



Taking Better Advantage of Fold Axis Data to Characterize Anisotropy of Complex Folded Structures in the Implicit Modeling Framework

Laure Pizzella, Robin Alais, Simon Lopez, Xavier Freulon, Jacques Rivoirard

► To cite this version:

Laure Pizzella, Robin Alais, Simon Lopez, Xavier Freulon, Jacques Rivoirard. Taking Better Advantage of Fold Axis Data to Characterize Anisotropy of Complex Folded Structures in the Implicit Modeling Framework. Mathematical Geosciences, 2021, 10.1007/s11004-021-09950-0 . hal-03278336

HAL Id: hal-03278336

<https://hal.science/hal-03278336>

Submitted on 22 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Taking better advantage of fold axis data to characterize anisotropy of complex folded structures in the implicit modeling framework

Laure Pizzella (✉ laure.pizzella@gmail.com)

MINES ParisTech <https://orcid.org/0000-0001-6295-8431>

Robin Alais

Mines ParisTech

Simon Lopez

Bureau de Recherches Geologiques et Minieres

Xavier Freulon

Mines ParisTech

Jacques Rivoirard

Mines ParisTech

Research Article

Keywords: 3-D structural geological modeling, Implicit modeling, Potential-field method, Non-Stationarity, Cokriging

Posted Date: May 7th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-485111/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Mathematical Geosciences on July 5th, 2021. See the published version at <https://doi.org/10.1007/s11004-021-09950-0>.

Taking better advantage of fold axis data to characterize anisotropy of complex folded structures in the implicit modeling framework

Laure Pizzella · Robin Alais · Simon Lopez ·
Xavier Freulon · Jacques Rivoirard

Received: date / Accepted: date

Abstract When too few field measurements are available for the geological modeling of complex folded structures, results of implicit methods typically exhibit an unsatisfactory bubbly aspect. However, in such cases, anisotropy data are often readily available but not fully exploited. Among them, fold axis data are a straightforward indicator of this local anisotropy direction. Focusing on the so-called potential field method, this work aims to evaluate the effect of the incorporation of such data into the modeling process. Given locally sampled fold axis data, this paper proposes to use the second-order derivatives of the scalar field in addition to the existing first-order ones. The mathematical foundation of the approach is developed and the respective efficiencies of both kinds of constraints are tested. Their integration and impact are discussed based on a synthetic case study, thereby providing practical guidelines to geomodeling tool users on the parsimonious use of data for the geological modeling of complex folded structures.

Laure Pizzella (✉)

Geostatistics Research Group, Center of Geosciences, MINES ParisTech , PSL Research University, Fontainebleau, France
E-mail: laure.pizzella@mines-paristech.fr

Robin Alais

Center for Mathematical Morphology, MINES ParisTech, PSL Research University, Fontainebleau, France
E-mail: robin.alais@mines-paristech.fr

Simon Lopez

Bureau de Recherches Géologiques et Minières, Georesources Division, Orléans, France
E-mail: slopez@brgm.fr

Xavier Freulon

Geostatistics Research Group, Center of Geosciences, MINES ParisTech , PSL Research University, Fontainebleau, France
E-mail: xavier.freulon@mines-paristech.fr

Jacques Rivoirard

Geostatistics Research Group, Center of Geosciences, MINES ParisTech , PSL Research University, Fontainebleau, France
E-mail: jacques.rivoirard@mines-paristech.fr

Keywords 3-D structural geological modeling · Implicit modeling · Potential-field method · Non-Stationarity · Cokriging

1 Introduction

Structural geological modeling aims to achieve three-dimensional models of geological objects and of their relationships. It is a mandatory and determining step for most geological studies in various fields of application: resource estimation, reservoir study, risk assessment, environmental remediation. It relies on a conceptual model proposed by geologists based on field observations. Given these preliminary data, the main goal is to choose a suitable mathematical model with parameters that give an optimal representation of the reality [Wellmann and Caumon, 2018].

For this purpose, explicit and implicit modeling are the two leading approaches. The former directly builds geological interfaces by solving a least-square smooth approximation problem on a discrete mesh [Mallet, 1992]. In contrast, the latter relies on a preliminary interpolation of a three-dimensional scalar field from geological data. Only then are the surfaces of interest extracted: they are those holding the same scalar value, referred to as their potential.

Over the last decade, implicit modeling algorithms have gained popularity in geomodelling software. Two families of methods can be distinguished, depending on whether the interpolation is performed on a meshed grid [Frank et al., 2007, Caumon et al., 2012] or not [Houlding, 1994]. Focus here is on the meshless approach. In this context, different basis functions have been deployed in order to interpolate the scalar field: the Radial-Basis functions [Cowan et al., 2003, Turk and O’Brien, 2005, Hillier et al., 2014], the Moving Least Squares functions [Manchuk and Deutsch, 2019, Renaudeau et al., 2019] or the dual form of kriging with the use of covariance functions [Lajaunie et al., 1997, Calcagno et al., 2008]. Following in the footsteps of Lajaunie et al. [1997], the approach developed in this paper is of geostatistical nature. It exploits and investigates possible improvements of the so-called potential field method.

One drawback of this seminal and efficient method is that it is not suited to the representation of complex structures such as poly-deformed areas, where the available data can be too sparse to correctly describe geometries at the chosen complexity level [Hillier et al., 2014]. To deal with this issue, other methods typically introduce inferred or interpreted locally varying anisotropies [Boisvert et al., 2009]. Alternatively, when facing important uncertainties, experienced users can introduce a priori geological knowledge into the model in the form of additional control points. This workaround is usually time consuming, error prone and user dependent. By its deterministic nature, it also prohibits any uncertainty analysis or quantification.

In the specific case of poly-deformed areas, which are at the core of the present study, the most challenging aspect is to understand, represent and model folded structures [Ramsay and Huber, 1987]. They are typically complex when formed by a succession of folding events or impacted by salt intrusions. In fact, taking into account these consecutive events can help enhance the implicit modeling framework. For instance, Maxelon et al. [2009] interpret variations in the orientation of the dominant planar fabrics. Laurent et al. [2016] suggest to integrate additional field data describ-

ing folds, lineations and vergence. This approach and the inference of its parameters are discussed in Grose et al. [2017]. In another vein, de la Varga and Wellmann [2016] adopt a Bayesian point of view: field data and interpretations are taken as prior information in the model.

While all these geological measurements bring relevant information about fold geometry, one straightforward pointer of the preferential direction of continuity of the fold structure (later referred to as the anisotropy) is the hinge line (Fig. 1). In the implicit approach, this line is directly linked to the underlying scalar field, which stays constant along the hinge line [Hillier et al., 2014]. It turns out that working with the direction of this line has a practical asset: it allows to use isotropic covariance functions in the interpolation model, while taking into account the anisotropy of the fold structure. The hinge line is deduced in practice from geological observations and more particularly from fold axes data, which are typically readily available in most geological studies. However, they are not always fully exploited.

The main objective of this paper is to explore three possible ways of taking better advantage of these data. First, by adding a second derivative constraint along the hinge line. Under some mild assumptions on the fold structure, it is proven that this is not only an appropriate mathematical translation of the constant property of the scalar field along this line, but also an additional geometrical feature that can help to better define the model. Second, by optimizing the location and number of first order derivative data. Last, by further constraining the model using both first and second derivative data in two orthogonal directions.

We start off in Sec. 2 by summarizing the theoretical background of the potential-field method. A new implementation of the second-order derivative values of the potential is introduced in this framework. The mathematical advantages of using first and second derivatives are given, and proofs of their ability to improve fold geometry modeling are also provided when needed. Then, a challenging synthetic case study with sparse data is presented in Sec. 3. It is used to illustrate the assets and limitations of using these mathematical objects in Sec. 4, thereby leading to the consideration of the maximum curvature criterion. These results, finally discussed in Sec. 5, should allow geomodeling software developers or users to get more insight on fold axis incorporation for the anisotropy modeling process.

2 Mathematical Framework

2.1 Three-dimensional implicit function construction

The key challenge of three-dimensional structural modeling in geology boils down to a reconstruction problem. It consists in approximating a three-dimensional model of the reality from sparse data sampled on the geological field under study. In the implicit modeling approach, focus is on the scalar field function Z , called potential, that maps any point $\mathbf{p} \in \mathbb{R}^3$ to a real value

$$\mathbf{p} \mapsto Z(\mathbf{p}) \quad \mathbb{R}^3 \rightarrow \mathbb{R}.$$

Punctual, localized and sparse measurements of this function are used to estimate it on the entire domain of interest. This is achieved by means of an interpolation technique (co-kriging, radial basis functions). The primary purpose of this scalar field is the extraction of iso-surfaces S_α

$$S_\alpha = \{\mathbf{p} \in \mathbb{R}^3, Z(\mathbf{p}) = \alpha\}.$$

They are the surfaces of interest, corresponding to the geological interfaces of the three-dimensional structural model. The so-called potential field method [Lajaunie et al., 1997] directly considers increments of the scalar field sampled on iso-surfaces. The implicit function is then defined up to an additive constant. It can be set by choosing a reference iso-surface, the potential value of which is arbitrarily set to zero [Calcagno et al., 2008].

2.2 Typical input data characterization

In order to sample the scalar field, geologists have access to different categories of data. In this work, we focus on field measurement data (or surface data) but the following remarks can be easily generalized to borehole data, or manual or automatic surface interpretation from geophysical surveys.

Field measurements may be summarized into two main categories: position point information ($\mathbf{p} \in \mathbb{R}^3$) and orientation vector data ($\mathbf{d} \in \mathbb{R}^3$). With regards to the implicit function Z to build, these field data mathematically correspond to the local evaluation of the function $Z(\mathbf{p})$ or its directional derivative $D_{\mathbf{d}} Z(\mathbf{p})$, respectively:

- i. The assertion that two points, $\mathbf{p}_i$ and $\mathbf{p}_j$, belong to the same interface S_α . They consequently present same potential value, therefore the difference of their potential value (called increment) is null

$$Z(\mathbf{p}_i) - Z(\mathbf{p}_j) = 0 \quad \forall \mathbf{p}_i, \mathbf{p}_j \in S_\alpha.$$

These points are usually called contact points. Among implicit approaches, increment data constitute the peculiarity of the potential field method; other implicit methods directly integrate contact point information by itself.

Since the potential value of $Z(\mathbf{p})$ at a particular point $\mathbf{p}$ is not geological data, increment data are geologically the most convenient to sample, but direct values of potential z_i may be introduced alternatively

$$Z(\mathbf{p}_i) = z_i \quad \forall \mathbf{p}_i \in \mathbb{R}^3.$$

- ii. By definition, the first order derivative of Z at point $\mathbf{p}$ in a direction $\mathbf{d}$ is

$$D_{\mathbf{d}} Z(\mathbf{p}) = \lim_{h \rightarrow 0} \frac{Z(\mathbf{p} + h \mathbf{d}) - Z(\mathbf{p})}{h \|\mathbf{d}\|}.$$

As the potential values have no physical meaning, there is no way to sample directional derivatives other than identifying direction lines that lie in the plane tangent to geological interfaces. Indeed, as geological interfaces are supposed to

be iso-potential surfaces, all derivatives of the potential along such direction lines will vanish. Such directions will usually be sampled as unit vectors τ such that $D_\tau Z(\mathbf{p}) = 0$, which can also be written

$$D_\tau Z(\mathbf{p}) = \tau \cdot \nabla Z(\mathbf{p}) = \langle \nabla Z(\mathbf{p}), \tau \rangle = 0,$$

where $\nabla Z(\mathbf{p})$ denotes the gradient of the scalar field, that writes

$$\nabla Z(\mathbf{p}) = \begin{pmatrix} D_{\mathbf{e}_x} Z(\mathbf{p}) \\ D_{\mathbf{e}_y} Z(\mathbf{p}) \\ D_{\mathbf{e}_z} Z(\mathbf{p}) \end{pmatrix},$$

with $(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$ being the standard basis of $\mathbb{R}^3$.

Geologists usually sample the tangent plane of geological interfaces using two such tangent directions, respectively called strike and dip directions. Strike direction, given by the unit vector τ_s , is obtained as the intersection of the tangent plane with a horizontal plane that holds the observation point. Dip direction, given by the unit vector τ_d , is measured replacing the horizontal plane by a vertical plane. Last, but not least, geological interfaces can be attributed a polarity, as formations can usually be chronologically ordered. Assuming that the potential Z is a decreasing function of the geological age of the formations, and that $\varepsilon \in \{-1, 1\}$ is such that $\varepsilon \tau_s \times \tau_d$ points toward younger formations we then know that

$$(\varepsilon \tau_s \times \tau_d) \cdot \nabla Z > 0.$$

In practical applications and software packages, this inequality constraint is usually replaced by the (biased) assumption that the the gradient of the scalar field has unit norm with the consequence that one can write

$$\nabla Z(\mathbf{p}) = \varepsilon \tau_s \times \tau_d,$$

This trick will be denoted as the gradient norm bias in the following.

