

Article

Geometric Interpretation of Errors in Multi-Parametrical Fitting Methods Based on Non-Euclidean Norms

George Livadiotis

Southwest Research Institute, Space Science & Engineering, San Antonio, TX 78238, USA; glivadiotis@swri.edu; Tel.: +1-210-274-4028

Received: 29 September 2019; Accepted: 25 October 2019; Published: 29 October 2019



Abstract: The paper completes the multi-parametrical fitting methods, which are based on metrics induced by the non-Euclidean L_q -norms, by deriving the errors of the optimal parameter values. This was achieved using the geometric representation of the residuals sum expanded near its minimum, and the geometric interpretation of the errors. Typical fitting methods are mostly developed based on Euclidean norms, leading to the traditional least-square method. On the other hand, the theory of general fitting methods based on non-Euclidean norms is still under development; the normal equations provide implicitly the optimal values of the fitting parameters, while this paper completes the puzzle by improving understanding the derivations and geometric meaning of the optimal errors.

Keywords: fitting; non-Euclidean norm; fitting errors; optimization

1. Introduction

The keys to evaluating an experimental result—e.g., compare it with the result anticipated by theories—require first the right selection of potential statistical tools and techniques for correctly processing and analyzing this result. This “processing and analyzing” involves two general types of approximation problems: One problem concerns a function fitting to given set of data. The other problem arises when a function is given analytically by an explicit mathematical type but we would like to find an alternative function with simpler form.

Let $V(x; p_1, p_2, \dots, p_n)$, with $x \in D \subseteq \mathfrak{X}$ and $(p_1, p_2, \dots, p_n) \in \{D_{p_1} \otimes D_{p_2} \otimes \dots \otimes D_{p_n}\} \subseteq \mathfrak{X}^n$, denote a multi-parametrical approximating function [1–6], symbolized as $V(x; \{p_k\})$, for short.

The widely used, traditional fitting method of least squares involves minimizing the sum of the squares of the residuals, i.e., the squares of the differences between the function $f(x)$ and the approximating function that represents the statistical model, $V(x)$. However, the least-square method is not unique. For instance, the absolute deviations minimization can also be applied. Generally, as soon as the desired norm of the metric space is given, the respective method of deviations minimization is defined. The least-square method is based on the Euclidean norm, while the alternative absolute deviations method is based on the uniform or Taxicab norm. In general, an infinite number of fitting methods can be defined, based on the metric space induced by the L_q -norm; this case is studied here in detail.

Given the metric induced by the L_q -norm, the functional of the total L_q -normed residuals [7–12], noted also as total deviations (TD), between the fixed $f(x)$ and the approximating $V(x; \{p_k\})$ functions in the domain D , is given by:

$$TD_q(\{p_k\})^q = \int_{x \in D} |V(x; \{p_k\}) - f(x)|^q dx. \quad (1)$$

The functional of total deviations, $TD_q(\{p_k\})^q$, is expanded (Taylor series) near its local minimum:

$$TD_q(\{p_k\})^q = A_0(q) + \sum_{k_1, k_2=1}^n A_{2, k_1 k_2}(q) \cdot (p_{k_1} - p_{k_1}^*) \cdot (p_{k_2} - p_{k_2}^*) + O(|p_k - p_k^*|^3), \tag{2a}$$

where

$$A_0(q) = TD_q(\{p_k^*\})^q, \tag{2b}$$

is the total deviation function at its global minimum, while

$$A_{2, k_1 k_2}(q) = \frac{1}{2} \cdot \frac{\partial^2}{\partial p_{k_1} \partial p_{k_2}} TD_q(\{p_k\})^q \Big|_{\{p_k\}=\{p_k^*\}}, \tag{2c}$$

is the Hessian matrix at this minimum, where all the components are positive, i.e., $A_0, A_{2, k_1 k_2} \geq 0, \forall k_1, k_2 = 1, 2, \dots, n$.

By expanding the approximating function $V(x; \{p_k\})$ near the TD's minimum, [7] showed the following equations:

$$A_0(q) = \int_{x \in D} |u|^q dx, \tag{3a}$$

and

$$A_{2, k_1 k_2}(q) = \delta_{1q} \cdot \gamma_{k_1 k_2} + \frac{q}{2} \cdot \int_{x \in D} \left[|u|^{q-1} \cdot \text{sgn}(u) \cdot \frac{\partial^2 V(x; \{p_k^*\})}{\partial p_{k_1} \partial p_{k_2}} + (q-1) \cdot |u|^{q-2} \cdot \frac{\partial V(x; \{p_k^*\})}{\partial p_{k_1}} \cdot \frac{\partial V(x; \{p_k^*\})}{\partial p_{k_2}} \right] dx, \tag{3b}$$

where

$$\gamma_{k_1 k_2} \equiv \sum_{\substack{\forall i: \\ u(x_i) = 0}} \frac{1}{|u'(x_i)|} \cdot \frac{\partial V(x_i; \{p_k^*\})}{\partial p_{k_1}} \cdot \frac{\partial V(x_i; \{p_k^*\})}{\partial p_{k_2}}. \tag{3c}$$

The normal equations are given by

$$\int_{x \in D} |u|^{q-1} \cdot \text{sgn}(u) \cdot \frac{\partial V(x; \{p_k^*\})}{\partial p_l} dx = 0, \forall l = 1, 2, \dots, n. \tag{4}$$

where we set $u = u(x) \equiv V(x; \{p_k^*\}) - f(x)$ for short.

The purpose of this paper is to present the geometric interpretation of the errors of the optimal parameter values, derived from a multi-parametrical fitting, based on a metric induced by the non-Euclidean L_q -norm. In Section 2, we derive the smallest possible value of the variation of the total deviations from its minimum, δTD , also called, the error of the total deviations value. In Section 3, we describe the geometric interpretation of the errors of the optimal parameter values, while in Section 4, we use this geometry to derive the exact equations that provide these errors. In Section 5, we apply the developed formulation for the 1-dim and 2-dim cases. Finally, Section 6 summarizes the conclusions.

2. The Error of the Total Deviation Values

The total deviations functional, $TD_a(\{p_k\})^q$, has a minimum value $A_0(q)$. The difference between these functionals cannot be arbitrarily small. Here we derive the smallest possible value of the variation of the total deviations from its minimum, δTD , also called, the error of the total deviations value.

