaussian distribution. For each scenario, we performed 16 inversions, the first follows the conventional approach, Equation (1) and for the others we used the *q*-PSI with $q = 0.1, 0.3, 0.5, \dots, 2.9$, Equation (26). The observed data for these two numerical experiments are depicted in Figure 3. We emphasize that the initial model is the same for all numerical experiments.

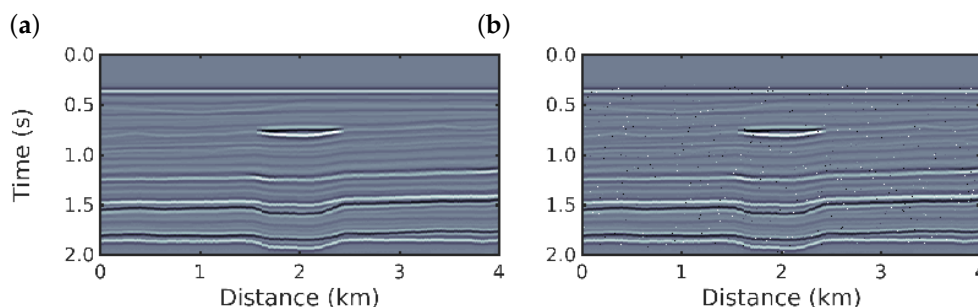


Figure 3. Observed post-stack data: (a) noiseless data (first scenario); and (b) spiky-noise data (second scenario).

The inversion results for the first scenario using conventional PSI and q -PSI with $q = 0.1, 0.3, 0.5, \dots, 2.9$ are shown in Figure 4. The results for both methods show a good image resolution of the subsurface, we then compare the results with the true reflectivity model (Figure 1b).

In the second scenario in which non-Gaussian noise is considered, the conventional PSI fails to obtain a good reconstruction of the reflectivity model, as depicted in Figure 5e. In contrast, the q -PSI seems to obtain results that are very close to the true reflectivity model as the q value increases in the $1 < q < 3$ case; that is, as the deviation from Gaussian behaviour is larger, as depicted in Figures 5f–o for $q = 1.1, 1.3, 1.5, 1.7, \dots, 2.9$, respectively. Although the q -PSI results for $q \leq 1.3$ show artefacts (see Figures 4a–d,f,g), they are still superior to the conventional PSI result (Figure 4e).

In addition, we compute three statistical measures to quantitatively compare the PSI results with the true reflectivity model: in the first measure, we computed the normalized root-mean-square (NRMS), which is defined as:

$$NRMS = \left[\frac{\sum_{i=1}^n (r_i^{true} - r_i^{inv})^2}{\sum_{i=1}^n (r_i^{true})^2} \right]^{1/2} \tag{32}$$

where r^{true} corresponds to the true reflectivity model and r^{inv} is the inversion result. A NRMS close to 0 means low error. The other two measures are similarity measures: Pearson’s coefficient (R) [37] and the structural similarity (SSIM) index [38]. Pearson’s coefficient measures the linear relationship between the reconstructed model and the true model, while the SSIM is a quality index of similarity between images. Both R and SSIM measures vary between -1 (anti correlation/similarity) to 1 (perfect correlation/similarity). The statistical measures for the two scenarios are summarized in Table 1.

In the analysis of the results presented in Table 1, we observe that the q -PSI inversion for the q -value around 2.1 shows the best result with the lowest NRMS and the highest similarity (highest R and SSIM) for all scenarios. However, we emphasize that our proposal with $q > 1.3$ also produces good results, is robust to non-Gaussian noise, and shows low sensitivity to outliers in the dataset, which are represented here by the spikes.

Furthermore, the q -PSI requires the same or fewer number of iterations than the conventional PSI to converge, as illustrated in Figure 6. In the first scenario, the convergence of both methods is quite similar, as depicted in Figures 6a and 7a. However, in the second scenario, the q -PSI with $q > 1.3$ and $q < 1$ converges more quickly than the others, in which the decay of the convergence curve slopes more as the q -value increases, as depicted in Figures 6b and 7b.