2.3 A new input data: the second-order derivative data

Just as the first-order derivative, the second-order derivative of the potential computed twice in direction $\mathbf{d}$ can be taken into account

$$D_{\mathbf{d}}^2 Z(\mathbf{p}) = \lim_{h \rightarrow 0} \frac{Z(\mathbf{p} + h \mathbf{d}) + Z(\mathbf{p} - h \mathbf{d}) - 2 Z(\mathbf{p})}{h^2 \|\mathbf{d}\|^2}. \quad (1)$$

Motivations for the addition of such data and its impact on the method are discussed in Sec. 2.5. Let us notice that the directional second-derivative of Z in a direction τ can be expressed using the Hessian matrix of Z , denoted by $D^2 Z(\mathbf{p})$ in the following

$$D_\tau^2 Z(\mathbf{p}) = {}^t \tau \cdot D^2 Z(\mathbf{p}) \cdot \tau,$$

where the left-hand side superscript t indicates the transpose of a vector or matrix.

2.4 The interpolation method: the potential field method

Using a co-kriging estimator [Wackernagel, 2013], the interpolated value of Z at point $\mathbf{p}$, $Z^*(\mathbf{p})$, is computed by the weighted sum of different input data

$$\begin{aligned} [Z(\mathbf{p}) - Z(\mathbf{p}_0)]^* = & \sum_{ii'} \lambda_{ii'} [Z(\mathbf{p}_i) - Z(\mathbf{p}_{i'})] + \sum_j \lambda_j \langle \nabla Z(\mathbf{p}_j), \tau_j \rangle \\ & + \sum_k \sum_d^{dim} \lambda_{k,d} D_{\mathbf{e}_d} Z(\mathbf{p}_k) + \sum_s \lambda_s ({}^t \tau_s \cdot D^2 Z(\mathbf{p}_s) \cdot \tau_s), \end{aligned} \quad (2)$$

where the $\lambda_{ii'}$, λ_j , λ_k and λ_s are coefficients to be determined and $Z(\mathbf{p}_0)$ is an arbitrary value of Z assigned at the arbitrary point $\mathbf{p}_0$ (most often chosen to be null).

In the universal co-kriging context [Chiles and Delfiner, 2009], the random function, here the potential function Z , is described as a combination of a residual random function with zero mean, Y , and a deterministic drift function, m

$$Z(\mathbf{p}) - Z(\mathbf{p}_0) = Y(\mathbf{p}) + m(\mathbf{p}). \quad (3)$$

The drift function m is usually expressed as a linear combination of n basis functions

$$m(\mathbf{p}) = \sum_{\ell=1}^n v_\ell f^\ell(\mathbf{p}), \quad (4)$$

where the f^ℓ are basis functions and v_ℓ unknown associated coefficients.

Kriging equations in this case, detailed in Appendix A, lead to a final co-kriging system

$$\begin{pmatrix} K & F \\ {}^t F & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} K_p \\ F_p \end{pmatrix}, \quad (5)$$

where K and K_p are the covariance matrices, respectively between data among themselves, and data and interpolated points; F and F_p are the universality conditions on data and interpolated points and λ and μ , the unknown variables to solve.

2.5 Motivations for second-order derivative data implementation

The main difference here, compared to the initial formulation of the method from Lajaunie et al. [1997], is the addition of second-order derivative values of potential in an established direction into the co-kriging equations. Original observations, which lead to the motivations of this development, are detailed below, as well as the mathematical proof of its merits.

2.5.1 Empirical observations

This new development has been motivated by the observation that, in case of folds with planar hinge line, in addition to the first-order derivative values of the potential (tangent), the second-order derivative values of the potential for a point on the hinge line are (Fig. 2):

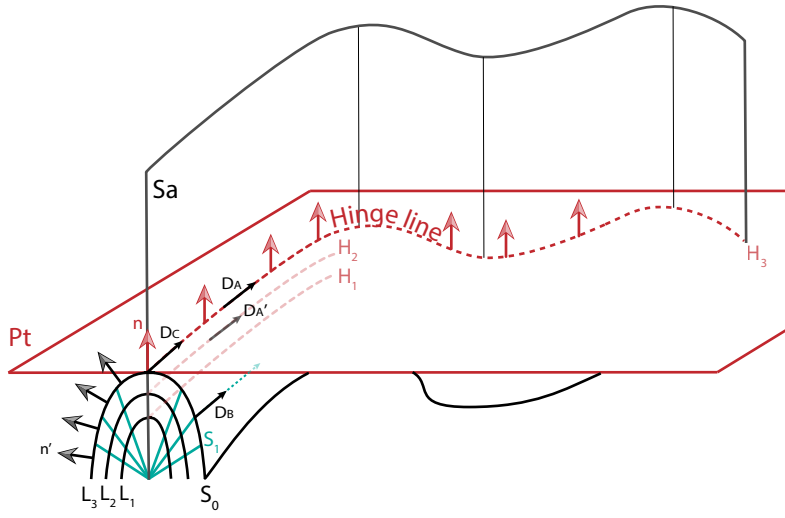


Fig. 1: Fold geometry and planes of interest. Sa: Axial surface. n : Pole vectors belonging to the axial plane; n' : Other pole vectors of the same profile. Pt : Tangent plane. S_0 : Bedding. S_1 : Cleavage. H_1 , H_2 , H_3 : Hinge lines of beddings L_1 , L_2 , L_3 . The hinge line H_3 and the crest line coincide in this case. D_A and $D_{A'}$: Fold axis measurements on hinge lines H_2 and H_3 . D_B : Intersection lineation between S_0 and S_1 measurement. D_C : Fold axis inference from stereographic projections using n' pole measurements.

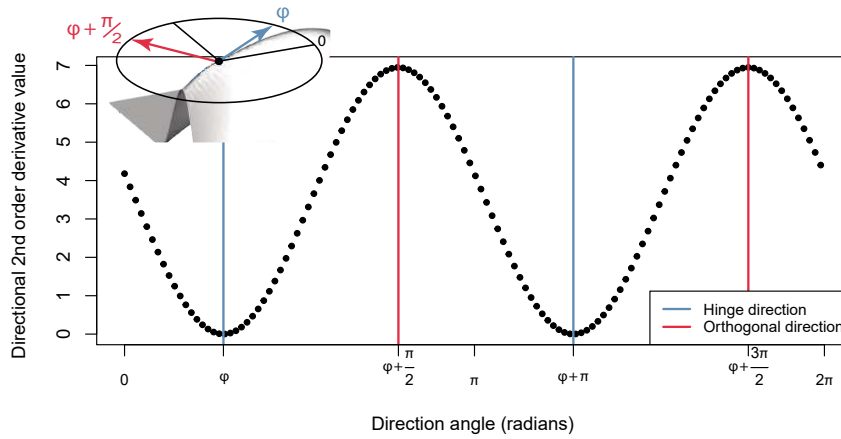


Fig. 2: Value of the second-order derivative of a point (Eq. 1) on the crest (hinge line) of the synthetic model (Sec. 3) in all directions on the tangent plane. ϕ and $\phi + \pi$ are the directions along the hinge line for this given point. $\phi + \frac{\pi}{2}$ and $\phi + \frac{3\pi}{2}$ are the orthogonal directions of the hinge line on the tangent plane.

- i. null in the direction of the hinge line,
- ii. maximal in the orthogonal direction (on the tangent plane).

To illustrate this, values of different directions for second-order derivative of an arbitrary point of a hinge line (crest here) of the synthetic case (Sec. 3) are shown on Fig. 2. The same pattern can be observed for all points on the crest, with a different value for the direction angle ϕ .

2.5.2 Demonstration and application scope

Concerning the first observation, this empirical observation is supported by a theoretical demonstration developed in Appendix C for particular folds (Fig. 1). The only requirement is that the axial poles n along the hinge line are colinear, or equivalently that there exists a plane to which they are all normal, defined here as the 'tangent plane' Pt .

Here the study focuses on an upright fold, with vertical axial surface Sa . In this case, the hinge line of the studied iso-potential surface coincides with the crest of the fold. However, the demonstration that the second-order directional derivative is null along a particular curve only relies on the existence of the tangent plane along this curve. This means that the conclusions are the same for inclined folds, as long as the axial poles n are still colinear. In the case of cylindrical folds, in particular, the second-order directional derivative is null along any line parallel to the hinge line.

The demonstration provided in Appendix C shows that the property stands for both syncline and anticline folds. Moreover, the axial surface and hinge line may freely vary in direction in the tangent plane and hence characterize anisotropy of the fold, irrespective of its shape. This opens up new constraints for fold axis data, restricted in the present work to particular types of poly-deformed folds, but that could conceivably be extended to more complex structures: even in cases where the hinge line is not a planar curve, if a portion satisfies the tangent plane requirement, we can incorporate the null second-order directional derivative information along it at that location.

The second-order derivative of a three-dimensional function gives information about the local behavior of this function, and thus intrinsically, for implicit approaches, on resulting isosurface curvature. This could be a significant information, especially in case of very curvy hinge lines. Indeed, the geometric meaning of these two previous observations is then:

1. a null value of second-order derivative of potential in the hinge line direction, implies local flatness in the hinge direction
2. the maximal value of directional second-order derivative of potential in the orthogonal direction to the hinge line on the tangent plane means that the curvature of the resulting isosurface is maximal orthogonal to the hinge direction.

2.5.3 Interest of second-order derivative data for assessing local geometry

Null first-derivative values of potential in all directions on the tangent plane already express local flatness in the hinge direction. However, the addition of a null second-order derivative value provides an additional information of 'smoothness'. Indeed,

assuming the potential Z is sufficiently smooth, its Taylor expansion at a point $\mathbf{p}$ on the hinge can be written

$$Z(\mathbf{p} + \mathbf{h}) = Z(\mathbf{p}) + \nabla Z(\mathbf{p}) \cdot \mathbf{h} + \mathbf{h}^t D^2 Z(\mathbf{p}) \mathbf{h} + O(\|\mathbf{h}\|^3).$$

For $\mathbf{h}$ in the tangent plane, the first-order term $\nabla Z(\mathbf{p}) \cdot \mathbf{h}$ is zero. If the hypotheses of Sec. 2.5.1 are satisfied and $\mathbf{h}$ is along the hinge, the second-order term $\mathbf{h}^t D^2 Z(\mathbf{p}) \mathbf{h}$ is null as well, so that $Z(\mathbf{p} + \mathbf{h}) - Z(\mathbf{p})$ is a third-order infinitesimal; by providing the information $D_\tau^2 Z(\mathbf{p}) = 0$ for τ along the hinge line, one can expect the resulting model to be smoother and then better constrained.

2.5.4 Fold axis data sampling

These mathematical properties are especially relevant for already existing field data, such as fold axis data. This data consists in direction measurements (azimuth and plunge values). In this section, we consider fold axis data sampled on the field. These data are often largely rather interpreted than measured on geological maps but rarely integrated, however, into final models. They represent unit localized vectors and do not necessarily belong to the hinge line. Indeed, they may belong to the axis of another sub-parallel surface of the folded system, but this information is still a relevant constraint about the folded structure for the scalar field.

This direction measurement may be sampled in different manners (Fig. 1):

1. Fold axis may be measured directly on the axial surface (D_A and $D_{A'}$), on one hinge line. This hinge zone is a fragility zone and is not often well preserved, these observations are then pretty rare, and thus valuable.
2. In fact, fold axis data come abundantly from intersection lineations measurements (D_B). These lines are the result of the intersection between bedding planes (S_0) and cleavage planes (S_1), which commonly takes place in folding processes.
3. Fold axis data may also be inferred from orientation measurements of fold limb normal vectors (gradient data), called pole vectors. Then, stereographic projection tools can be used to determine local fold axis data (D_C measurement from n and n' pole vectors).

The first measurement is directly positioned on one hinge line. In practice, even though the latter are measures at other locations, they are commonly replaced at the hinge line location. Then, this paper studies the use of these three raw measurements on the hinge line.

However, wherever they come from, these fold axis data measurements significantly contribute to the global hinge line interpretation over the domain. This global interpretation is performed by geologists from field data but also other observations with various interpretative values, such as micro-scale structures, parasitic folds, global regional context and expert knowledge. For a complete review of fold data see Hudleston and Treagus [2010], McClay [2013].

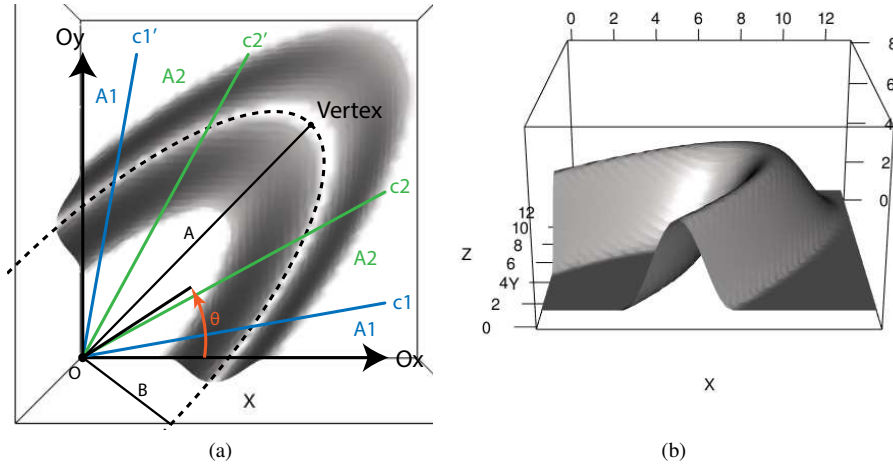


Fig. 3: Top (a) and side (b) view of the iso-surface $Z = 2.5$ of the potential of Eq. (7). The elliptic hinge line of the fold, with semi-major and semi-minor axes, A and B , is presented on (a) (dashed line). Random data for different samplings are generated with the angle θ around this structure. Data sampling zones, $A1$ and $A2$, with their respective cut off angles, $c_1 = \frac{2\pi}{24}$, $c_2 = \frac{4\pi}{24}$, $c'_1 = \frac{8\pi}{24}$, $c'_2 = \frac{10\pi}{24}$, are also shown, as well as the vertex of the ellipse.