First, we mention that the transition of the continuous to the discrete way for describing the values of x , can be realized as follows:

$$\int_a^b (\dots) dx = \lim_{N \rightarrow \infty} \frac{b-a}{N} \cdot \sum_{i=1}^N (\dots), \tag{5}$$

while the expression of the total deviations is given by

$$TD_q(\{p_k\})^q = \int_{x \in D} |V(x; \{p_k\}) - f(x)|^q dx \approx \frac{L}{N} \cdot \sum_{i=1}^N |V(x_i; \{p_k\}) - f(x_i)|^q = x_{res} \cdot \sum_{i=1}^N |V(x_i; \{p_k\}) - f(x_i)|^q, \tag{6}$$

for large values of N , where L is the total length of the domain D , and the resolution of x -values is $x_{res} = L/N$. In the discrete case, it is sufficient to express the total deviations simply by

$$TD_q(\{p_k\})^q = \sum_{i=1}^N |u_i|^q, \tag{7}$$

where we set $u_i \equiv u(x_i; \{p_k\}) \equiv V(x_i; \{p_k\}) - f(x_i)$, $V_i \equiv V(x_i; \{p_k\})$ and $y_i \equiv f(x_i)$.

Then, we calculate the error of the total deviations values, δTD , near the local minimum of $TD_q(\{p_k\})^q$, that is, for $\{p_k\} = \{p_k^*\}$, $\forall k = 1, 2, \dots, n$. Thus,

$$\delta TD = \left[\delta TD_q(\{p_k\})^q \right] \Big|_{\{p_k\}=\{p_k^*\}} = \left[\delta \sum_{i=1}^N |u_i|^q \right] \Big|_{\{p_k\}=\{p_k^*\}}. \tag{8}$$

In the case of a large number of sampling elements, we adopt the continuous description, i.e.,

$$\sum_{i=1}^N |u_i|^q \approx N \cdot \langle |u|^q \rangle, \text{ with } \langle |u|^q \rangle = \int_{u \in D_u} |u|^q S(u) du, \tag{9}$$

and $S(u)$ is the distribution of u -values in their domain D_u , that is,

$$\langle |u|^q \rangle = \frac{1}{L} \cdot \int_{x \in D} |V(x; \{p_k\}) - f(x)|^q dx, \tag{10}$$

since x -values are equidistributed in their domain D . Therefore,

$$\delta \sum_{i=1}^N |u_i|^q \approx \delta N \cdot \langle |u|^q \rangle + N \cdot \delta \langle |u|^q \rangle = \langle |u|^q \rangle + N \cdot \delta \langle |u|^q \rangle \approx \frac{1}{N} \sum_{i=1}^N |u_i|^q + N \cdot \delta \langle |u|^q \rangle, \tag{11}$$

where the number of the sampling elements, N , can be varied by 1, thus $\delta N = 1$. Hence,

$$\left[\delta \sum_{i=1}^N |u_i|^q \right] \Big|_{\{p_k\}=\{p_k^*\}} \approx \frac{1}{N} \left[\sum_{i=1}^N |u_i|^q \right] \Big|_{\{p_k\}=\{p_k^*\}} + N \cdot [\delta \langle |u|^q \rangle] \Big|_{\{p_k\}=\{p_k^*\}}, \tag{12}$$

where

$$\left[\sum_{i=1}^N |u_i|^q \right] \Big|_{\{p_k\}=\{p_k^*\}} = A_0(q). \tag{13}$$

Moreover, we show that the far right part of Equation (12) is zero. Indeed:

$$\delta\langle |u|^q \rangle = \frac{1}{L} \cdot \delta \int_{x \in D} |V(x; \{p_k\}) - f(x)|^q dx = \frac{1}{L} \cdot \sum_{k=1}^n A_{1,k}(q) \delta p_k, \tag{14}$$

where

$$A_{1,k}(q) = q \cdot \int_{x \in D} |V(x; \{p_k\}) - f(x)|^{q-1} \text{sign}[V(x; \{p_k\}) - f(x)] \cdot \frac{\partial V(x; \{p_k\})}{\partial p_k} dx, \tag{15}$$

and thus, we obtain:

$$[\delta\langle |u|^q \rangle]_{\{p_k\}=\{p_k^*\}} = \frac{1}{L} \cdot \sum_{k=1}^n A_{1,k}(q)_{\{p_k\}=\{p_k^*\}} \delta p_k^* = 0, \tag{16}$$

leading to the set of the following n normal equations:

$$0 = A_{1,k}(q)_{\{p_k\}=\{p_k^*\}} = q \cdot \int_{x \in D} |V(x; \{p_k^*\}) - f(x)|^{q-1} \text{sign}[V(x; \{p_k^*\}) - f(x)] \cdot \frac{\partial V(x; \{p_k\})}{\partial p_k} \Big|_{\{p_k\}=\{p_k^*\}} dx. \tag{17}$$

Finally,

$$\left[\delta \sum_{i=1}^N |u_i|^q \right]_{\{p_k\}=\{p_k^*\}} = \frac{1}{N} A_0(q), \tag{18}$$

hence:

$$\delta TD(q) = \frac{1}{N} A_0(q). \tag{19}$$

Similarly, for the continuous way of x -values, we have:

$$TD_q(\{p_k\})^q = \int_{x \in D} |u|^q dx \approx x_{res} \cdot \sum_{i=1}^N |u_i|^q, \tag{20a}$$

$$\delta TD \approx x_{res} \cdot \left[\delta \sum_{i=1}^N |u_i|^q \right]_{\{p_k\}=\{p_k^*\}} \approx \frac{1}{N} x_{res} \cdot \left[\sum_{i=1}^N |u_i|^q \right]_{\{p_k\}=\{p_k^*\}} = \frac{1}{N} A_0(q), \tag{20b}$$

$$A_0(q) = \left[\int_{x \in D} |u|^q dx \right]_{\{p_k\}=\{p_k^*\}} \approx x_{res} \cdot \left[\sum_{i=1}^N |u_i|^q \right]_{\{p_k\}=\{p_k^*\}}, \tag{20c}$$

$$\delta TD(q) = \frac{1}{N} A_0(q) \approx \frac{x_{res}}{L} A_0(q). \tag{20d}$$

The result of Equation (20d) will be used in Section 4 on the expression of the optimal errors.