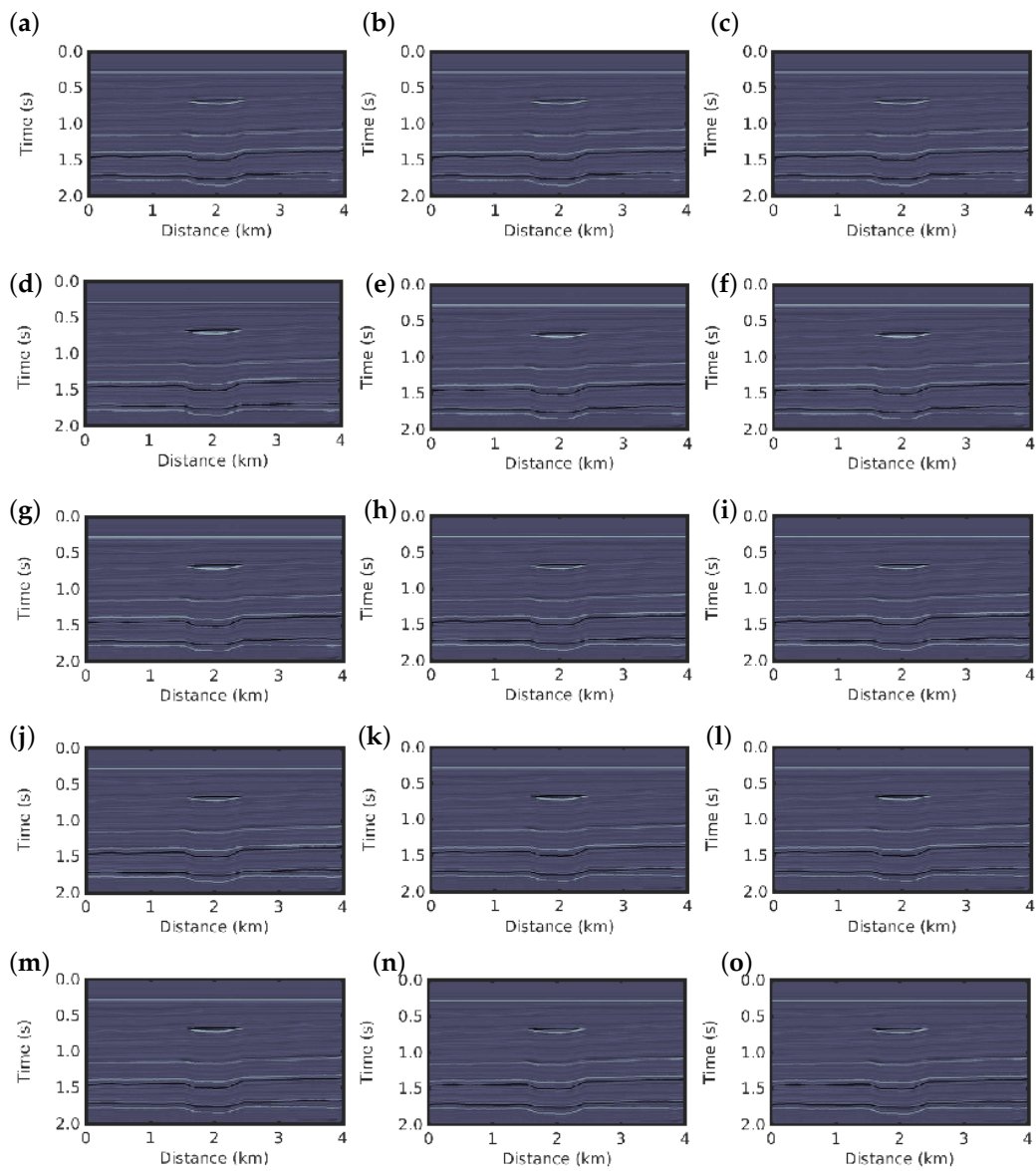


Figure 4. Inversion results for the first scenario: reflectivity models for the q -PSI with (a) $q = 0.3$, (b) $q = 0.5$, (c) $q = 0.7$, (d) $q = 0.9$, (e) conventional PSI, and q -PSI with (f) $q = 1.1$, (g) $q = 1.3$, (h) $q = 1.5$, (i) $q = 1.7$, (j) $q = 1.9$, (k) $q = 2.1$, (l) $q = 2.3$, (m) $q = 2.5$, (n) $q = 2.7$, (o) $q = 2.9$.