2.5.5 Covariance function requirements

Implementation of second-order derivative value of potential leads to the need of a four-times differentiable covariance function (see Appendix B.2 for more details on successive derivations of the covariance function). Thus, the classically implemented cubic covariance model is not applicable; here the considered model is one of the ‘piecewise polynomial, positive definite and compactly supported radial functions of minimal degree’ introduced by Wendland [1995], denoted by Wendland C4 in the following, and which may be expressed as

$$C(r) = \frac{1}{3}(1-r)_+^6(35r^2 + 18r + 3). \quad (6)$$

3 Case study

3.1 Presentation of the considered potential field

In this paper, an explicit theoretical potential function has been chosen as case study. Shaped to represent an overturned anticline structure, the hinge line – or crest, since they coincide here – describes an ellipse trajectory of semi-major and semi-minor

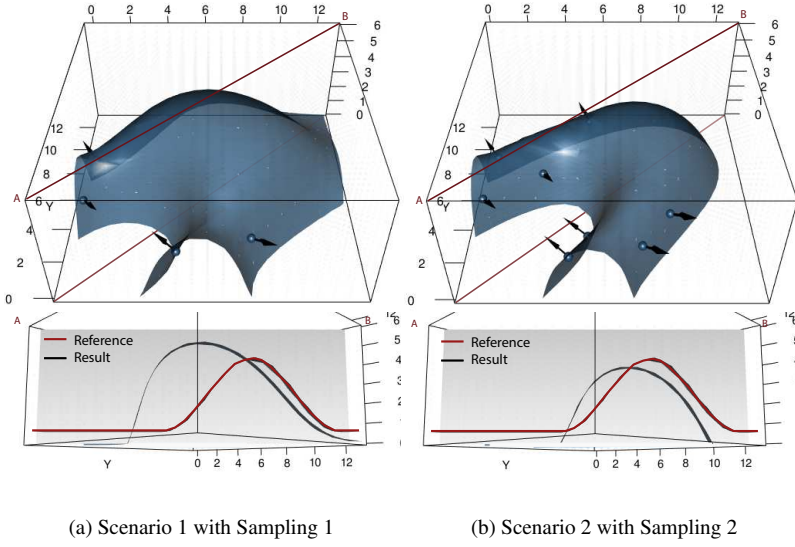


Fig. 4: Initial results of the potential field method on two different initial samplings. Sampling 1 consists in 4 contact points (blue spheres) and 4 gradient data (black arrows). Sampling 2 consists in 8 contact points and 8 gradient data.

axes A and B (Fig. 3). Therefore, this trajectory represents the main continuity direction of the structure, defined as anisotropy, here significantly varying in space. The explicit potential function can be expressed as

$$Z(x,y,z) = \begin{cases} z - \alpha \sin(\omega d) & \text{if } d \in [l - \varepsilon ; l + \varepsilon], \\ z & \text{otherwise,} \end{cases} \quad (7)$$

where

- $d = \sqrt{a(x-y)^2 + b(x+y)^2}$,
- $\alpha = 1.9$, $\omega = 2$, $a = 0.5$ and $b = 0.05$, with $a = \frac{1}{B^2}$ and $b = \frac{1}{A^2}$,
- l and ε being chosen so that $\sin(\omega d) = 0$ for $d = l - \varepsilon$ and $d = l + \varepsilon$, leading to continuous isosurfaces: $\varepsilon = \frac{\pi}{\omega}$ and $l = \frac{5\pi}{2\omega}$.

Only one iso-potential surface of this field at level $Z(x,y,z) = 2.5$ is shown on Fig. 3: this surface is defined as our reference of study thereafter.

3.2 Initial results on different samplings

In the implicit modeling framework, an anisotropic structure like this one requires a dense sampling, in terms of quantity and localization of data, in order to obtain an admissible model. The following section will highlight interpolation issues encountered in such a sparse data environment.

3.2.1 Random sampling strategy

In basic geological structural studies, sampled data are usually contact points and gradient data. In that respect, considering the initial potential field method, Fig. 4 presents two different resulting interpolated surfaces, based on two different initial samplings. These samplings have been randomly generated in different zones of the structure (denoted A1 and A2 zones on Fig. 3). Resulting models of the potential field method on each sampling is presented in Fig. 4.

- Scenario 1 (Fig. 4a): The first sampling consists in 4 random contact points (with 4 associated gradient data) generated in the A1 zone.
- Scenario 2 (Fig. 4b): The second sampling is defined by 4 random contact points (with 4 associated gradient data) generated in the A1 zone, and 4 random contact points/gradient data generated in the A2 zone.

As a consequence, both sampling density and area coverage of data are higher in Sampling 2 compared to Sampling 1.

3.2.2 Interpolation parameters

In the following, the implemented covariance function is the Wendland C4 (Eq. 6), the range of the model is equal to the diagonal of the model (15) and a vertical drift is considered here

$$m(\mathbf{p}) = v z(\mathbf{p}).$$

The interpolation is performed on a 20-by-20-by-20 grid (although for numerical reasons the error score V_e (Sec. 3.2.3) is computed on a 20-by-20-by-80 grid). Although true values of gradient data are known in this synthetic example, gradients are defined unitary as a default choice, as it is implemented that way in software packages (Sec. 2.2).

3.2.3 Modeling quality indicators

This work focuses mainly on the qualitative aspect of the resulting three dimensional models, as it is generally performed in geological studies. For this purpose, the three-dimensional isosurface of interest is presented for each experiment thereafter. In addition, in order to help visualization, the resulting surface of interpolation is compared to the original theoretical surface using a vertical cross section along the ‘first diagonal’ ($x = y$, profile A-B on Fig. 4a). This cross section is shown underneath each corresponding three dimensional model, even though the cross section direction is not always displayed. However, a quantitative measure, the volume error V_e , is provided. This indicator estimates the volume difference between the theoretical and modeled surfaces in the region near the structure of interest; its computation is detailed in Appendix D. This measure, made possible here by the relative simplicity of the synthetic model, can be distorted by numerical errors, notably due to the grid resolution in z . An uncertainty value ΔV is computed and provided for each model, and has

to be taken into account when comparing values of the indicator V_e . In the end, the investigations performed in this work can be seen as attempts to minimize this error score. These different scores are reported in appendix D. To express improvements of the resulting model compared to initial results in the same scenario i (with its own volume error V_{e_i}), the relative volumic improvement (RVI) ηV_e is introduced

$$\eta V_e = \frac{V_{e_i} - V_e}{V_{e_i}} \quad \text{where } V_{e_i} \text{ is the volume error of scenario } i; \quad i \in (1, 2)$$

This signed indicator ηV_e quantifies the reduction of error compared to the initial model. It is expressed as a percentage. Negative values mean deterioration compared to initial results, 100% RVI means that the error is now null, whereas 0% RVI means no change in volume error value. More details are provided in Appendix D.

These criteria, both qualitative and quantitative, have to be handled simultaneously in order to evaluate the results and then provide better insight on the scalar field variations. All resulting values for the different experiments are summarized in appendix D.4.

3.2.4 Critical case versus well-constrained scenario

Resulting surfaces of the potential field method with the two different samplings of Sec. 3.2.1 are shown in Fig. 4. This interpolation is performed with the choice of parametrization detailed in Sec. 3.2.2. These two starting assumptions lead to two different results:

The resulting surface with Sampling 1 (Fig. 4a) presents a ‘bubbly’ aspect far from the data, as it is generally observed with implicit modeling approaches in sparse data situations. With no data around, the potential value tends towards the drift (mean) value. This result is actually not satisfactory for geologists. The volumetric error is $V_{e_1} = 104.68$. Moreover, this phenomena takes place here with use of Wendland C4 covariance model, but also occurs with the more usual cubic covariance one. This example may be considered as a critical case scenario, where fold anisotropy introduction would significantly benefit the model. Proposed solutions will be evaluated in comparison with this vertical error score, which should never exceed this critical value.

On the other hand, the resulting surface obtained with the better sampling, Sampling 2, is shown in Fig. 4b. This sampling choice leads to a better result, as could be expected. The volumic error is $V_{e_2} = 48.36$, which is 56.32 better than the previous scenario. This change of scenario improves the result by 54%. Hence this scenario may be considered as a better constrained model. Moreover, despite this prevailing error, this result may be qualitatively considered as satisfying and thus, this scenario considered as a well constrained model. This can be defined as a benchmark for further comparisons with improvements made on the previous case. Still, the impact of additional anisotropy information data will also be tested on this model in order to see the impact of these new data on a satisfying well constrained model, but where there is still possible room for improvement.

These two sampling starting assumptions are mentioned as Scenario 1 and Scenario 2 in the following.

3.2.5 Exploratory methodology

The first scenario is not satisfying, especially considering the fact that in such situations we observe that geologists actually already do hold anisotropy information: either localized fold axis data or a final interpretation of the hinge line (Sec. 2.5.4). The parsimonious deployment of both kinds of inputs is discussed in Sec. 5.

However, the following results on the exploitation of fold axis data fit for both approaches. Indeed, the integration of the hinge line may be done by discretizing it and then use punctual fold axis data, which is the correct mathematical expression of this information. Nevertheless these conclusions raise different outstanding issues:

- i. Whereas in theory null second-order derivatives yield additional smoothness information compared to first order derivative values, in practice for hinge line conditioning, is it a real relevant constraint to add?
- ii. Although the use of tangent data was already possible in the initial method of [Lajaunie et al., 1997], in practice, the use of such constraints for fold geometry modeling is not necessarily always adopted by geologists, even though data of different types mentioned in Sec. 2.5.4 are available. When tangent data are taken into account, there is still uncertainty about the necessary amount of them to well constrain poly-deformed folds in sparse data environment, as well as a geographical distribution preference. These understandings would allow to define priorities for either the direct geological sampling strategy of fold axis data or the parametrization of an interpreted hinge line.
- iii. Considering the tangent plane, for a given data field measurement (fold axis) on the hinge line, different geometric constraints may be taken into account: at least two tangent data, a null second-order derivative value and a maximal second-order derivative value. The relevance of each one should be investigated.

4 Investigations on better fold axis data integration

In the following, these three issues are addressed on the synthetic case regarding the two previous scenarios proposed in Sec. 3. The hinge line direction is known; different constraints are imposed upon it depending on the experiment. All the resulting surfaces are compared to the theoretical one (Fig. 3) through cross sections just as presented in Sec. 3.2, even though its orientation is no longer indicated. The different volumetric error indicators are listed in appendix D.4.

4.1 Use of fold axis data as tangent data, null second-order derivative data or both

On the hinge line, a null first-order derivative value of potential constraint may be imposed for fold axis vector data. New observations made in Sec. 2.5.1 motivated the addition of a null second-order derivative value of potential constraint as well.