3. The Uncertainty Manifold

We define the deviation of the total deviations functional from its minimum, $\Delta TD \equiv TD_q(\{p_k\})^q - TD_q(\{p_k^*\})^q > 0$, which is expressed with the quadratic form:

$$\Delta TD(\{\delta p_k\}) = \sum_{k_1, k_2=1}^n A_{2,k_1 k_2}(q) \cdot \delta p_{k_1} \delta p_{k_2}, \tag{21}$$

where we set $\delta p_k \equiv p_k - p_k^*, \forall k = 1, \dots, n$.

Given a particular value of ΔTD , each of these parameter deviations, e.g., the k -th component δp_k , has a maximum value $\delta p_{k, \max}$. This maximum value $\delta p_{k, \max}$ of each parameter deviation δp_k , depends

on the value of ΔTD . The smallest possible value of $\delta p_{k,\max}$ is deduced when ΔTD also reaches its smallest value. The smallest possible value of $\delta p_{k,\max}$ interprets the error δp_k^* of the optimal parameter values $p_k^*, \forall k = 1, 2, \dots, n$; this is achieved when the particular value ΔTD is given by the smallest possible value of a deviation from the TD 's minimum, δTD . In Section 2, we showed that δTD equals:

$$\delta TD(q) = \frac{1}{N} \cdot A_0(q). \tag{22}$$

There are cases, where the total deviations value is subject to an experimental, reading, or any other type of a non-statistical error; this is, in general, called the resolution value T_{res} . Then, the smallest possible value δTD is meaningful only when it stays above the threshold of T_{res} ; in other words, $\delta TD \geq T_{res}$ or, if $A_0/N \leq T_{res}$, then $\delta TD = T_{res}$. Hence,

$$\delta TD(q) = \begin{cases} \frac{1}{N} \cdot A_0(q), & \text{if } \frac{A_0}{N} > T_{res} , \\ T_{res} , & \text{if } \frac{A_0}{N} \leq T_{res} . \end{cases} \tag{23}$$

The quadratic form in Equation (21) is positive definite, and thus it defines an n -dimensional paraboloid (hypersurface with a local minimum) immersed into an $(n+1)$ -dimensional space. The corresponding $n + 1$ axes are given by the n parameter deviations $\{\delta p_k\}_{k=1}^n$ and the deviation ΔTD , describing thus, the $(n+1)$ -dimensional space as

$$(\delta p_1, \delta p_2, \dots, \delta p_n; \Delta TD) \in \{D_{p_1} \otimes D_{p_2} \otimes \dots \otimes D_{p_n}\} \otimes D_{\Delta TD} \subseteq \mathfrak{R}^{n+1}, \tag{24}$$

where $D_{\Delta TD} = \{\Delta TD \geq \delta TD > 0 \mid \Delta TD \in \mathfrak{R}\}$ is the domain of the deviation values, ΔTD .

Given a fixed value of ΔTD , and that can be the value of the smallest deviation, i.e., $\delta TD = \Delta TD(\{\delta p_k\})$, the set of the parameter deviations $\{\delta p_k\}_{k=1}^n$ defines a locus of an n -dimensional ellipsoid, rotated with respect to the axes $\{\delta p_k\}_{k=1}^n$. This n -dimensional ellipsoid is bounded by the $(n - 1)$ -dimensional locus of intersection between the n -dimensional paraboloid $\Delta TD = \Delta TD(\{\delta p_k\})$ and the n -dimensional hyperplane $\Delta TD = \delta TD$.

The n -dimensional ellipsoid is called uncertainty manifold, denoted by U_n , for short. This is a manifold with an edge, meaning thus, its boundary, denoted by ∂U_n . In general, the edge of an n -dimensional manifold is an $(n - 1)$ -dimensional manifold. Here, the edge ∂U_n involves the $(n - 1)$ -dimensional locus of intersection between the n -dimensional paraboloid $\Delta TD = \Delta TD(\{\delta p_k\})$ and the n -dimensional hyperplane $\Delta TD = \delta TD$. The n -dimensional cuboid, which encloses the uncertainty manifold's edge ∂U_n , is also a manifold with an edge and is denoted by Uc_n . Its edge is an $(n - 1)$ -dimensional manifold denoted by ∂Uc_n .

For example, consider the case of two-parametrical approximating functions, $V(x; p_1, p_2)$. Then, the quadratic form of Equation (21) defines the two-dimensional paraboloid $\Delta TD = \Delta TD(\delta p_1, \delta p_2)$, immersed into the three-dimensional space with Cartesian axes given by $(x \equiv \delta p_1, y \equiv \delta p_2, z \equiv \Delta TD)$. The two-dimensional ellipsoid is defined by the space bounded by the locus $\delta TD = \Delta TD(\delta p_1, \delta p_2)$, which is the intersection of the two-dimensional paraboloid $\Delta TD = \Delta TD(\delta p_1, \delta p_2)$ and the two-dimensional hyperplane $\Delta TD = \delta TD$. For visualizing this example, see Figure 1.