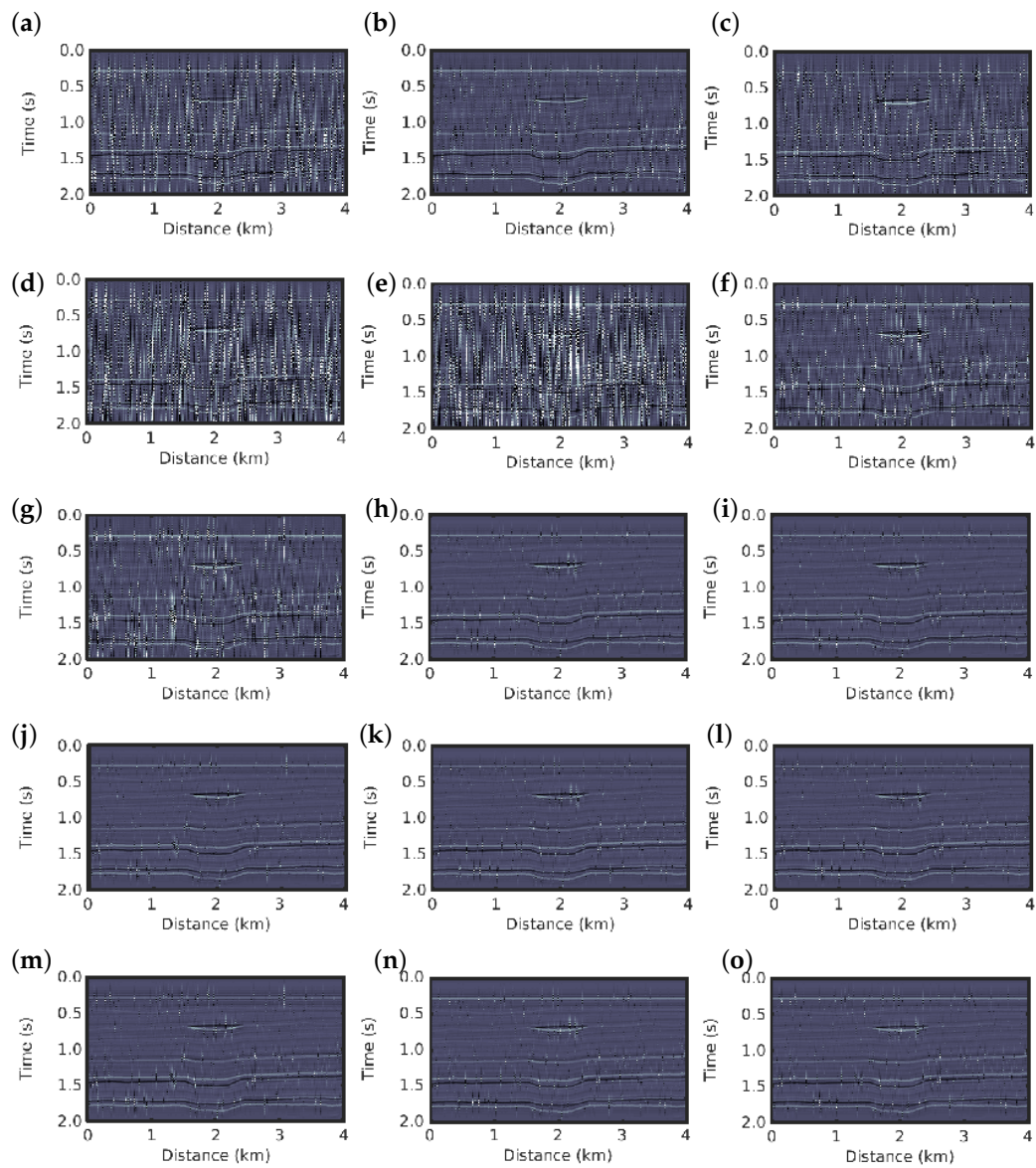


Figure 5. Inversion results for the second scenario: reflectivity models for the q -PSI with (a) $q = 0.3$, (b) $q = 0.5$, (c) $q = 0.7$, (d) $q = 0.9$, (e) conventional PSI, and q -PSI with (f) $q = 1.1$, (g) $q = 1.3$, (h) $q = 1.5$, (i) $q = 1.7$, (j) $q = 1.9$, (k) $q = 2.1$, (l) $q = 2.3$, (m) $q = 2.5$, (n) $q = 2.7$, (o) $q = 2.9$.

Table 1. Main statistics of the PSI results.