In order to characterize the contribution of each constraint in practice, Fig. 5 shows the impact of such data on the synthetic case, regarding the two different starting scenarios. 15 homogeneously distributed vectors are sampled along the hinge line

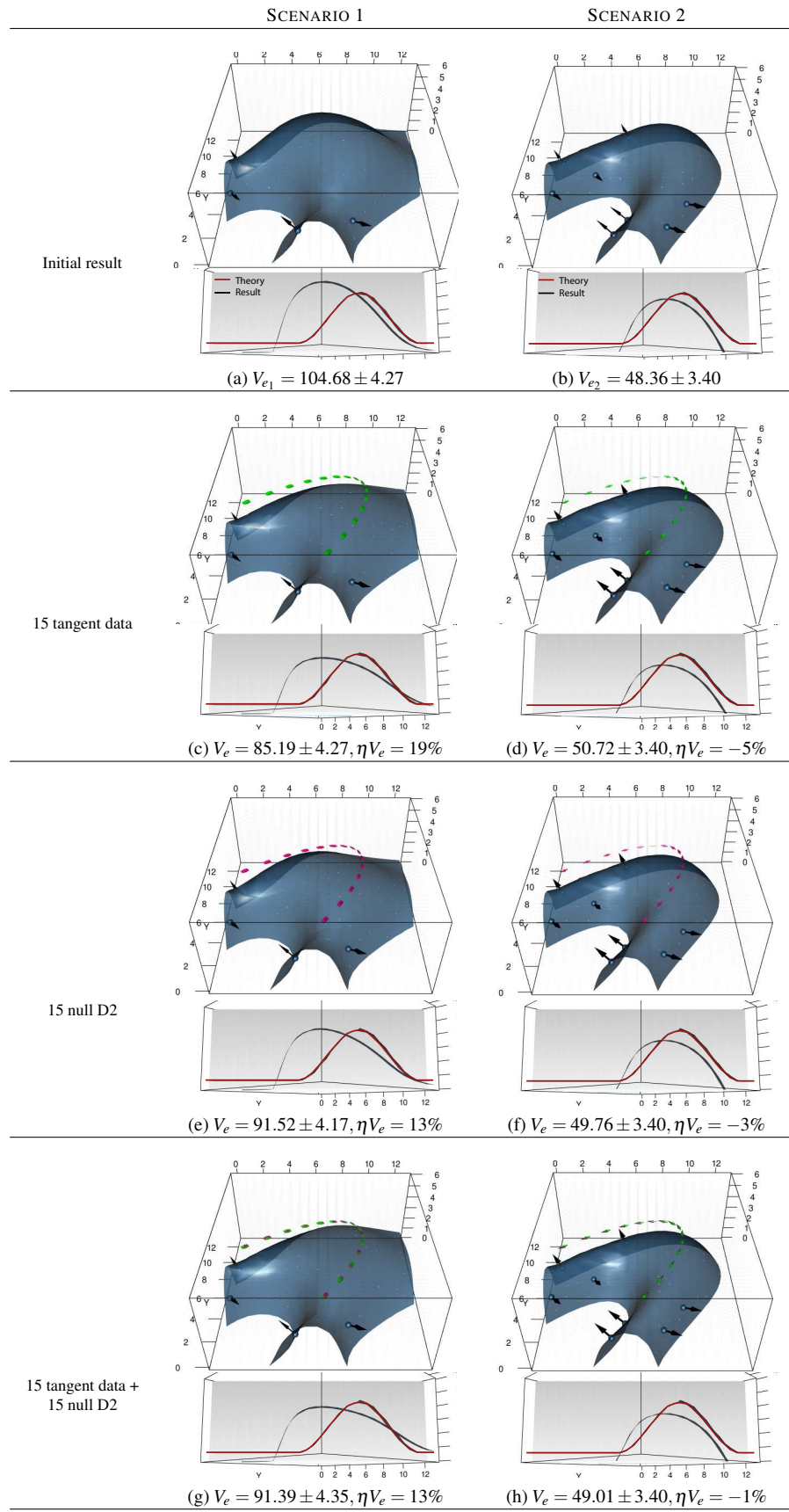


Fig. 5: Impact of (c)(d) 15 tangent data (green arrows), (e)(f) 15 null second-order derivative data (D2, pink arrows), (g)(h) both data on the resulting surface compared to (a)(b) the initial result with Sampling 1 or Sampling 2

and expressed in terms of tangent (Figures 5c, 5d), null second-order derivative (D2) (Figures 5e, 5f) or both (Figures 5g, 5h). For comparison, the initial results from Sec. 3.2.4 are also recalled (Figures 5a, 5b). The conclusions are the following:

- Scenario 1: The introduction of null second-order derivative constraints does have an impact on the resulting model (Fig. 5e), in comparison with the initial result of this scenario (Fig. 5a). These constraints appear to flatten the unwanted bulging as expected; the resulting surface draws near the theoretical one. V_e is 92 ± 4 , RVI is around 13%. However, the introduction of tangent data constraints alone (Fig. 5c) visually leads to a better result: the modeled hinge line seems to be localized more accurately. In addition, V_e is 85 ± 4 , RVE is around 19%. Finally, the combination of both constraints (Figures 5g) leads to a qualitatively similar result that the one with tangent data constraints only. However, V_e is 91 ± 4 , RVI is around 13%, which is closer to the result with null second order derivatives. The differences between these three results fall within the uncertainty bounds given by ΔV : direct comparisons should be qualitative in nature.
- Scenario 2: In this specific well constrained case, the RVI scores of these four comparisons indicate that the four models are equivalent (within the uncertainty range, at this resolution). Qualitatively, these models do not show perceptible differences either. As a conclusion, in this well constrained scenario, the addition of either tangent or null second-order derivative data does not seem necessary.

In the end, these different results do not bring to light the interest of the introduction of second-order derivative value made on Sec. 2.5.3. This could possibly be explained by the pre-existence of local flatness, or at least smoothness, of the model preceding its introduction. Actually this smoothness is generally inherent in geological surfaces. As a matter of fact, this local feature is the main assumption of geological modeling interpolation methods. However, second-order derivative values could bring relevant information in ambiguous surfaces in other surface modeling applications.

4.2 Quantity or spatial distribution of tangent data impact

In response to observations made on Sec. 4.1, it still appears relevant to sample fold axis data and incorporate them in terms of tangent data alone at least. The impact of the modalities of this integration, such as quantity and spatial distribution remain ill-known. For this purpose, the impact of quantity and spatial distribution of these data has been studied on the synthetic example.

4.2.1 Impact of the quantity of tangent data

The impact of the quantity of tangent data along the hinge line is described in Fig. 6. 29 initial tangent data regularly spaced along the hinge line have been decreased to 8, then 4 data points only.

- Scenario 1: The results show that the quality of the resulting model remains the same for these three numbers of tangent data values. The relative improvement

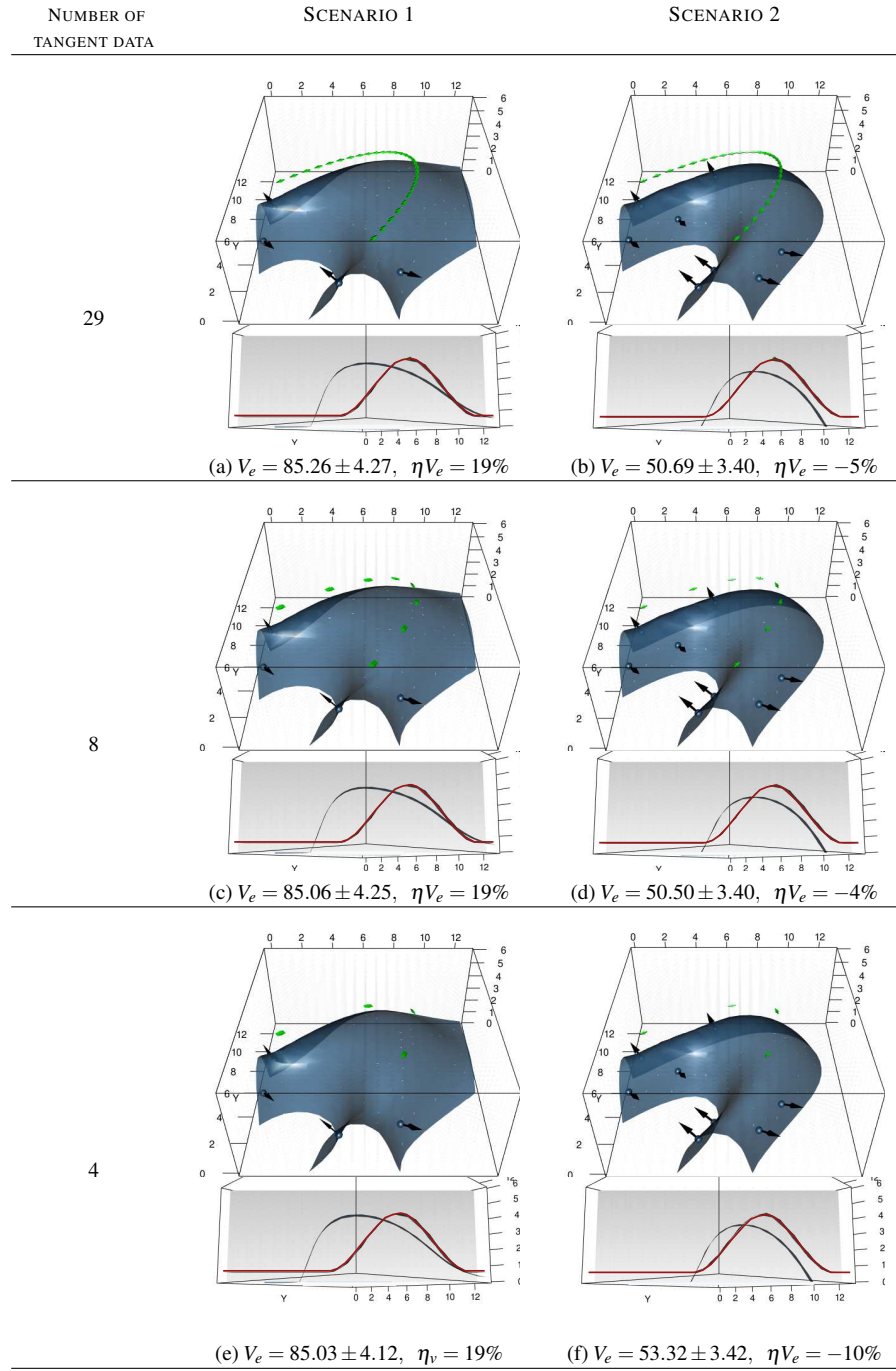


Fig. 6: Impact of the number of tangent data along the hinge line in the two different scenarios

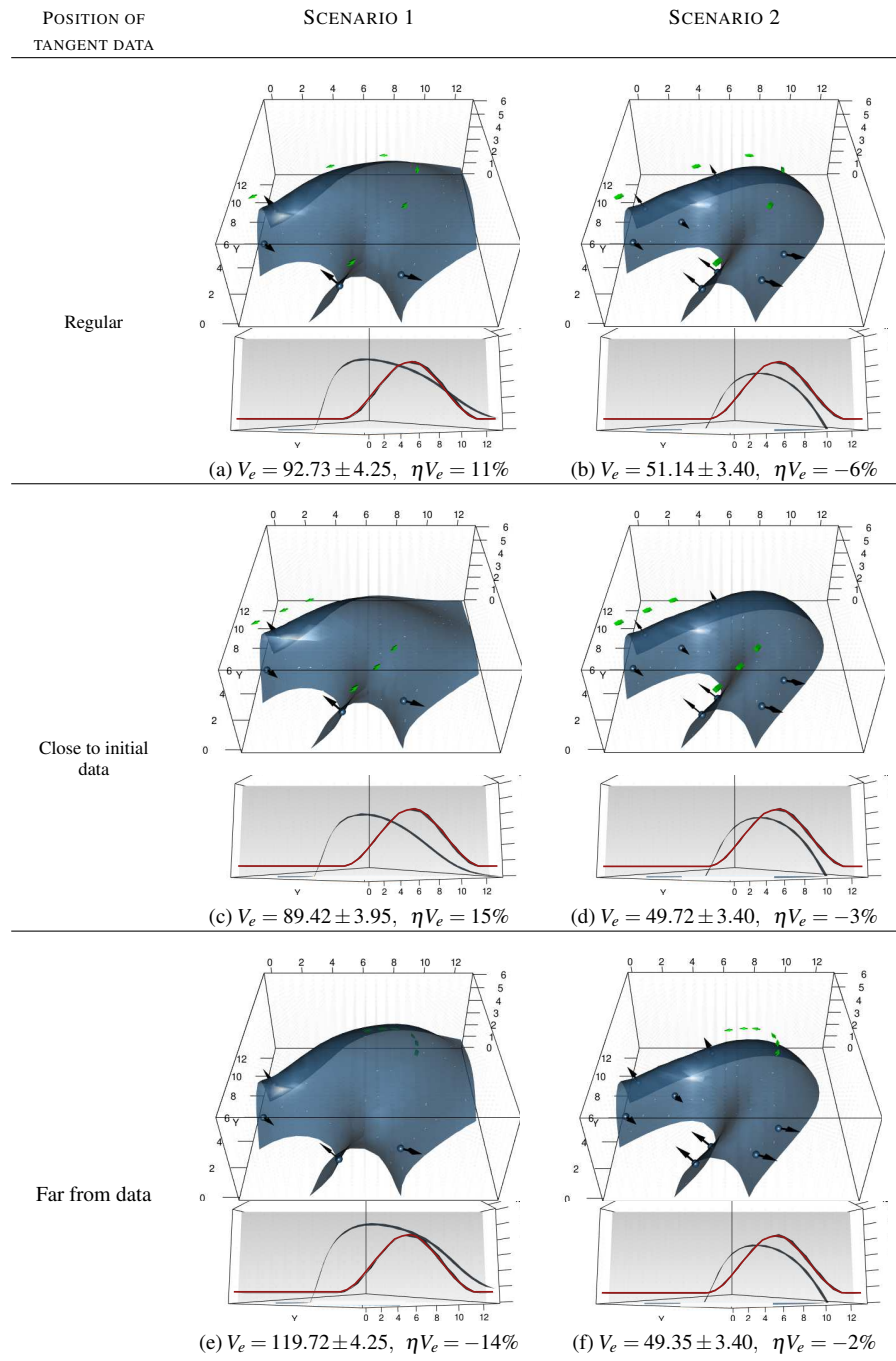


Fig. 7: Impact of the spatial distribution of the tangent data along the hinge line in the two different scenarios. 6 tangent data are poised in different ways: (a)(b) regularly spaced along the hinge line, (c)(d) close to the initial data, (e)(f) far from the initial data

(RVI) is about 19% for 29 tangent data (Fig. 6a), 8 data (Fig. 6c) or 4 data (Fig. 6e). This principally suggests that the integration of a limited number of tangent data already shows a positive impact on the result, thus a fine distribution of these tangent data is not necessary, at least in this particular case.