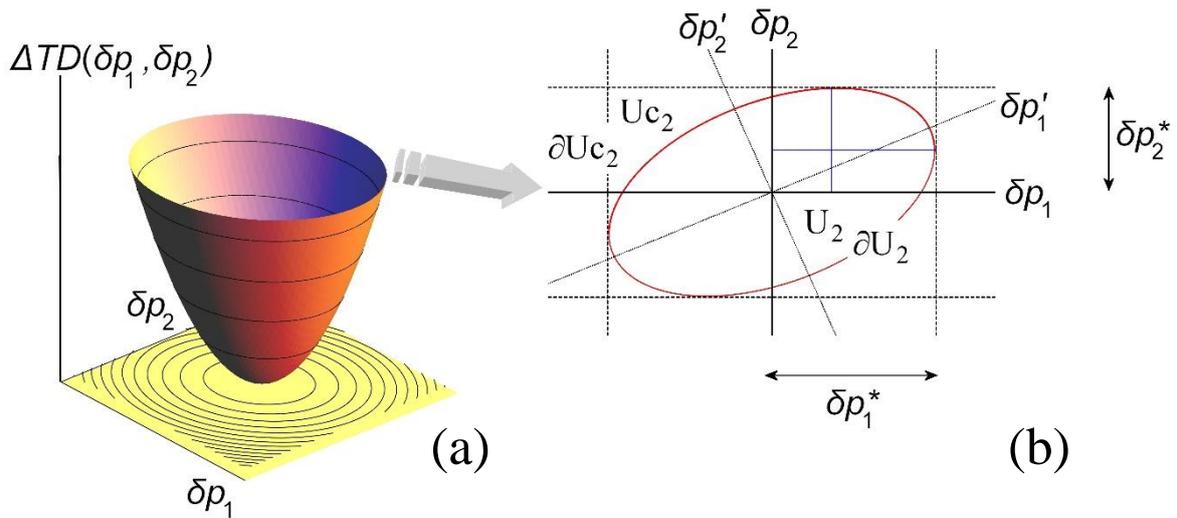


Figure 1. (a) Scheme of the n -dimensional paraboloid, $\Delta TD(\{\delta p_k\})$, which is constructed as the deviation of the total residuals or deviations functional, total deviations (TD), from its minimum value, A_0 , as shown in Equation (21); the illustrated example is for $n = 2$. (b) The intersection between the paraboloid and a constant hyper-plane $\Delta TD = \delta TD$ is a rotated n -dimensional ellipsoid, or a rotated ellipsis for the case of $n = 2$, that is the uncertainty manifold ∂U_2 (enclosing U_2). The rectangular adjusted on the extrema of ∂U_2 denotes the manifold ∂U_{c2} (enclosing U_{c2}).

Next, we will use the concept of the hyper-dimensional uncertainty manifold to derive the expressions of the errors of the optimal parameter values.

4. Derivation of the Errors of the Optimal Parameter Values

The expressions of the errors of the optimal parameter values—or simply, optimal errors—are well-known in the case of the least-square and other Euclidean based fitting methods. In [7], we have used the error expression, which is caused by the curvature, in order to have an estimate of the optimal errors (for applications, see [11–18]). Here, we will see the formal geometric derivation of the optimal error expressions.

First, we note that the edge of the uncertainty manifold, ∂U_n , has a number of n extrema, denoted by $\{C^{(k)}\}_{k=1}^n$, and they are related to the errors of the parameters optimal values, $\{\delta p_k^*\}_{k=1}^n$, as follows:

The position vector $\Delta^{\rightarrow(\mu)}$ of the corresponding point $C^{(\mu)}$, $\forall \mu = 1, 2, \dots, n$, consists of n components each, i.e., $\Delta^{\rightarrow(\mu)} = (\Delta^{(\mu)}_1, \Delta^{(\mu)}_2, \dots, \Delta^{(\mu)}_n)$. Thus,

$$\begin{aligned}
 C^{(1)} : \quad & (\Delta^{(1)}_1, \Delta^{(1)}_2, \dots, \Delta^{(1)}_n) = \Delta^{\rightarrow(1)} \\
 C^{(2)} : \quad & (\Delta^{(2)}_1, \Delta^{(2)}_2, \dots, \Delta^{(2)}_n) = \Delta^{\rightarrow(2)} \\
 & \dots \\
 C^{(n)} : \quad & (\Delta^{(n)}_1, \Delta^{(n)}_2, \dots, \Delta^{(n)}_n) = \Delta^{\rightarrow(n)}.
 \end{aligned} \tag{25}$$

These components are given by the condition:

$$\left. \left\{ \frac{\partial \Delta TD(\{\delta p_k\})}{\partial \delta p_v} \right\} \right|_{\{\delta p_k\}=\{\Delta^{(\mu)}_k\}} = \left. \left\{ \frac{\partial}{\partial \delta p_v} \sum_{k_1, k_2=1}^n A_{2, k_1 k_2}(q) \cdot \delta p_{k_1} \cdot \delta p_{k_2} \right\} \right|_{\{\delta p_k\}=\{\Delta^{(\mu)}_k\}} = 0, \tag{26}$$

$\forall v = 1, 2, \dots, n$, with $v \neq \mu$.

The above $(n - 1)$ equations, given in Equation (26), together with

$$\delta TD = \Delta TD(\{\delta p_k\} = \{\Delta^{(\mu)}_k\}) = \sum_{k_1, k_2=1}^n A_{2, k_1 k_2}(q) \cdot \Delta^{(\mu)}_{k_1} \Delta^{(\mu)}_{k_2}, \tag{27}$$

are sufficient for the calculation of the n unknown components of $\Delta^{\rightarrow(\mu)} = (\Delta^{(\mu)}_1, \Delta^{(\mu)}_2, \dots, \Delta^{(\mu)}_n)$.

The $(n - 1)$ equations, given in Equation (26), arise from the fact that each of the errors of parameters optimal values, i.e., the k -th, δp_k^* is derived from the maximum value of the corresponding component δp_k , that is, $\delta p_{k, \max}$. Hence, the existence of maximum values $\{\delta p_k^*\}_{k=1}^n$, i.e., $\forall k = 1, 2, \dots, n, \exists \delta p_k^* \in U_n: \forall \delta p_k \in U_n, \delta p_k^* \geq \delta p_k$, leads to the errors estimation, $\{\delta p_k^*\}_{k=1}^n = \{\delta p_{k, \max}\}_{k=1}^n$. These maximum values, are located on the edge of the uncertainty manifold ∂U_n , that is the hypersurface $\delta TD = \sum_{k_1, k_2=1}^n A_{2, k_1 k_2}(q) \cdot \delta p_{k_1} \delta p_{k_2}$, i.e., $\exists \delta p_k^* \in \partial U_n$.

The maximization of the μ -th parameter deviation δp_μ within uncertainty manifold is derived as follows: Since $\delta TD = \Delta TD(\{\delta p_k\})$, we can express δp_μ in terms of $\delta p_\nu, \forall \nu = 1, 2, \dots, n$ with $\nu \neq \mu$, i.e., $\delta p_\mu = \delta p_\mu(\{\delta p_k\}_{k=1, k \neq \mu}^n)$. Then, the procedure of finding the maximum value of δp_μ , that is, $\delta p_\mu^* = \delta p_\mu|_{\max}$, involves finding all the derivatives $\partial/\partial \delta p_\nu, \forall \nu = 1, 2, \dots, n$ with $\nu \neq \mu$ of $\delta p_\mu = \delta p_\mu(\{\delta p_k\}_{k=1, k \neq \mu}^n)$, or equivalently, of $\delta TD = \Delta TD(\{\delta p_k\})$ (implicit derivatives). The $(n - 1)$ equations $\partial \delta p_\mu(\{\delta p_k\}_{k=1, k \neq \mu}^n) / \partial \delta p_\nu = 0$ lead to the $(n - 1)$ relationships $\delta p_\nu = \delta p_\nu(\delta p_\mu), \forall \nu = 1, 2, \dots, n$ with $\nu \neq \mu$, which together with $\delta p_\mu = \delta p_\mu(\{\delta p_k\}_{k=1, k \neq \mu}^n)$, leads to the specific values of $\{\Delta^{(\mu)}_k = \delta p_k\}_{k=1}^n$. Yet, only the μ -th component $\Delta^{(\mu)}_\mu = \delta p_\mu$ gives the error $\delta p_\mu^* = \delta p_\mu|_{\max}$, i.e.,