Strategy	First Scenario			Second Scenario		
	NRMS	R	SSIM	NRMS	R	SSIM
Our proposal ($q = 0.1$)	0.8369	0.8282	0.8277	2.4821	0.4302	0.2756
Our proposal ($q = 0.3$)	0.8366	0.8293	0.8288	2.7015	0.4128	0.2523
Our proposal ($q = 0.5$)	0.8365	0.8287	0.8282	1.3092	0.6187	0.6215
Our proposal ($q = 0.7$)	0.8367	0.8292	0.8286	1.3186	0.6129	0.6152
Our proposal ($q = 0.9$)	0.8369	0.8295	0.8289	4.5209	0.3395	0.1495
Conventional PSI ($q \rightarrow 1.0$)	0.8373	0.8292	0.8286	6.5366	0.3118	0.1222
Our proposal ($q = 1.1$)	0.8370	0.8296	0.8290	2.0517	0.5514	0.4328
Our proposal ($q = 1.3$)	0.8371	0.8293	0.8288	2.1844	0.5362	0.4085
Our proposal ($q = 1.5$)	0.8366	0.8294	0.8288	1.0115	0.7015	0.6934
Our proposal ($q = 1.7$)	0.8374	0.8293	0.8287	1.0057	0.7040	0.6971
Our proposal ($q = 1.9$)	0.8366	0.8289	0.8284	0.9896	0.7083	0.7037
Our proposal ($q = 2.1$)	0.8376	0.8293	0.8287	0.9884	0.7085	0.7041
Our proposal ($q = 2.3$)	0.8371	0.8296	0.8290	0.9966	0.7074	0.7018
Our proposal ($q = 2.5$)	0.8373	0.8290	0.8284	1.0370	0.6951	0.6833
Our proposal ($q = 2.7$)	0.8373	0.8293	0.8287	1.0051	0.7040	0.6971
Our proposal ($q = 2.9$)	0.8376	0.8280	0.8274	1.0178	0.7035	0.6946

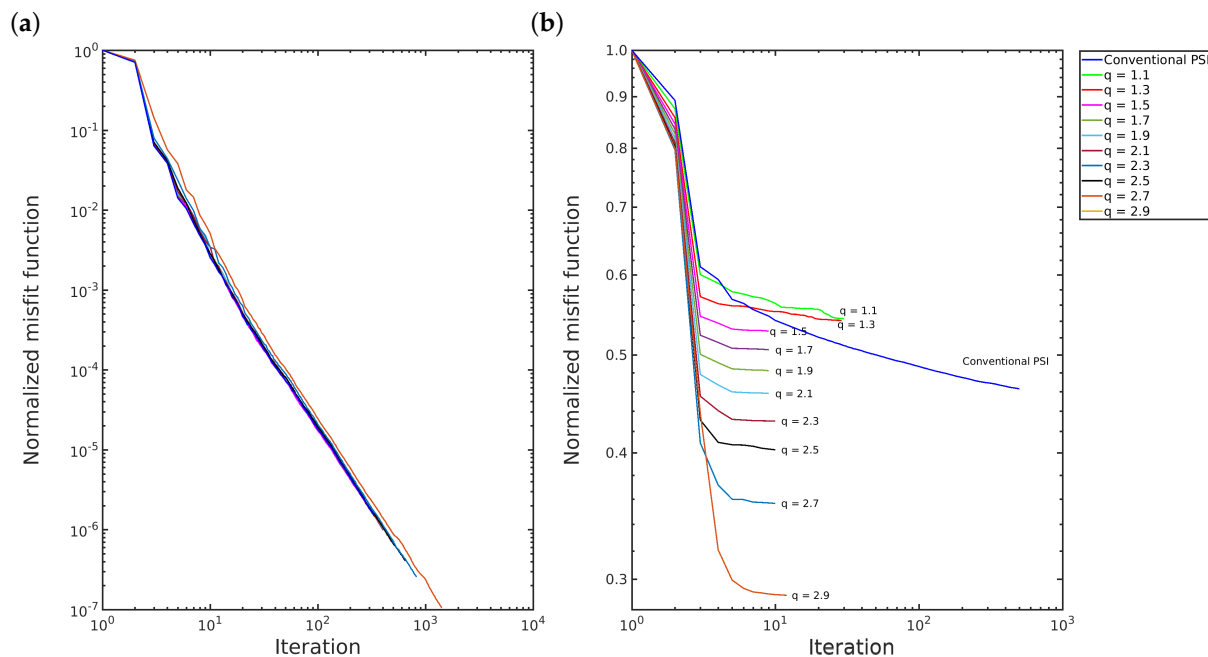


Figure 6. Convergence: (a) first scenario; and (b) second scenario for $1 < q < 3$ case.