- Scenario 2: In the well-constrained scenario, the number of tangent data does not seem to have a significant impact on the result. These different constraints lead to similar results (Fig. 6b, Fig. 6d, Fig. 6f) on the edge of the indicator's resolution. These statements are also similar to the ones observed in the previous experiment with derivative constraints (Sec. 4.1).

4.2.2 Impact of the spatial distribution of tangent data

Now considering a fixed number of tangent data, the impact of the localization of such data along the hinge line is described by Fig. 7. Three major sampling strategies have been adopted here: new tangent data regularly spaced along the hinge line, new data close to the initial inputs (contact points and gradient data) and new data far from the initial inputs. Adding anisotropy information in the last location could potentially be of some specific interest, since this zone is the ellipse's vertex area (Fig. 3) where the anisotropy direction varies strongly, in comparison to the whole structure.

- Scenario 1: With a constant number of tangent data, the results show that the use of these data close to initial contact points and gradient data (Fig. 7c), or regularly spaced (Fig. 7a), qualitatively give the best results, by flattening the budging around the hinge line. The quantitative improvement of these strategies is around 11% and 15%, respectively. However, regularly spaced data visually seem to improve the result mainly around the vertex area, whereas tangent data close to initial data improve the result around these well-constrained zones. This highlights that local information of flatness is given. On the contrary, additional tangent data placed far from contact points and gradient data (Fig. 7e) make the modeling worse by -14% . This could be explained by the contradiction between the forced local flatness given through six tangent data constraints at the vertex position here and the forced unitary norm given by gradient data which may lead to some unexpected results (Sec. ii.).
- Scenario 2: The addition of tangent data does not significantly affect the resulting model whatever position of these data is (Figures 7b, 7d, 7f). RVI values are similar, within the limit of their uncertainties.

As a conclusion, with respect to this only case study, it seems that a parsimonious regularly spaced sampling of tangent data is the most cautious strategy to adopt. Adding anisotropy information where it is particularly needed could, counter-intuitively, worsen the modeling, possibly because of the gradient norm bias.

4.3 Insights on orthogonal constraints of fold axis data on the tangent plane

The previous sections, 4.1 and 4.2, have considered only fold axis data directions, i.e. the hinge line direction. However, it is possible to take into account the orthogonal

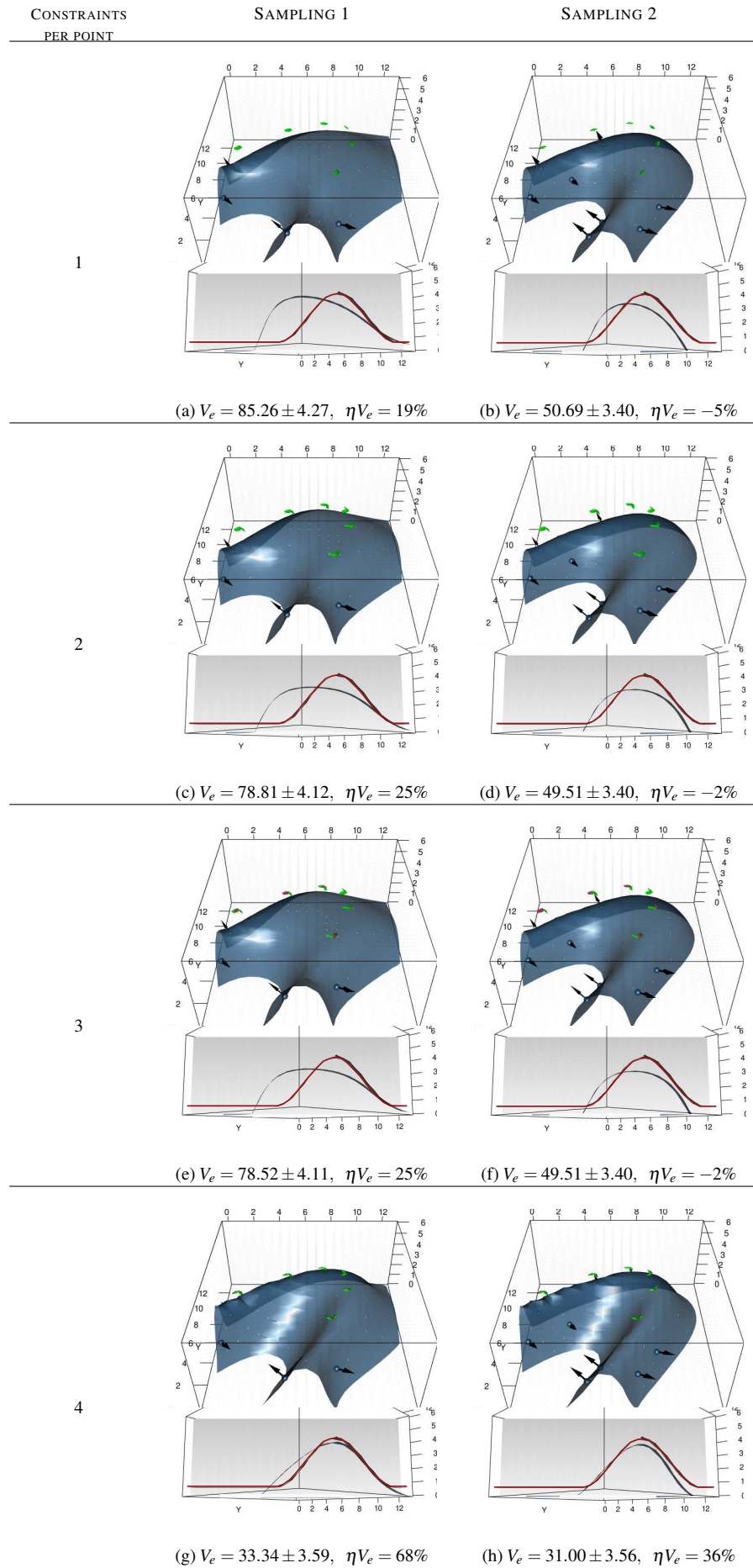


Fig. 8: Comparison of the impact of (first row) 1, (second row) 2, (third row) 3 and (last row) 4 constraints per point at the same 6 locations on the fold axis of the theoretical surface

direction on the local tangent plane (Fig. 1) of these fold axis data. This consideration would be relevant for two major reasons. First, the first-order derivative value is null in this direction. In practice this second tangent data is sometimes already integrated in the model, completed with a younging direction, which indicates the polarity of the normal of the tangent plane. This is performed in order to express pole vectors (Fig. 1) on the hinge line. Indeed, this vector shall not be provided to the model as a gradient input, because of the gradient norm bias. In the end, more insights on this tangent constraint impact would benefit to the understanding of its relevance. Second, the observations made in Sec. 2.5.1 revealed that this direction is a maximum value of the second-order derivative. This new constraint could have potential interest in search of the valuation of fold axis data in the enhancement of geometric control of fold structures.

For this purpose, Fig. 8 shows the consecutive impact of these four constraints, first and second-order derivatives in the two mentioned directions, for 6 fold axis data on the crest line:

- Scenario 1: The use of tangent data in the hinge line direction only (Fig. 8a) leads to the same observations previously made in this case (Sec. 4.1,4.2). However, the addition of the second tangent data (Fig. 8c) seems to conduct to a flatter hinge line, but not especially better positioned. The addition of null-second order derivative in the hinge line direction (Fig. 8e) does not add an appreciable qualitative or quantitative difference. Finally, the incorporation of second-order derivative values in the orthogonal direction (Fig. 8g) leads to noticeably better results; the resulting surface appears to be closer to the theoretical one than it has ever been so far, and the RVI score is around 68%. The volumic error V_e is even better than the initial result in the well-constrained scenario S2 (33.34 instead of 48.36).
- Scenario 2: In this well constrained scenario, there is no significant difference between the first three constraint influences, i.e. the addition of two tangent data (Figures 8b, 8d) and null second-order derivative value (Fig. 8f). Just as in the case of Scenario 1, the final addition of second-order derivative values in the orthogonal direction (Fig.8h) conducts to a very satisfying model. Qualitatively, the resulting surface is very close to the theoretical one, closest than it has been at any time for the different attempts in this work. RVI score is around 36%, whereas the previous results with this same sampling did not show positive improvement scores.

The observations performed on this single synthetic case lead to several conclusions. In addition to tangent data in the hinge line direction, the integration of orthogonal tangent data does not appear in this case to add sufficient information to constrain the fold structure as expected. The conclusions on the addition of null second-order derivative data are the same as the ones previously made (Sec. 4.1: these data do not affect significantly the model. However, the addition of the second-order derivative values in the direction where they are maximal, orthogonal to the hinge line, seems to affect in a very significant way the results; at least as it could be expected by users in such situation.

Unfortunately, in practice it is not possible to directly sample these values on the field. Then, the immediate integration of this constraint on real cases for now would only be possible by an empirical approach. Indeed, the value of this constraint controls the orthogonal curvature, hence the local ‘pinch’ of the fold. This value could potentially be estimated by a trial-and-error approach with visualization tools, for example. However, these observations are worth exploring, and should open a path for further mathematical developments on the integration of a local maximum as a criterion input for implicit methods.

5 Conclusion and discussion

This paper tackles a well-known issue encountered in the potential field method: when the data is sparse, classical models can lead to unsatisfactory results (e.g. presenting a ‘bubbly’ aspect). This is particularly the case when the geological field is made of complex folded structures. In such situations, geologists are forced to add artificial control points. The alternative proposed here is to incorporate information about the anisotropy. It can be characterized by fold axis data, which have the advantage of being readily available in most fold structural studies. This approach is motivated from a mathematical point of view in Sec. 2, using the potential method [Lajaunie et al., 1997] as a reference. This allows to describe poly-phased structures in a purely geometric, parsimonious manner, without having to make strong assumptions on the underlying succession of geological events. Another asset of this method is that the covariance function used for the interpolation of the potential scalar field can remain isotropic.

A challenging case study is described in Sec. 3. It serves in Sec. 4 as a founding example to gain more insight into the usage of fold axis data for characterizing the anisotropic geometries of folds. Three major conclusions are drawn from this exploration.

- i. From a mathematical point of view, the second-order derivative values of the three-dimensional scalar field are shown to be of particular interest. They bring theoretical information on the curvature of the surface, notably where it is flat and might coincide with the hinge line. However, in the chosen example, adding null second-order derivative constraints to the model only slightly improves the results. When combined with tangent data, almost no improvement is observed. Obviously, no generalization can be drawn from a single example. Nonetheless, it might be conjectured that the effect of these mathematical objects are bound to be limited in other geological studies too. Indeed, one of the most important features characterizing complex iso-surfaces is their smoothness. It is at the foundation of geological interpolation methods, which try to minimize the curvature. When dealing with fold structures, curvatures turn out to be quite regular, especially in the hinge line direction. Therefore, it is our belief that tangent data suffice to capture this key characteristic.
- ii. Tangent data turn out to have a significant impact on the quality of the results; when ignored, undesirable effects are observed. Even with a very small number of such data, fold anisotropy integration is notably improved. It is also observed

that their effect is all the more positive as fold axis vectors are sampled close to the initial gradient and contact point data.

- iii. The value of the second-order derivative in the orthogonal direction of the hinge line (i.e. the maximum value of second-order derivatives) appears to be particularly useful. Models incorporating this information in a synthetic example show promising results. Unfortunately, this value is hardly reachable in practice, although it could conceivably be estimated if a cross-section is available. Alternatively, an arbitrary value could be chosen, which would control the local pinch of the fold. Such observations suggest to investigate the incorporation of a Gaussian curvature condition in the model; this is left for future research.

All these remarks hold for discrete fold axis data. They usually come from raw measurements, but could also be generated from the discretization of an interpreted hinge line. Indeed, the ontological boundary between these two types of model input is not immediate: the interpretation of the hinge line stems from a combination of fold axis data and other geological observations such as micro-scale structures or the regional context. In practice, the interpreted hinge line could be directly integrated as a sketch (expert interpretation) or automatically interpolated from some direction measurements (automatic interpretation). Then, it could be discretized with synthetic fold axis data. A non-negligible limitation of this technique is that the uncertainty linked to such artificial input is generally intractable. Nonetheless, interpreted data hold as much geological value as field measurements [Frodeman, 1995]. Therefore, favoring either one over the other should remain an expert choice, based on practical considerations such as the desired model complexity, the state of play, or the targeted applications. It goes beyond the methodological scope proposed here.

In any case, using a discretized version of the interpreted hinge line remains more desirable, in our opinion, than introducing purely arbitrary control points (e.g. contact points or gradient data). Indeed, the latter generate complex, non-parsimonious models of the anisotropy. In particular, contact points force the surface to pass through a particular point; this condition can be viewed as too restrictive. In addition, gradient data control the contraction of the scalar field (isosurfaces get closer to each other with the gradient norm) and thus introduce the gradient norm bias. The interpreted hinge line, on the other hand, gives some degree of freedom to the model and solves the gradient norm bias by adding tangent data. By focusing on the two leading types of information that characterize the anisotropy, namely the null and maximum curvatures, it produces a parsimonious realistic model. The number of tangent or second-derivative data to add to the model must be pondered relative to the induced computational complexity, since each constraint adds one line to the co-kriging system.