$$\delta p_\mu^* = \Delta^{(\mu)}_\mu. \tag{28}$$

In Appendix A, we solve Equations (26) and (27), where, we concluded that

$$\Delta^{(m_1)}_{m_2} = \frac{\sigma^2_{m_1 m_2}}{\sqrt{\sigma^2_{m_1 m_1}}} = \sqrt{\delta TD} \cdot \frac{(A_2(q)^{-1})_{m_1 m_2}}{\sqrt{(A_2(q)^{-1})_{m_1 m_1}}}, \tag{29a}$$

or

$$\Delta^{(m_1)}_{m_2} = \begin{cases} \sqrt{\frac{1}{N} \cdot A_0(q) \cdot \frac{(A_2(q)^{-1})_{m_1 m_2}}{\sqrt{(A_2(q)^{-1})_{m_1 m_1}}}}, & \text{if } \frac{A_0}{N} > T_{res}, \\ \sqrt{T_{res} \cdot \frac{(A_2(q)^{-1})_{m_1 m_2}}{\sqrt{(A_2(q)^{-1})_{m_1 m_1}}}}, & \text{if } \frac{A_0}{N} \leq T_{res}, \end{cases} \tag{29b}$$

and thus, from Equation (28), we finally derive the errors:

$$\delta p_k^* = \Delta^{(k)}_k = \sqrt{\delta TD(q) \cdot (A_2(q)^{-1})_{kk}} = \begin{cases} \sqrt{\frac{1}{N} \cdot A_0(q) \cdot (A_2(q)^{-1})_{kk}}, & \frac{A_0}{N} > T_{res}, \\ \sqrt{T_{res} \cdot (A_2(q)^{-1})_{kk}}, & \frac{A_0}{N} \leq T_{res}, \end{cases} \tag{30a}$$

$\forall k = 1, 2, \dots, n$.

Finally, taking into account the resolution of each parameter value, $\{p_{res, k}\}_{k=1}^n$, we have:

$$\delta p_k^* = \text{Max}\left(p_{res, k}, \sqrt{\delta TD(q) \cdot (A_2(q)^{-1})_{kk}}\right), \tag{30b}$$

where

$$\sqrt{\delta TD(q) \cdot (A_2(q)^{-1})_{kk}} = \begin{cases} \sqrt{\frac{1}{N} \cdot A_0(q) \cdot (A_2(q)^{-1})_{kk}}, & \text{if } \frac{A_0}{N} > T_{res}, \\ \sqrt{T_{res} \cdot (A_2(q)^{-1})_{kk}}, & \text{if } \frac{A_0}{N} \leq T_{res}, \end{cases} \tag{30c}$$

$\forall k = 1, 2, \dots, n$.

5. Formulation of the Cases of $n = 1$ and $n = 2$ Dimensional Uncertainty Manifold

5.1. The Case of $n = 1$

Let us begin with the case of a one-dimensional paraboloid, given simply by the parabola

$$\Delta TD(\delta p) = A_2(q) \cdot \delta p^2, \tag{31}$$

corresponding to uni-parametrical approximating functions. The locus of intersection between this parabola and the line $\Delta TD = \delta TD$ (that is, the one-dimensional hyperplane) are the two points $\delta p_{\pm} = \pm \sqrt{\delta TD / A_2(q)}$. The uncertainty manifold U_1 is the one-dimensional ellipsoid, defined by the line segment $\delta p_- \leq \delta p \leq \delta p_+$, which is enclosed by the points δp_{\pm} . In this case, the edge of the uncertainty manifold ∂U_1 is restricted to the zero-dimensional space composed only by the two points δp_{\pm} . The manifolds U_1 and U_{C1} coincide (similarly with their edges, ∂U_1 and ∂U_{C1} , respectively). Hence,

$$\delta p^* = \sqrt{\frac{\delta TD(q)}{A_2(q)}} = \begin{cases} \sqrt{\frac{1}{N} \cdot A_0(q) / A_2(q)}, & \frac{A_0}{N} > T_{res}, \\ \sqrt{T_{res} / A_2(q)}, & \frac{A_0}{N} \leq T_{res}. \end{cases} \tag{32}$$

5.2. The Case of $n = 2$

The case of bi-parametrical approximating functions is characterized by the two-dimensional paraboloid,

$$\Delta TD(\delta p_1, \delta p_2) = A_{11}(q) \cdot \delta p_1^2 + 2A_{12}(q) \cdot \delta p_1 \delta p_2 + A_{22}(q) \cdot \delta p_2^2, \tag{33}$$

which is illustrated in Figure 1. The locus of intersection between this paraboloid and the plane $\Delta TD(\delta p_1, \delta p_2) = \delta TD$ is given by the rotated ellipse:

$$\delta TD = A_{11}(q) \cdot \delta p_1^2 + 2A_{12}(q) \cdot \delta p_1 \delta p_2 + A_{22}(q) \cdot \delta p_2^2, \tag{34}$$

written suitably as

$$\left(\frac{\delta p'_1}{b_1}\right)^2 + \left(\frac{\delta p'_2}{b_2}\right)^2 = 1, \tag{35}$$

after the rotation transformation

$$\begin{pmatrix} \delta p_1 \\ \delta p_2 \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \delta p'_1 \\ \delta p'_2 \end{pmatrix}, \text{ or} \tag{36}$$

$$\vec{\delta p} = R \vec{\delta p'}, \tag{37}$$

where

$$R = R(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \vec{\delta p} \equiv \begin{pmatrix} \delta p_1 \\ \delta p_2 \end{pmatrix}, \vec{\delta p'} \equiv \begin{pmatrix} \delta p'_1 \\ \delta p'_2 \end{pmatrix}. \tag{38}$$