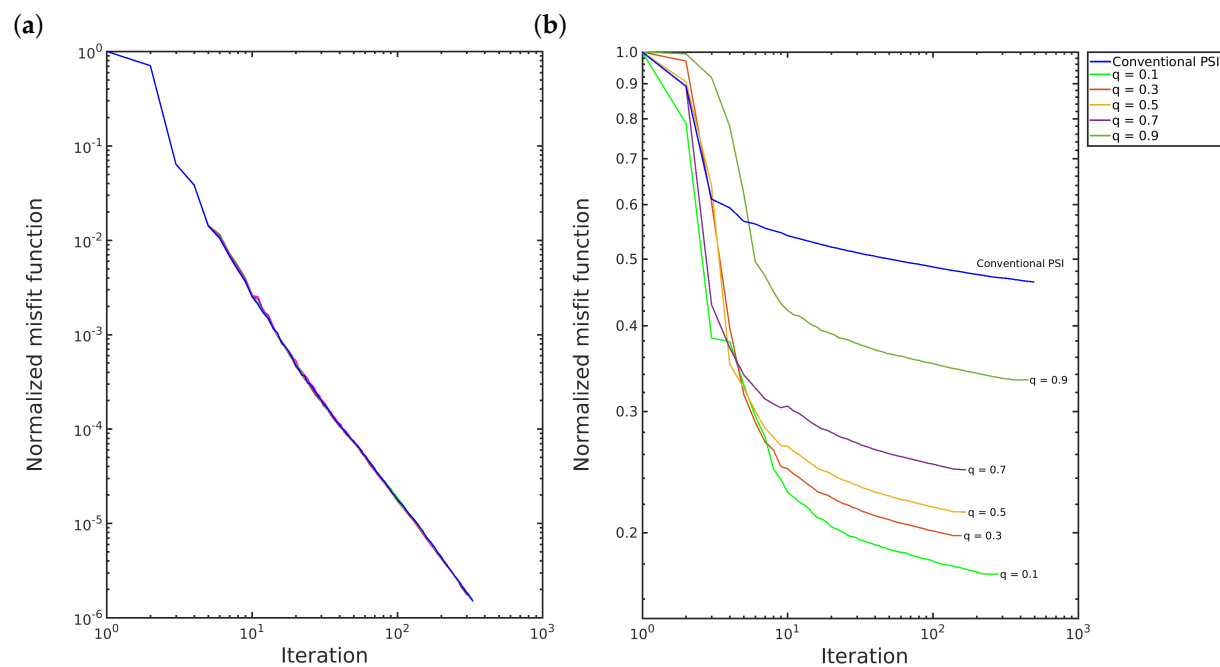


Figure 7. Convergence: (a) first scenario; and (b) second scenario for $0 < q < 1$ case.

5. Conclusions

We investigate the portability of Tsallis nonextensive entropy to compute physical parameters for a geophysical data inverse problem. In this context, we have presented a robust misfit function to mitigate the seismic inversion sensitivity to noise in the reconstruction of subsurface reflectivity models. Given that we employ a PSI method for the geophysical inversion, we refer to our proposal with the abbreviation q -PSI—in reference to Tsallis statistics (or q -statistics).

To illustrate the stability and robustness of our proposal, we considered two numerical experiments: the first experiment is an ideal noiseless data and in the second experiment we added spikes in 1% of the data simulating a non-Gaussian noise. The numerical experiments demonstrate that the q -PSI outperforms the conventional PSI and is a promising tool for seismic exploration, especially for use with non-Gaussian noisy-data. Furthermore, the inclusion of Tsallis statistics in solving the inverse problem under study accelerates the algorithmic convergence of the inversion method, especially for the $1.3 < q < 3.0$ cases.

The choice of the q -parameter is an important aspect of our methodology and, after a statistical analysis of the PSI results, we conclude that q -values around 2.1 produce more reliable estimates of the physical parameters. To conclude, the q -PSI is a valuable tool in exploration geophysics. Furthermore, we believe that our proposal can be extended to other inverse problems that require robust methods.