Another way of incorporating the hinge line into the model, which was not considered here, would be to use it as a drift function. This would require an explicit analytic formula describing the hinge line, instead of a collection of discrete values. In practice, it is possible to create an explicit fold potential from an initial three-dimensional curve sketch using the distance to a skeleton [Henrion et al., 2010, Garcia et al., 2018]. This approach brings geological modeling and computer graphics

techniques closer together, but quantifying the uncertainty of such data remains an open question. This is left for future research.

Acknowledgements The authors would like to thank Gabriel Courrioux of BRGM for being a pioneer of the geomodeler software and for initiating this research work with his colleague, Bernard Bourguine. We are particularly grateful to Christian Lajaunie, inventor of the geostatistical potential field method, who largely participated in the development of the consecutive derivatives of the covariance function. Support and advice from the Geostatistics Research Group, especially Emilie Chautru for her careful editing, were also helpful, just as the ones from the Geomodeling Team of BRGM. Authors thank Michael Hillier, Guillaume Caumon and the anonymous reviewer for their fruitful comments.

A Formulation of the potential field method in terms of kriging equations

Whereas the seminal paper by [Lajaunie et al., 1997] adopted a dual (co-)kriging formulation, we retained here the universal (co-)kriging formulation.

The interested reader may consult the chapter 3.4 of the book by [Chiles and Delfiner, 2009] to have more details about this derivation.

Equation 2 writes:

$$[Z(\mathbf{p}) - Z(\mathbf{p}_0)]^* = \sum_{i,i'}^{i'_n} \lambda_{i,i'} [Z(\mathbf{p}_i) - Z(\mathbf{p}_{i'})] + \sum_g^{g_n} \lambda_g Z'_g(\mathbf{p}_g) + \sum_s^{s_n} \lambda_s Z''_s(\mathbf{p}_s) \quad (\text{A.8})$$

where the estimate $[Z(\mathbf{p}) - Z(\mathbf{p}_0)]^*$ is a linear combination of:

- i'_n increment data,
- g_n first-order directional derivative of potential data (tangent and/or components of gradient data),
- s_n second-order directional derivative of potential data,

and where for the sake of simplicity, we also index the directional derivatives Z' and Z'' with g and s rather than with directions τ_g and τ_s :

$$\begin{aligned} Z'_g(\mathbf{p}_g) &= D_{\tau_g} Z(\mathbf{p}_g), \\ Z''_s(\mathbf{p}_s) &= D_{\tau_s}^2 Z(\mathbf{p}_s). \end{aligned} \quad (\text{A.9})$$

Similarly, in the following, in summations of first and second-order directional derivatives of any function, we will drop the directions and index by g and s as well.

The estimation error ε is defined by

$$\begin{aligned} \varepsilon &= [Z(\mathbf{p}) - Z(\mathbf{p}_0)]^* - [Z(\mathbf{p}) - Z(\mathbf{p}_0)] \\ &= \sum_{i,i'}^{i'_n} \lambda_{i,i'} [Z(\mathbf{p}_i) - Z(\mathbf{p}_{i'})] + \sum_g^{g_n} \lambda_g Z'_g(\mathbf{p}_g) + \sum_s^{s_n} \lambda_s Z''_s(\mathbf{p}_s) - [Z(\mathbf{p}) - Z(\mathbf{p}_0)]. \end{aligned} \quad (\text{A.10})$$

A.1 Universality conditions

Universality conditions ensure that the estimation is unbiased in the sense that the expectation of the estimation error is null *i.e.* $E(\varepsilon) = 0$.

Noting that

$$\begin{aligned} E[Z(\mathbf{p}) - Z(\mathbf{p}_0)] &= m(\mathbf{p}) = \sum_{\ell}^n v_{\ell} f^{\ell}(\mathbf{p}), \\ E[Z'_g(\mathbf{p}_g)] &= m'_g(\mathbf{p}_g) = \sum_{\ell}^n v_{\ell} f^{\ell'}_g(\mathbf{p}_g), \\ E[Z''_s(\mathbf{p}_s)] &= m''_s(\mathbf{p}_s) = \sum_{\ell}^n v_{\ell} f^{\ell''}_s(\mathbf{p}_s), \end{aligned} \quad (\text{A.11})$$

with the same abbreviation logic as previously and $m'_g(\mathbf{p}_g)$ (*resp.* $f_g^{\ell'}$ ($\mathbf{p}_g$)) standing for $D_{\tau_g} m(\mathbf{p}_g)$ (*resp.* $D_{\tau_g} f^\ell(\mathbf{p}_g)$).

Then, universality conditions write:

$$\begin{aligned}
E[\varepsilon] &= E\left[\sum_{ii'}^{i' n} \lambda_{ii'} [Z(\mathbf{p}_i) - Z(\mathbf{p}_{i'})] + \sum_g^{g n} \lambda_g Z'_g(\mathbf{p}_g) + \sum_s^{s n} \lambda_s Z''_s(\mathbf{p}_s) - [Z(\mathbf{p}) - Z(\mathbf{p}_0)]\right] \\
&= \sum_{ii'}^{i' n} \lambda_{ii'} [E[Z(\mathbf{p}_i)] - E[Z(\mathbf{p}_{i'})]] + \sum_g^{g n} \lambda_g E[Z'_g(\mathbf{p}_g)] + \sum_s^{s n} \lambda_s E[Z''_s(\mathbf{p}_s)] - E[Z(\mathbf{p}) - Z(\mathbf{p}_0)] \\
&= \sum_{ii'}^{i' n} \lambda_{ii'} \left[\sum_\ell^n v_\ell f^\ell(\mathbf{p}_i) - \sum_\ell^n v_\ell f^\ell(\mathbf{p}_{i'}) \right] + \sum_g^{g n} \lambda_g \sum_\ell^n v_\ell f_g^{\ell'}(\mathbf{p}_g) + \sum_s^{s n} \lambda_s \sum_\ell^n v_\ell f_s^{\ell''}(\mathbf{p}_s) \\
&\quad - \sum_\ell^n v_\ell f^\ell(\mathbf{p}) \\
&= \sum_\ell^n v_\ell \left(\sum_{ii'}^{i' n} \lambda_{ii'} [f^\ell(\mathbf{p}_i) - f^\ell(\mathbf{p}_{i'})] + \sum_g^{g n} \lambda_g f_g^{\ell'}(\mathbf{p}_g) + \sum_s^{s n} \lambda_s f_s^{\ell''}(\mathbf{p}_s) - f^\ell(\mathbf{p}) \right) \\
&= 0,
\end{aligned} \tag{A.12}$$

which must be true for all values of v_ℓ and implies n equations (universality constraints), as many as there are drift functions:

$$\sum_{ii'}^{i' n} \lambda_{ii'} [f^\ell(\mathbf{p}_i) - f^\ell(\mathbf{p}_{i'})] + \sum_g^{g n} \lambda_g f_g^{\ell'}(\mathbf{p}_g) + \sum_s^{s n} \lambda_s f_s^{\ell''}(\mathbf{p}_s) - f^\ell(\mathbf{p}) = 0 \quad \forall \ell \in \{1, \dots, n\} \tag{A.13}$$

denoted $T^\ell(\lambda_{ii'}, \lambda_g, \lambda_s) = 0$ in the following.

A.2 Optimality conditions

Optimality conditions ensure that $\text{Var}(\varepsilon)$ is minimal.

In this part, the following convention is adopted:

$$\begin{aligned}
\text{Cov}(Z(\mathbf{p}_i), Z(\mathbf{p}_j)) &= C_{i,j}^{\bullet\bullet}, \\
\text{Cov}(Z'_g(\mathbf{p}_g), Z(\mathbf{p}_j)) &= C_{g,j}^{\bullet'\bullet}, \\
\text{Cov}(Z''_s(\mathbf{p}_s), Z(\mathbf{p}_j)) &= C_{s,j}^{\bullet''\bullet}.
\end{aligned} \tag{A.14}$$

When double indices are used, the notation will refer to increment values, for example $C_{ii',j}^{\bullet\bullet}$ will refer to $\text{Cov}(Z(\mathbf{p}_i) - Z(\mathbf{p}_{i'}), Z(\mathbf{p}_j))$.

The variance of the estimation error (A.10) is

$$\begin{aligned}
\text{Var}(\varepsilon) &= \text{Var} \left[\sum_{ii'}^{i' n} \lambda_{ii'} [Z(\mathbf{p}_i) - Z(\mathbf{p}_{i'})] + \sum_g^{g n} \lambda_g Z'_g(\mathbf{p}_g) + \sum_s^{s n} \lambda_s Z''_s(\mathbf{p}_s) - Z(\mathbf{p}) \right] \\
&= \sum_{ii', i_2 i_3} \lambda_{ii'} \lambda_{i_2 i_3} C_{ii', i_2 i_3}^{\bullet\bullet} + \sum_{g, g'} \lambda_g \lambda_{g'} C_{g, g'}^{\bullet'\bullet} + \sum_{s, s'} \lambda_s \lambda_{s'} C_{s, s'}^{\bullet''\bullet} - \text{Var}[Z(\mathbf{p})] \\
&\quad + 2 \left(\sum_{ii', g} \lambda_{ii'} \lambda_g C_{ii', g}^{\bullet'\bullet} + \sum_{ii', s} \lambda_{ii'} \lambda_s C_{ii', s}^{\bullet''\bullet} + \sum_{g, s} \lambda_g \lambda_s C_{g, s}^{\bullet'\bullet} - \sum_{ii', p} \lambda_{ii'} C_{ii', p}^{\bullet\bullet} - \sum_{g, p} \lambda_g C_{g, p}^{\bullet'\bullet} - \sum_{s, p} \lambda_s C_{s, p}^{\bullet''\bullet} \right)
\end{aligned} \tag{A.15}$$

Lagrange multipliers are used to minimize the quadratic form $\text{Var}[\varepsilon(\lambda_{ii'}, \lambda_g, \lambda_s)]$ under the constraint that all universality conditions hold *i.e.* $T^\ell(\lambda_{ii'}, \lambda_g, \lambda_s) = 0$ for all ℓ .

This implies that

$$\text{Var}[\mathcal{E}(\lambda_{ii'}, \lambda_g, \lambda_s)] + 2 \sum_l \mu_l T^\ell(\lambda_{ii'}, \lambda_g, \lambda_s) \quad (\text{A.16})$$

should have its partial derivatives with respect to each $\lambda_{ii'}, \lambda_g, \lambda_s$ and ℓ Lagrangian multipliers μ_ℓ to be nul.

The Lagrangian L is then given by

$$\begin{aligned} L(\lambda_{ii'}, \lambda_g, \lambda_s, \mu_l) = & \sum_{ii', i_2 i_3} \lambda_{ii'} \lambda_{i_2 i_3} C_{ii', i_2 i_3}^{\bullet\bullet} + \sum_{g, g'} \lambda_g \lambda_{g'} C_{g, g'}^{\bullet\bullet'} + \sum_{s, s'} \lambda_s \lambda_{s'} C_{s, s'}^{\bullet\bullet''} \\ & + 2 \left(\sum_{ii', g} \lambda_{ii'} \lambda_g C_{ii', g}^{\bullet\bullet'} + \sum_{ii', s} \lambda_{ii'} \lambda_s C_{ii', s}^{\bullet\bullet''} + \sum_{g, s} \lambda_g \lambda_s C_{g, s}^{\bullet\bullet''} \right) \\ & - 2 \left(\sum_{ii'} \lambda_{ii'} C_{p, ii'}^{\bullet\bullet} + \sum_g \lambda_g C_{p, g}^{\bullet\bullet'} + \sum_s \lambda_s C_{p, s}^{\bullet\bullet''} \right) \\ & - 2 \sum_\ell \mu_\ell T^\ell(\lambda_{ii'}, \lambda_g, \lambda_s) \end{aligned} \quad (\text{A.17})$$

Derivating this Lagrangian according to $\lambda_{ii'}, \lambda_g, \lambda_s$ and μ_l gives

$$\left\{ \begin{array}{l} \sum_{i_2 i_3} \lambda_{i_2 i_3} C_{ii', i_2 i_3}^{\bullet\bullet} + \sum_g \lambda_g C_{ii', g}^{\bullet\bullet'} + \sum_s \lambda_s C_{ii', s}^{\bullet\bullet''} + \sum_\ell \mu_\ell [f^\ell(\mathbf{p}_i) - f^\ell(\mathbf{p}_{i'})] = C_{p, ii'}^{\bullet\bullet} \quad \forall (i, i') \\ \sum_{g'} \lambda_{g'} C_{g, g'}^{\bullet\bullet'} + \sum_{ii'} \lambda_{ii'} C_{ii', g}^{\bullet\bullet'} + \sum_s \lambda_s C_{g, s}^{\bullet\bullet''} + \sum_\ell \mu_\ell f_g^{\ell'}(\mathbf{p}_g) = C_{p, g}^{\bullet\bullet'} \quad \forall g \\ \sum_{s'} \lambda_{s'} C_{s, s'}^{\bullet\bullet''} + \sum_{ii'} \lambda_{ii'} C_{ii', s}^{\bullet\bullet''} + \sum_g \lambda_g C_{g, s}^{\bullet\bullet''} + \sum_\ell \mu_\ell f_s^{\ell''}(\mathbf{p}_s) = C_{p, s}^{\bullet\bullet''} \quad \forall s \\ \sum_{ii'} \lambda_{ii'} [f^\ell(\mathbf{p}_i) - f^\ell(\mathbf{p}_{i'})] + \sum_g \lambda_g f_g^{\ell'}(\mathbf{p}_g) + \sum_s \lambda_s f_s^{\ell''}(\mathbf{p}_s) = f^\ell(\mathbf{p}) \quad \forall \ell \end{array} \right. \quad (\text{A.18})$$

which defines kriging equations and writes in matrix form:

$$\begin{pmatrix} K & F \\ F^T & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} K_p \\ F_p \end{pmatrix} \quad (\text{A.19})$$