Then,

$$\delta TD = \vec{\delta p}^t A_2 \vec{\delta p} = \vec{\delta p'}^t (R^t A_2 R) \vec{\delta p'} = \vec{\delta p'}^t B \vec{\delta p'}, \tag{39}$$

where the diagonal matrix

$$B = B(q) = \begin{pmatrix} B_{11}(q) & 0 \\ 0 & B_{22}(q) \end{pmatrix}, \tag{40}$$

has the following elements

$$B_{\pm} = \frac{1}{2} \left[A_{11} + A_{22} \pm \sqrt{(A_{11} - A_{22})^2 + 4A_{12}^2} \right], B_{11} = B_+, B_{22} = B_-, \tag{41}$$

which are the eigenvalues of the matrix $A_2(q)$. The ellipsis' major/minor axes in Equation (35) are:

$$b_1 = \sqrt{\frac{\delta TD}{B_{11}}}, b_2 = \sqrt{\frac{\delta TD}{B_{22}}}. \tag{42}$$

while the rotation angle θ in Equation (36) is given by

$$\tan 2\theta = \frac{2A_{12}}{A_{22} - A_{11}}. \tag{43}$$

The uncertainty manifold U_2 is the rotated 2-dim ellipsoid in the $(\delta p_1, \delta p_2)$ axes, defined by:

$$A_{11}(q) \cdot \delta p_1^2 + 2A_{12}(q) \cdot \delta p_1 \delta p_2 + A_{22}(q) \cdot \delta p_2^2 \leq \delta TD, \tag{44}$$

or, in the rotated axes $(\delta p'_1, \delta p'_2)$, is simply given by:

$$\left(\frac{\delta p'_1}{b_1} \right)^2 + \left(\frac{\delta p'_2}{b_2} \right)^2 \leq 1, \tag{45}$$

which is enclosed by the ellipse corresponding to the equal sign of Equation (45), that is the edge of the uncertainty manifold, ∂U_2 .

Finally, the errors are

$$\begin{aligned} \delta p_1^* &= \sqrt{\delta TD(q) \cdot (A_2(q)^{-1})_{11}} = \sqrt{\delta TD(q) \cdot \frac{(A_2)_{22}(q)}{D_{A_2}(q)}} \\ &= \begin{cases} \sqrt{\frac{1}{N} \cdot A_0(q) \cdot \frac{(A_2)_{22}(q)}{D_{A_2}(q)}}, & \frac{A_0}{N} > T_{res}, \\ \sqrt{T_{res} \cdot \frac{(A_2)_{22}(q)}{D_{A_2}(q)}}, & \frac{A_0}{N} \leq T_{res}, \end{cases} \end{aligned} \tag{46}$$

and

$$\begin{aligned} \delta p_2^* &= \sqrt{\delta TD(q) \cdot (A_2(q)^{-1})_{22}} = \sqrt{\delta TD(q) \cdot \frac{(A_2)_{11}(q)}{D_{A_2}(q)}} \\ &= \begin{cases} \sqrt{\frac{1}{N} \cdot A_0(q) \cdot \frac{(A_2)_{11}(q)}{D_{A_2}(q)}}, & \frac{A_0}{N} > T_{res}, \\ \sqrt{T_{res} \cdot \frac{(A_2)_{11}(q)}{D_{A_2}(q)}}, & \frac{A_0}{N} \leq T_{res}. \end{cases} \end{aligned} \tag{47}$$

6. Conclusions

The paper presented the geometric interpretation of the errors of the optimal parameter values, derived from a multi-parametrical fitting, based on a metric induced by the non-Euclidean L_q -norm. Typical fitting methods are mostly developed based on Euclidean norms, leading to the traditional least-square method. On the other hand, the theory of general fitting methods based on non-Euclidean norms, is still under development; the normal equations can provide the optimal values of the fitting parameters, while this paper completed the puzzle by improving understanding the derivations and geometric meaning of the errors.

In particular, we showed that the statistical errors of the optimal parameter values are given by the axes of the ellipsoid called uncertainty manifold, that is, the intersection of the paraboloid

of the residuals' expansion $\Delta TD(\{\delta p_k\}) \equiv TD_q(\{p_k = p_k^* + \delta p_k\})^q - TD_q(\{p_k^*\})^q$ along the deviations $\{\delta p_k\}_{k=1}^n$, with the hyperplane $\Delta TD(\{\delta p_k\}) = \delta TD = const$. The constant δTD represents the smallest possible value of a deviation from the TD's minimum, also mentioned as an error of the value of the total deviations.

In summary, the L_q -normed fitting involves minimizing:

$$TD_q(\{p_k\})^q = A_0(q) + \sum_{k_1, k_2=1}^n A_{2, k_1 k_2}(q) \cdot (p_{k_1} - p_{k_1}^*) \cdot (p_{k_2} - p_{k_2}^*) + O(|p_k - p_k^*|^3), \tag{48a}$$

where

$$A_0(q) = \int_{x \in D} |u|^q dx, \tag{48b}$$

$$A_{2, k_1 k_2}(q) = \delta_{1q} \cdot \gamma_{k_1 k_2} + \frac{q}{2} \cdot \int_{x \in D} \left[|u|^{q-1} \cdot \text{sgn}(u) \cdot \frac{\partial^2 V(x; \{p_k^*\})}{\partial p_{k_1} \partial p_{k_2}} + (q-1) \cdot |u|^{q-2} \cdot \frac{\partial V(x; \{p_k^*\})}{\partial p_{k_1}} \cdot \frac{\partial V(x; \{p_k^*\})}{\partial p_{k_2}} \right] dx, \tag{48c}$$

$$\gamma_{k_1 k_2} \equiv \sum_{\forall i : u(x_i) = 0} \frac{1}{|u'(x_i)|} \cdot \frac{\partial V(x_i; \{p_k^*\})}{\partial p_{k_1}} \cdot \frac{\partial V(x_i; \{p_k^*\})}{\partial p_{k_2}}. \tag{48d}$$

The normal equations are given by:

$$\int_{x \in D} |u|^{q-1} \cdot \text{sgn}(u) \cdot \frac{\partial V(x; \{p_k^*\})}{\partial p_l} dx = 0, \forall l = 1, 2, \dots, n, \tag{48e}$$

where we set $u = u(x) \equiv V(x; \{p_k^*\}) - f(x)$.