Author Contributions: Conceptualization, I.P.d.L., S.L.E.F.d.S., G.C., and J.M.d.A.; methodology, I.P.d.L., S.L.E.F.d.S., G.C., and J.M.d.A.; software, S.L.E.F.d.S.; validation, S.L.E.F.d.S.; formal analysis, I.P.d.L., S.L.E.F.d.S., G.C., and J.M.d.A.; investigation, I.P.d.L., S.L.E.F.d.S., G.C., and J.M.d.A.; writing—original draft preparation, I.P.d.L., S.L.E.F.d.S., G.C., and J.M.d.A.; visualization, I.P.d.L., S.L.E.F.d.S., G.C., and J.M.d.A.; supervision, G.C., and J.M.d.A.; writing—review and editing, S.L.E.F.d.S., G.C., and J.M.d.A. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001*.

Acknowledgments: We thank the anonymous reviewers for the time and effort dedicated to carefully reading the original article. Their many comments and suggestions greatly improved the presentation and clarity of this article.

G.C. acknowledges *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) for support through productivity fellowship (grant no. 304421/2015-4). J.M.d.A. thanks CNPq for his productivity fellowship (grant no. 313431/2018-3).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BGS	Boltzmann-Gibbs-Shannon
MEP	Maximum entropy principle
NRMS	Normalized root mean square
PSI	Post-stack inversion
SSIM	Structural similarity

References

1. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [CrossRef]
2. Bekenstein, J.D. Some comments on Boltzmann-Gibbs statistical mechanics. *Phys. Rev. D* **1973**, *7*, 2333–2346. [CrossRef]
3. Tsallis, C. Some comments on Boltzmann-Gibbs statistical mechanics. *Chaos Soliton. Fract.* **1995**, *6*, 539–559. [CrossRef]
4. Tsallis, C. Non-extensive thermostatistics: Brief review and comments. *Phys. A Stat. Mech. Appl.* **1995**, *221*, 277–290. [CrossRef]
5. Picoli Jr., S.; Mendes, R.S.; Malacarne, L.C.; Santos, R.P.B. q-distributions in complex systems: A brief review. *Braz. J. Phys.* **2009**, *39*, 468–474. [CrossRef]
6. Jaynes, E.T. Information Theory and Statistical Mechanics. *Phys. Rev.* **1957**, *106*, 620–630. [CrossRef]
7. Jaynes, E.T. Information Theory and Statistical Mechanics II. *Phys. Rev.* **1957**, *108*, 171–190. [CrossRef]
8. Tsallis, C.; Mendes, R.; Plastino, A. The role of constraints within generalized nonextensive statistics. *Phys. A Stat. Mech. Appl.* **1998**, *261*, 534–554. [CrossRef]
9. Du, J. Nonextensivity in nonequilibrium plasma systems with Coulombian long-range interactions. *Phys. Lett. A* **2004**, *329*, 262–267. [CrossRef]
10. Abe, S.; Suzuki, N. Scale-free statistics of time interval between successive earthquakes. *Phys. A Stat. Mech. Appl.* **2005**, *350*, 588–596. [CrossRef]
11. Silva, R.; França, G.S.; Vilar, C.S.; Alcaniz, J.S. Nonextensive models for earthquakes. *Phys. Rev. E* **2006**, *73*, 026102. [CrossRef] [PubMed]
12. Nonextensive Statistical Mechanics and Thermodynamics: Bibliography. Available online: <http://tsallis.cat.cbpf.br/TEMUCO.pdf> (accessed on 19 March 2020).
13. Ayón-Beato, E.; García, A.; Mansilla, R.; Terrero-Escalante, C.A. Stewart-Lyth inverse problem. *Phys. Rev. D* **2000**, *62*, 103–112. [CrossRef]
14. Huang, J.; Supaongprapa, T.; Terakura, I.; Wang, F.; Ohnishi, N.; Sugie, N. A model-based sound localization system and its application to robot navigation. *Rob. Aut. Syst.* **1999**, *27*, 199–209. [CrossRef]
15. da Silva, S.L.E.F.; Juliã, J.; Bezerra, F.H.R. Deviatoric Moment Tensor Solutions from Spectral Amplitudes in Surface Network Recordings: Case Study in São Caetano, Pernambuco, Brazil. *Bull. Seism. Soc. Am.* **2017**, *107*, 1495–1511. [CrossRef]
16. Bertero, M.; Piana, M. *Inverse Problems in Biomedical Imaging: Modeling and Methods of Solution*; Springer-Verlag Italia: Milano, Italy, 2006.
17. Prato, M.; Zanni, L. Inverse problems in machine learning: An application to brain activity interpretation. *J. Phys. Conf. Ser.* **2008**, *135*, 012085. [CrossRef]
18. Menke, W. *Geophysical Data Analysis: Discrete Inverse Theory*, 3rd ed.; Academic Press: London, UK, 2012.
19. Tarantola, A. *Inverse Problem Theory and Methods for Model Parameter Estimation*; Society for Industrial and Applied Mathematics (SIAM): Philadelphia, PA, USA, 2005.