Denoting $N = (ii'_n + g_n + s_n)$, the number of data, and p , the number of points to estimate : K is $N \times N$, K_p is $N \times p$, F is $N \times n$, F_p is $n \times p$ and 0 is a block zero matrix.

$$\begin{aligned} K = \begin{pmatrix} K_{ii', i_2 i_3}^{\bullet\bullet} & K_{ii', g}^{\bullet\bullet'} & K_{ii', s}^{\bullet\bullet''} \\ K_{g, ii'}^{\bullet\bullet'} & K_{g, g'}^{\bullet\bullet'} & K_{g, s}^{\bullet\bullet''} \\ K_{s, ii'}^{\bullet\bullet''} & K_{s, g}^{\bullet\bullet''} & K_{s, s'}^{\bullet\bullet''} \end{pmatrix} \quad K_p = \begin{pmatrix} K_{ii', p}^{\bullet\bullet} \\ K_{g, p}^{\bullet\bullet'} \\ K_{s, p}^{\bullet\bullet''} \end{pmatrix} \quad F_p = \begin{pmatrix} f^1(\mathbf{p}_1) & \dots & f^1(\mathbf{p}_p) \\ \dots & & \dots \\ f^n(\mathbf{p}_1) & \dots & f^n(\mathbf{p}_p) \end{pmatrix} \\ F = \begin{pmatrix} f^1(\mathbf{p}_i) - f^1(\mathbf{p}_{i'}) & \dots & f^n(\mathbf{p}_i) - f^n(\mathbf{p}_{i'}) \\ f_g^{1'}(\mathbf{p}_g) & \dots & f_g^{n'}(\mathbf{p}_g) \\ f_s^{1''}(\mathbf{p}_s) & \dots & f_s^{n''}(\mathbf{p}_s) \end{pmatrix} \end{aligned} \quad (\text{A.20})$$

with the same abbreviation logic $f_g^{\ell'}(\mathbf{p}_g) = D_{\tau_g} f^\ell(\mathbf{p}_g)$ and $f_s^{\ell''}(\mathbf{p}_s) = D_{\tau_s}^2 f^\ell(\mathbf{p}_s)$.

B Covariance function derivation developments

B.1 Definitions

Covariance function of Z between point $\mathbf{p}$ and point $\mathbf{q}$

$$(\mathbf{p}, \mathbf{q}) \mapsto \text{Cov}(Z(\mathbf{p}), Z(\mathbf{q})) \quad \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}. \quad (\text{B.21})$$

Stationarity case of order two assumption

$$\text{Cov}(Z(\mathbf{p}), Z(\mathbf{q})) = C_Z(\mathbf{p} - \mathbf{q}) = C_Z(\mathbf{h}_{pq}), \quad (\text{B.22})$$

with

$$\mathbf{h}_{pq} \mapsto C_Z(\mathbf{h}_{pq}) \quad \mathbb{R}^3 \rightarrow \mathbb{R}. \quad (\text{B.23})$$

Case of constant anisotropic distance (global anisotropy)

$$C_Z(\mathbf{h}_{pq}) = K(r(\mathbf{h}_{pq})), \quad (\text{B.24})$$

with

$$K : \mathbb{R} \rightarrow \mathbb{R} \quad \text{the chosen covariance model function}, \quad (\text{B.25})$$

$$r(\mathbf{h}_{pq}) = \|\mathbf{h}_{pq}\|_A = \sqrt{(\mathbf{p} - \mathbf{q})^T A (\mathbf{p} - \mathbf{q})} \quad \mathbb{R}^3 \rightarrow \mathbb{R} \quad \text{the anisotropic norm}, \quad (\text{B.26})$$

where A is a fixed rotation/dilatation matrix of $\mathbb{R}^3$.

In the following, $r(\mathbf{h}_{pq})$ may be denoted r for readability.

B.2 Notations

B.2.1 Differentiation with respect to variable $\mathbf{p}$ or $\mathbf{q}$

Although by assumption the covariance function of Z between points $\mathbf{p}$ and $\mathbf{q}$ only depends on the anisotropic distance r between them, in order to compute the partial derivatives needed in our model, we still have to consider it as a function of two variables $\mathbf{p}$ and $\mathbf{q}$ in $\mathbb{R}^3$. In the following, we will need to differentiate with respect to either of them, which we will denote with a superscript indicating the variable we differentiate upon.

For instance, for a function $F : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$, we note the partial derivative with respect to $\mathbf{p}$ along the vector $\boldsymbol{\tau}$ as

$$D_{\boldsymbol{\tau}}^p F(\mathbf{p}, \mathbf{q}) = \lim_{h \rightarrow 0} \frac{1}{h} (F(\mathbf{p} + h\boldsymbol{\tau}, \mathbf{q}) - F(\mathbf{p}, \mathbf{q})).$$

Denoting $(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$ the standard basis of $\mathbb{R}^3$, we can then define the gradient with respect to $\mathbf{p}$ as

$$\nabla^p F(\mathbf{p}, \mathbf{q}) = \begin{pmatrix} D_{\mathbf{e}_x}^p F(\mathbf{p}, \mathbf{q}) \\ D_{\mathbf{e}_y}^p F(\mathbf{p}, \mathbf{q}) \\ D_{\mathbf{e}_z}^p F(\mathbf{p}, \mathbf{q}) \end{pmatrix}.$$

By differentiating this with respect to $\mathbf{q}$, we can define $D^{p,q} F(\mathbf{p}, \mathbf{q})$ and by further differentiating with respect to one or the other variable, we can define higher-order differentials, like $D^{p,p,q} F(\mathbf{p}, \mathbf{q})$ or $D^{p,p,q,q} F(\mathbf{p}, \mathbf{q})$.

The dimensions of these quantities must be kept in mind in the following equations: $D_{\boldsymbol{\tau}}^p F(\mathbf{p}, \mathbf{q})$ is a scalar, $\nabla^p F(\mathbf{p}, \mathbf{q})$ is a 3D vector, $D^{p,q} F(\mathbf{p}, \mathbf{q})$ a 3 by 3 matrix, $D^{p,p,q} F(\mathbf{p}, \mathbf{q})$ is a 3x3x3 third-order tensor, and $\nabla^{p,p,q,q} F(\mathbf{p}, \mathbf{q})$ a 3x3x3x3 fourth-order tensor.

Successive derivations of the covariance model K , which is a scalar function, are classically denoted by K' for the first derivative, and by $K^{(n)}$ for the n -th derivative.

B.2.2 Antisymmetry property

In the particular case where F can be written as a function of $\mathbf{p} - \mathbf{q}$, i.e. there exists a function $G : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that $F(\mathbf{p}, \mathbf{q}) = G(\mathbf{p} - \mathbf{q})$, differentiating with respect to $\mathbf{p}$ is the same as differentiating with respect to $\mathbf{q}$ and changing the sign of the result.

In the present case, under our second-order stationary assumption, the covariance function of Z and the anisotropic distance satisfy this antisymmetry property, so we can substitute $\mathbf{p}$ for $\mathbf{q}$ in their differentials simply by changing the sign. The final sign of the substituted covariance function will then depend on the parity of the number of substitutions in the following way: $(-1)^n D^{p,p,\dots} C_Z(\mathbf{h}_{pq}) = D^{q,q,\dots} C_Z(\mathbf{h}_{pq})$ where n is the number of substitutions.

B.3 First derivative of covariance function

Directional derivative at point $\mathbf{p}$ according to $\mathbf{e}_x$

$$\begin{aligned}
 \text{Cov}(D_{\mathbf{e}_x} Z(\mathbf{p}), Z(\mathbf{q})) &= D_{\mathbf{e}_x}^p \text{Cov}(Z(\mathbf{p}), Z(\mathbf{q})) \\
 &= \text{Cov}\left(\lim_{\alpha \rightarrow 0} \frac{Z(\mathbf{p} + \alpha \mathbf{e}_x) - Z(\mathbf{p})}{\alpha}, Z(\mathbf{q})\right) \\
 &= \lim_{\alpha \rightarrow 0} \frac{1}{\alpha} [\text{Cov}(Z(\mathbf{p} + \alpha \mathbf{e}_x), Z(\mathbf{q})) - \text{Cov}(Z(\mathbf{p}), Z(\mathbf{q}))] \\
 &= \lim_{\alpha \rightarrow 0} \frac{1}{\alpha} [C_Z(\mathbf{h}_{pq} + \alpha \mathbf{e}_x) - C_Z(\mathbf{h}_{pq})] = D_{\mathbf{e}_x}^p C_Z(\mathbf{h}_{pq}) \\
 &= \lim_{\alpha \rightarrow 0} \frac{1}{\alpha} [K(r(\mathbf{h}_{pq} + \alpha \mathbf{e}_x)) - K(r(\mathbf{h}_{pq}))] = D_{\mathbf{e}_x}^p K(r(\mathbf{h}_{pq})).
 \end{aligned} \tag{B.27}$$

Thus, we obtain

$$D_{\mathbf{e}_x}^p \text{Cov}(Z(\mathbf{p}), Z(\mathbf{q})) = D_{\mathbf{e}_x}^p C_Z(\mathbf{h}_{pq}) = D_{\mathbf{e}_x}^p K(r(\mathbf{h}_{pq})). \tag{B.28}$$

The gradient of C_Z with respect to $\mathbf{p}$ is

$$\nabla^p C_Z(\mathbf{h}_{pq}) = \begin{pmatrix} D_{\mathbf{e}_x}^p C_Z(\mathbf{h}_{pq}) \\ D_{\mathbf{e}_y}^p C_Z(\mathbf{h}_{pq}) \\ D_{\mathbf{e}_z}^p C_Z(\mathbf{h}_{pq}) \end{pmatrix}. \tag{B.29}$$

First derivative specification of C_Z with respect to $\mathbf{p}$

$$\boxed{\nabla^p C_Z(\mathbf{h}_{pq}) = \nabla^p r(\mathbf{h}_{pq}) \cdot K'(r(\mathbf{h}_{pq}))}, \tag{B.30}$$

with

$$\nabla^p r(\mathbf{h}_{pq}) = \frac{A \mathbf{h}_{pq}}{r(\mathbf{h}_{pq})}. \tag{B.31}$$

Let us notice again that

$$\nabla^p r(\mathbf{h}_{pq}) = -\nabla^q r(\mathbf{h}_{pq}). \tag{B.32}$$

Equation B.30 leads to 2 conditions

- K should be differentiable,
- For $\nabla^p C_Z(\mathbf{h}_{pq})$ to be defined where $\mathbf{p} = \mathbf{q}$, we need $K'(0) = 0$.

B.3.1 Generalization: first-order derivative of the covariance function in a specific direction

$$\begin{aligned}
 \text{Cov}(D_{\tau} Z(\mathbf{p}), Z(\mathbf{q})) &= \text{Cov}(\tau_p \cdot \nabla Z(\mathbf{p}), Z(\mathbf{q})) \\
 &= \tau_p \cdot \nabla^p \text{Cov}(Z(\mathbf{p}), Z(\mathbf{q})) \\
 &= \tau_p \cdot \nabla^p C_Z(\mathbf{h}_{pq}).
 \end{aligned} \tag{B.33}$$

B.4 Second-order derivative of the covariance function

Covariance of partial derivatives w.r.t. $\mathbf{p}$ according to $\mathbf{e}_x$ and w.r.t. $\mathbf{q}$ according to $\mathbf{e}_y$

$$\text{Cov}(D_{\mathbf{e}_x} Z(\mathbf{p}), D_{\mathbf{e}_y} Z(\mathbf{q})) = {}^t \mathbf{e}_x . D^{p,q} C_Z(\mathbf{h}_{pq}) . \mathbf{e}_y. \quad (\text{B.34})$$

Second-order derivative specification of C_Z with respect to $\mathbf{p}$ then $\mathbf{q}$

$$D^{p,q} C_Z(\mathbf{h}_{pq}) = K^{(2)}(r(\mathbf{h}_{pq})) \nabla^q r(\mathbf{h}_{pq}) . {}^t \nabla^p r(\mathbf{h}_{pq}) + K'(r(\mathbf{h}_{pq})) D^{p,q} r(\mathbf{h}_{pq}), \quad (\text{B.35})$$

with

$$D^{p,q} r(\mathbf{h}_{pq}) = -\frac{A}{r(\mathbf{h}_{pq})} + \frac{A \mathbf{h}_{pq}}{r^3(\mathbf{h}_{pq})} {}^t (A \mathbf{h}_{pq}) \in \mathcal{M}_3(\mathbb{R}). \quad (\text{B.36})$$

Let us notice again that

$$D^{p,q} C_Z(\mathbf{h}_{pq}) = -D^{q,q} C_Z(\mathbf{h}_{pq}) = -D^{p,p} C_Z(\mathbf{h}_{pq}). \quad (\text{B.37})$$

Equation B.35 leads to 2 conditions

- K should be twice differentiable,
- For $D^{p,q} C_Z(\mathbf{h}_{pq})$ to be defined at $\mathbf{h}_{pq} = 0$, we need $K'(0) = 0$.