Finally, we summarize the concluding relationships of the paper:

$$\Delta TD(\{\delta p_k\}) \equiv TD_q(\{p_k = p_k^* + \delta p_k\})^q - TD_q(\{p_k^*\})^q = \sum_{k_1, k_2=1}^n A_{2, k_1 k_2}(q) \cdot \delta p_{k_1} \delta p_{k_2}, \tag{49a}$$

$$\Delta TD_{\min} \equiv \text{Max}(\delta TD(q), T_{res}), \delta TD(q) = \frac{1}{N} \cdot A_0(q), A_0(q) \equiv TD_q(\{p_k^*\})^q, \tag{49b}$$

$$\delta p_k^* = \text{Max}\left(p_{res, k}, \sqrt{\delta TD(q) \cdot (A_2(q)^{-1})_{kk}}\right), \tag{49c}$$

with special cases:

- For $n = 1$:

$$\delta p^* = \sqrt{\frac{\delta TD(q)}{A_2(q)}} = \begin{cases} \sqrt{\frac{1}{N} \cdot A_0(q) / A_2(q)}, & \frac{A_0}{N} > T_{res}, \\ \sqrt{T_{res} / A_2(q)}, & \frac{A_0}{N} \leq T_{res}. \end{cases} \tag{50a}$$

- For $n = 2$:

$$\begin{aligned} \delta p_1^* &= \sqrt{\delta TD(q) \cdot (A_2(q)^{-1})_{11}} = \sqrt{\delta TD(q) \cdot \frac{(A_2)_{22}(q)}{D_{A_2}(q)}} \\ &= \begin{cases} \sqrt{\frac{1}{N} \cdot A_0(q) \cdot \frac{(A_2)_{22}(q)}{D_{A_2}(q)}}, & \frac{A_0}{N} > T_{res}, \\ \sqrt{T_{res} \cdot \frac{(A_2)_{22}(q)}{D_{A_2}(q)}}, & \frac{A_0}{N} \leq T_{res}, \end{cases} \end{aligned} \tag{50b}$$

and

$$\begin{aligned} \delta p_2^* &= \sqrt{\delta TD(q) \cdot (A_2(q)^{-1})_{22}} = \sqrt{\delta TD(q) \cdot \frac{(A_2)_{11}(q)}{D_{A_2}(q)}} \\ &= \begin{cases} \sqrt{\frac{1}{N} \cdot A_0(q) \cdot \frac{(A_2)_{11}(q)}{D_{A_2}(q)}}, & \frac{A_0}{N} > T_{res}, \\ \sqrt{T_{res} \cdot \frac{(A_2)_{11}(q)}{D_{A_2}(q)}}, & \frac{A_0}{N} \leq T_{res}. \end{cases} \end{aligned} \tag{50c}$$

Funding: This research was funded by NASA’s HGI Program, grant number NNX17AB74G.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. Extrema of the Uncertainty Manifold

Here we calculate the position vectors $\left\{ \overset{\rightarrow}{\Delta}^{(k)} \right\}_{k=1}^n$ and the maximum points $\left\{ C^{(k)} \right\}_{k=1}^n$, used for the derivations of errors of the optimal parameter values in Section 4.

The following $(n - 1)$ equations:

$$\left[\frac{\partial}{\partial \delta p_\nu} \Delta TD(\{\delta p_k\}) \right] \Big|_{\{\delta p_k\} = \{\Delta^{(\mu)}_k\}} = 0, \tag{A1}$$

$\forall v = 1, 2, \dots, n$, with $v \neq \mu$, together with the one of

$$\delta TD = \Delta TD(\{\delta p_k\} = \{\Delta^{(\mu)}_k\}) = \sum_{k_1, k_2=1}^n A_{2, k_1 k_2}(q) \cdot \Delta^{(\mu)}_{k_1} \Delta^{(\mu)}_{k_2}, \tag{A2}$$

are sufficient for the calculation of the n unknown components of $\overset{\rightarrow}{\Delta}^{(\mu)} = (\Delta^{(\mu)}_1, \Delta^{(\mu)}_2, \dots, \Delta^{(\mu)}_n)$. Then,

$$\begin{aligned} &\left\{ \frac{\partial}{\partial \delta p_\nu} [\Delta TD(\{\delta p_k\})] \right\} \Big|_{\{\delta p_k\} = \{\Delta^{(\mu)}_k\}} = \left\{ \frac{\partial}{\partial \delta p_\nu} \left[\sum_{k_1, k_2=1}^n A_{2, k_1 k_2}(q) \cdot \delta p_{k_1} \cdot \delta p_{k_2} \right] \right\} \Big|_{\{\delta p_k\} = \{\Delta^{(\mu)}_k\}} \\ &= \left\{ \sum_{k_1, k_2=1}^n A_{2, k_1 k_2}(q) \cdot \delta p_{k_1} \cdot \delta_{k_2 \nu} + \sum_{k_1, k_2=1}^n A_{2, k_1 k_2}(q) \cdot \delta p_{k_2} \cdot \delta_{k_1 \nu} \right\} \Big|_{\{\delta p_k\} = \{\Delta^{(\mu)}_k\}} \\ &= 2 \left\{ \sum_{k=1}^n A_{2, \nu k}(q) \cdot \delta p_k \right\} \Big|_{\{\delta p_k\} = \{\Delta^{(\mu)}_k\}} = 2 \cdot \sum_{k=1}^n A_{2, \nu k}(q) \cdot \Delta^{(\mu)}_k, \end{aligned} \tag{A3}$$

(where we used the Kronecker’s delta, $\delta_{mn} = 1$ for $m = n$ and 0 for $m \neq n$). Hence,

$$\sum_{k=1}^n A_{2, \nu k}(q) \cdot \Delta^{(\mu)}_k = 0, \tag{A4}$$

$\forall v = 1, 2, \dots, n$, with $v \neq \mu$. Setting:

$$\sum_{k=1}^n A_{2, \mu k}(q) \cdot \Delta^{(\mu)}_k \equiv \zeta_\mu, \tag{A5}$$

we have

$$\zeta_\mu \cdot \hat{e}_\mu = (A_2) \overset{\rightarrow}{\Delta}^{(\mu)}, \tag{A6}$$

where

$$\hat{e}_\mu \equiv \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \left(\{\hat{e}_{\mu,k} = 0\}_{k=1, k \neq \mu}^n; \hat{e}_{\mu,\mu} = 1 \right). \quad (\text{A7})$$