20. Hadamard, J. Sur les problèmes aux dérivés partielles et leur signification physique. *Princ. Univ. Bull.* **1902**, *13*, 49–52.
21. Wang, Y. *Seismic Inversion: Theory and Applications*; Wiley Blackwell: Hoboken, NJ, USA, 2016.
22. Yilmaz, O. *Seismic Data Analysis: Processing, Inversion and Interpretation of Seismic Data*; Society of Exploration (SEG): Tulsa, OK, USA, 2000.
23. Karlı, H.; Güney, R.; Senkaya, M. Post-stack high-resolution deconvolution using Cauchy norm regularization with FX filter weighting. *Arab. J. Geosci.* **2017**, *10*, 551. [[CrossRef](#)]
24. Ferrari, D.; Yang, Y. Maximum likelihood estimation. *Ann. Stat.* **2010**, *38*, 753–783. [[CrossRef](#)]
25. Prato, D.; Tsallis, C. Nonextensive foundation of levy distributions. *Phys. Rev. E* **1999**, *60*, 2398. [[CrossRef](#)]
26. Sato, A.-H. q-gaussian distributions and multiplicative stochastic processes for analysis of multiple financial time series. *J. Phys. Conf. Ser.* **2010**, *201*, 012008. [[CrossRef](#)]
27. Suyari, H.; Tsukada, M. Law of error in Tsallis statistics. *IEEE Trans. Inf. Theory* **2005**, *51*, 753–757. [[CrossRef](#)]
28. Amundsen, L. Comparison of the least-squares criterion and the Cauchy criterion in frequency-wavenumber inversion. *Geophysics* **1991**, *56*, 2027–2035. [[CrossRef](#)]
29. Russell, B. *Introduction to Seismic Inversion Methods*; Society of Exploration (SEG): Tulsa, OK, USA, 1988.
30. Tirnakli, U.; Borges, E. The standard map: From Boltzmann-Gibbs statistics to Tsallis statistics. *Sci. Rep.* **2016**, *6*, 23644. [[CrossRef](#)] [[PubMed](#)]
31. Nocedal, J.; Wright, S.J. *Numerical Optimisation*, 2nd ed.; Springer: New York, NY, USA, 2006.
32. Byrd, R.H.; Lu, P.; Nocedal, J.; Zhu, C. A limited memory algorithm for bound constrained optimization. *J. Sci. Comp.* **1995**, *16*, 1190–1208. [[CrossRef](#)]
33. Versteeg, R. The Marmousi experience: Velocity model determination on a synthetic complex data set. *Lead. Edge* **1994**, *13*, 927–936. [[CrossRef](#)]
34. Martin, G.S.; Wiley, R.; Marfurt, K.J. Marmousi2: An elastic upgrade for Marmousi. *The Lead. Edge* **2006**, *25*, 156–166. [[CrossRef](#)]
35. Stolt, R.H.; Weglein, A.B. Migration and inversion of seismic data. *Geophysics* **1985**, *12*, 2458–2472. [[CrossRef](#)]
36. Ricker, N. Further developments in the wavelet theory of seismogram structure. *Bull. Seism. Soc. Am.* **1943**, *3*, 197–228.
37. Pearson, K.; Henrici, O.M.F.E. VII. Mathematical contributions to the theory of evolution.; III. Regression, heredity, and panmixia. *Phil. Trans. R. Soc. Lond.* **1896**, *187*, 253–318.
38. Zhou, W.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Im. Proc.* **2004**, *13*, 600–612.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).