Inversing Equation (A6), we obtain

$$\overset{\rightarrow}{\Delta}^{(\mu)} = \zeta_\mu \cdot (A_2^{-1}) \hat{e}_\mu. \quad (\text{A8})$$

or

$$\Delta^{(\mu)}_m = (A_2^{-1})_{m\mu} \cdot \zeta_\mu, \quad (\text{A9})$$

$\forall m = 1, 2, \dots, n$. Then,

$$\begin{aligned} \delta TD &= \sum_{k_1, k_2=1}^n A_{2, k_1 k_2} \cdot \Delta^{(\mu)}_{k_1} \Delta^{(\mu)}_{k_2} = \sum_{k_1, k_2=1}^n A_{2, k_1 k_2} \cdot (A_2^{-1})_{k_1 \mu} \cdot \zeta_\mu \cdot (A_2^{-1})_{k_2 \mu} \cdot \zeta_\mu \\ &= \zeta_\mu^2 \sum_{k_1, k_2=1}^n A_{2, k_1 k_2} \cdot (A_2^{-1})_{k_1 \mu} (A_2^{-1})_{k_2 \mu} = \zeta_\mu^2 \sum_{k_2=1}^n \delta_{k_2 \mu} \cdot (A_2^{-1})_{k_2 \mu} = \zeta_\mu^2 \cdot (A_2^{-1})_{\mu \mu}, \end{aligned} \quad (\text{A10})$$

thus,

$$\zeta_\mu = \sqrt{\frac{\delta TD}{(A_2^{-1})_{\mu \mu}}}, \quad (\text{A11})$$

or

$$\Delta^{(\mu)}_m = (A_2^{-1})_{m\mu} \cdot \sqrt{\frac{\delta TD}{(A_2^{-1})_{\mu \mu}}}, \quad (\text{A12})$$

$\forall m = 1, 2, \dots, n$, and by refreshing the indices, we end up with

$$\Delta^{(m_1)}_{m_2} = \sqrt{\delta TD} \cdot \frac{(A_2(q)^{-1})_{m_1 m_2}}{\sqrt{(A_2(q)^{-1})_{m_1 m_1}}}. \quad (\text{A13})$$

References

1. Kenney, J.F.; Keeping, E.S. Linear Regression and Correlation. In *Mathematics of Statistics*, 3rd ed.; Part 1; Van Nostrand: Princeton, NJ, USA, 1962; pp. 252–285.
2. McCullagh, P. What is statistical model? *Ann. Stat.* **2002**, *30*, 1225–1310. [[CrossRef](#)]
3. Adèr, H.J. Modelling. In *Advising on Research Methods: A Consultant's Companion*; Adèr, H.J., Mellenbergh, G.J., Eds.; Johannes van Kessel Publishing: Huizen, The Netherlands, 2008; pp. 271–304.
4. Melissinos, A.C. *Experiments in Modern Physics*; Academic Press Inc.: London, UK, 1966; pp. 438–464.
5. Burden, R.L.; Faires, J.D. *Numerical Analysis*; PWS Publishing Company: New Orleans, LA, USA, 1993; pp. 437–438.
6. Livadiotis, G.; McComas, D.J. Fitting method based on correlation maximization: Applications in Astrophysics. *J. Geophys. Res.* **2013**, *118*, 2863–2875. [[CrossRef](#)]
7. Livadiotis, G. Approach to general methods for fitting and their sensitivity. *Physica A* **2007**, *375*, 518–536. [[CrossRef](#)]
8. Livadiotis, G. Expectation values and variance based on Lp norms. *Entropy* **2012**, *14*, 2375–2396. [[CrossRef](#)]

9. Livadiotis, G. Chi-p distribution: Characterization of the goodness of the fitting using L_p norms. *J. Stat. Distr. Appl.* **2014**, *1*, 4. [[CrossRef](#)]
10. Livadiotis, G. Non-Euclidean-normed Statistical Mechanics. *Physica A* **2016**, *445*, 240–255. [[CrossRef](#)]
11. Livadiotis, G. On the convergence and law of large numbers for the non-Euclidean L_p -means. *Entropy* **2017**, *19*, 217. [[CrossRef](#)]
12. Livadiotis, G.; Moussas, X. The sunspot as an autonomous dynamical system: A model for the growth and decay phases of sunspots. *Physica A* **2007**, *379*, 436–458. [[CrossRef](#)]
13. Frisch, P.C.; Bzowski, M.; Livadiotis, G.; McComas, D.J.; Möbius, E.; Mueller, H.-R.; Pryor, W.R.; Schwadron, N.A.; Sokól, J.M.; Vallergera, J.V.; et al. Decades-long changes of the interstellar wind through our solar system. *Science* **2013**, *341*, 1080. [[CrossRef](#)] [[PubMed](#)]
14. Schwadron, N.A.; Mobius, E.; Kucharek, H.; Lee, M.A.; French, J.; Saul, L.; Wurz, P.; Bzowski, M.; Fuselier, S.; Livadiotis, G.; et al. Solar radiation pressure and local interstellar medium flow parameters from IBEX low energy hydrogen measurements. *Astrophys. J.* **2013**, *775*, 86. [[CrossRef](#)]
15. Livadiotis, G.; McComas, D.J. Evidence of large scale phase space quantization in plasmas. *Entropy* **2013**, *15*, 1118–1132. [[CrossRef](#)]
16. Fuselier, S.A.; Allegrini, F.; Bzowski, M.; Dayeh, M.A.; Desai, M.; Funsten, H.O.; Galli, A.; Heirtzler, D.; Janzen, P.; Kubiak, M.A.; et al. Low energy neutral atoms from the heliosheath. *Astrophys. J.* **2014**, *784*, 89. [[CrossRef](#)]
17. Zirnstein, E.J.; McComas, D.J. Using kappa functions to characterize outer heliosphere proton distributions in the presence of charge-exchange. *Astrophys. J.* **2015**, *815*, 31. [[CrossRef](#)]
18. Fuselier, S.A.; Dayeh, M.A.; Livadiotis, G.; McComas, D.J.; Ogasawara, K.; Valek, P.; Funsten, H.O. Imaging the development of the cold dense plasma sheet. *Geophys. Res. Lett.* **2015**, *42*, 7867–7873. [[CrossRef](#)]



© 2019 by the author. 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/>).