B.4.1 Second-order derivative of the covariance function in a specific direction

$$\begin{aligned} \text{Cov}(D_{\tau}^2 Z(\mathbf{p}), Z(\mathbf{q})) &= \text{Cov}({}^t \tau . D^2 Z(\mathbf{p}) . \tau, Z(\mathbf{q})) \quad \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R} \\ &= {}^t \tau . D^{p,p} C_Z(\mathbf{h}_{pq}) . \tau \\ &= \sum_i \sum_j \tau[i] \tau[j] D^{p,p} C_Z(\mathbf{h}_{pq})[i, j]. \end{aligned} \quad (\text{B.38})$$

B.5 Third-order derivative of the covariance function

$$\begin{aligned} \text{Cov}(D_{\tau}^2 Z(\mathbf{p}), D_{\mu} Z(\mathbf{q})) &= \text{Cov}({}^t \tau . D^2 Z(\mathbf{p}) . \tau, {}^t \mu . \nabla Z(\mathbf{q})) \\ &= \sum_i \sum_j \sum_k \tau[i] \tau[j] \mu[k] D^{p,p,q} C_Z(\mathbf{h}_{pq})[i, j, k]. \end{aligned} \quad (\text{B.39})$$

Reminding that

$$D^{p,p,q} C_Z(\mathbf{h}_{pq}) = -D^{p,p,p} C_Z(\mathbf{h}_{pq}), \quad (\text{B.40})$$

where $D^{p,p,p} C_Z(\mathbf{h}_{pq})$ is a third-order tensor (3x3x3) given by

$$\begin{aligned} D^{p,p,p} C_Z(\mathbf{h}_{pq})[i, j, k] &= K'(r) D^{p,p,p} r[i, j, k] \\ &\quad + K^{(2)}(r) (\nabla^p r[i] D^{p,p}[j, k] + \nabla^p r[j] D^{p,p}[i, k] + \nabla^p r[k] D^{p,p}[i, j]) \\ &\quad + K^{(3)}(r) (\nabla^p r[i] \nabla^p r[j] \nabla^p r[k]). \end{aligned} \quad (\text{B.41})$$

The third-order derivative of r is also a third-order tensor (3x3x3) given by

$$\begin{aligned} D^{p,p,p} r(\mathbf{h}_{pq})[i, j, k] &= -\frac{1}{r^3} [A_{i,j} \mathbf{t}_k + A_{i,k} \mathbf{t}_j + A_{j,k} \mathbf{t}_i] \\ &\quad + \frac{3}{r^5} [\mathbf{t}_i \mathbf{t}_j \mathbf{t}_k], \end{aligned} \quad (\text{B.42})$$

where $\mathbf{t} = A \mathbf{h}_{p,q}$

B.6 Fourth-order derivative of the covariance function

$$\begin{aligned} \text{Cov}(D_{\tau}^2 Z(\mathbf{p}), D_{\mu}^2 Z(\mathbf{q})) &= \text{Cov}({}^t \tau . D^2 Z(\mathbf{p}). \tau, {}^t \mu . D^2 Z(\mathbf{q}). \mu) \\ &= \sum_i \sum_j \sum_k \sum_l \tau[i] \tau[j] \mu[k] \mu[l] D^{p,p,q,q} C_Z(\mathbf{h}_{pq})[i, j, k, l]. \end{aligned} \quad (\text{B.43})$$

Reminding that

$$D^{p,p,q,q} C_Z(\mathbf{h}_{pq}) = D^{p,p,p,p} C_Z(\mathbf{h}_{pq}), \quad (\text{B.44})$$

$D^{p,p,p,p} C_Z(\mathbf{h}_{pq})$ is a 3x3x3x3 fourth-order tensor

$$D^{p,p,p,p} C_Z(\mathbf{h}_{pq}) = K'(r) M_1 + K^{(2)}(r) M_2 + K^{(3)}(r) M_3 + K^{(4)}(r) M_4, \quad (\text{B.45})$$

with

$$\begin{aligned} M_1 &= D^{p,p,p,p} r \\ M_2[i, j, k, l] &= D^{p,p} r[i, j] D^{p,p} r[k, l] + D^{p,p} r[i, k] D^{p,p} r[j, l] + D^{p,p} r[i, l] D^{p,p} r[j, k] \\ &\quad + \nabla^p r[i] D^{p,p,p} r[j, k, l] + \nabla^p r[j] D^{p,p,p} r[i, k, l] + \nabla^p r[k] D^{p,p,p} r[i, j, l] + \nabla^p r[l] D^{p,p,p} r[i, j, k] \\ M_3[i, j, k, l] &= D^{p,p} r[i, j] \nabla^p r[k] \nabla^p r[l] + D^{p,p} r[i, k] \nabla^p r[j] \nabla^p r[l] + D^{p,p} r[i, l] \nabla^p r[j] \nabla^p r[k] \\ &\quad + D^{p,p} r[j, k] \nabla^p r[i] \nabla^p r[l] + D^{p,p} r[j, l] \nabla^p r[i] \nabla^p r[k] + D^{p,p} r[k, l] \nabla^p r[i] \nabla^p r[j] \\ M_4[i, j, k, l] &= \nabla^p r[i] \nabla^p r[j] \nabla^p r[k] \nabla^p r[l] \end{aligned} \quad (\text{B.46})$$

and

$$\begin{aligned} D^{p,p,p,p} r(\mathbf{h}_{pq})[i, j, k, l] &= -\frac{1}{r^3} (A_{ij} A_{kl} + A_{ik} A_{jl} + A_{il} A_{jk}) \\ &\quad + \frac{3}{r^5} (A_{ij} A_{kl} \mathbf{t}_k \mathbf{t}_l + A_{ik} A_{jl} \mathbf{t}_j \mathbf{t}_l + \dots + A_{kl} A_{ij} \mathbf{t}_i \mathbf{t}_j) \\ &\quad - \frac{15}{r^7} (\mathbf{t}_i \mathbf{t}_j \mathbf{t}_k \mathbf{t}_l), \end{aligned} \quad (\text{B.47})$$

where $\mathbf{t} = A \mathbf{h}_{p,q}$.

C Demonstration of null second-order derivative along the hinge line

If we consider a twice-differentiable parametric curve $\gamma : t \mapsto \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \in \mathbb{R}^3$ and define $g : t \mapsto Z(\gamma(t))$,

where Z is our potential field ($Z : \mathbb{R}^3 \rightarrow \mathbb{R}$), we have in all generality

$$g'(t) = \nabla Z(\gamma(t)) \cdot \overrightarrow{\gamma'(t)} \quad (\text{C.48})$$

and

$$g''(t) = \nabla Z(\gamma(t)) \cdot \overrightarrow{\gamma''(t)} + {}^t \gamma''(t) \cdot D^2 Z(\gamma(t)) \cdot \overrightarrow{\gamma'(t)}. \quad (\text{C.49})$$

Let us assume that γ is a parametrization of a planar curve included in an isopotential of Z , and that furthermore, there exists a plane that is tangent to this isopotential along the curve defined by γ . This is the case, for instance, for any parametrization of the H_3 curve of Fig. 1.

Under this assumption, g is constant, therefore in particular $g'' = 0$. Since $\overrightarrow{\gamma''(t)}$ is in the same plane as γ (plane P_t in the case of Fig. 1), that is tangent to the isopotential, it is orthogonal to the gradient $\nabla Z(\gamma(t))$.

Equation C.49 then reduces to ${}^t \gamma'(t) \nabla^2 Z(\gamma(t)) \overrightarrow{\gamma'(t)} = 0$, i.e. $\overrightarrow{\gamma'(t)} \in \text{Ker}(D^2 Z(\gamma(t)))$. In other words, the directional second derivative of Z is null along $\overrightarrow{\gamma'(t)}$, which is the direction of the crest/hinge line H_3 in the particular case of Fig. 1.

D Volumic error indicator

D.1 Volumic error (V_e) computation

The indicator V_e provides a quantitative measure of the ability of the model to reconstruct the structure of interest of our synthetic example. It can be thought of as the volume between the two surfaces near the fold.

Let us denote by $z(x, y)$ and $\hat{z}(x, y)$ the respective elevations of the theoretical surface and the modeled surface at location (x, y) . The isopotential considered in this work, described in Sec. 3 has a constant elevation of $z_{floor} = 0.6$ on a large part of the domain, but no information is provided to any model in this zone. In order not to penalize errors made far from the region of interest, instead of considering $\hat{z}$, we consider $\min(\hat{z}, z_{floor})$ when calculating the volume

$$V_e = \|z - \min(\hat{z}, z_{floor})\|_1 = \int |z(x, y) - \min(\hat{z}(x, y), z_{floor})|$$

D.2 Uncertainty (ΔV) on this computation

In practice, the interpolated value of the potential $Z^*(x, y, z)$ is known only on the grid points (x_i, y_j, z_k) ; at each location (x_i, y_j) the elevation $\hat{z}(x_i, y_j)$ is estimated by taking z_k such that $Z^*(x_i, y_j, z_k)$ is closest to 2.5, which is the value of the iso-surface considered here. This leads to an estimation uncertainty ΔV that depends on the vertical resolution (the finer the grid, the lower this uncertainty). For each model, this uncertainty is computed along with the volumic error indicator (it is almost constant but may vary a little, because this uncertainty is zero for locations where $\hat{z} \leq z_{floor}$, since we use the value z_{floor} anyway). Several models are close in performance, in terms of this indicator; this uncertainty needs to be taken into account when comparing them.

D.3 Relative volumic improvement (RVI, ηV_e)

$S1$: Scenario 1 ; $S2$: Scenario 2

V_e : Volumic error (u^3)

V_{e1} : Volumic error of initial result of Scenario 1 (u^3)

V_{e2} : Volumic error of initial result of Scenario 2 (u^3)

where u is the arbitrary length unit used here.

ηV_e (RVI) : Signed relative difference of V_e compared to initial result in the same scenario (%)

$$\eta V_e = \frac{V_{ei} - V_e}{V_{ei}} \quad i \in (1, 2)$$

D.4 Results

This part summarizes all the intermediate computations for the results concerning the volumic error of the different investigations of Sec. 4.

References

- Boisvert, J., Manchuk, J., and Deutsch, C. (2009). Kriging in the presence of locally varying anisotropy using non-euclidean distances. *Math. Geosci.*, 41(5):585–601.
- Calcagno, P., Chilès, J.-P., Courrioux, G., and Guillen, A. (2008). Geological modelling from field data and geological knowledge: Part i. modelling method coupling 3d potential-field interpolation and geological rules. *Phys. Earth Planet. Inter.*, 171(1-4):147–157.
- Caumon, G., Gray, G., Antoine, C., and Titeux, M.-O. (2012). Three-dimensional implicit stratigraphic model building from remote sensing data on tetrahedral meshes: Theory and application to a regional model of la popa basin, ne mexico. *IEEE Trans Geosci Remote Sens*, 51(3):1613–1621.
- Chiles, J.-P. and Delfiner, P. (2009). *Geostatistics: modeling spatial uncertainty*, volume 497. John Wiley & Sons.
- Cowan, E., Beatson, R., Ross, H., Fright, W., McLennan, T., Evans, T., Carr, J., Lane, R., Bright, D., Gillman, A., et al. (2003). Practical implicit geological modelling. In *Fifth international mining geology conference*, pages 17–19. Australian Institute of Mining and Metallurgy Bendigo, Victoria.
- de la Varga, M. and Wellmann, J. F. (2016). Structural geologic modeling as an inference problem: A bayesian perspective. *Interpretation*, 4(3):SM1–SM16.

Initial V_e in S1 (u^3)	Initial V_e in S2 (u^3)
104.68 ± 4.12	48.36 ± 3.40

Initial values

	V_e in S1 (u^3)	V_e in S2 (u^3)	ηV_e to S1 (%)	ηV_e to S2 (%)
Tangent	85.19 ± 4.27	50.72 ± 3.40	18.62	−4.89
D2	91.52 ± 4.17	49.76 ± 3.40	12.58	−2.91
Tangent + D2	91.39 ± 4.35	49.01 ± 3.40	12.70	−1.35

First experiment (Fig. 5)

	V_e in S1 (u^3)	V_e in S2 (u^3)	ηV_e to S1 (%)	ηV_e to S2 (%)
29	85.26 ± 4.27	50.69 ± 3.40	18.56	−4.83
8	85.06 ± 4.25	50.50 ± 3.40	18.74	−4.43
4	85.03 ± 4.12	53.32 ± 3.42	18.78	−10.26

Second experiment (Fig. 6)

	V_e in S1 (u^3)	V_e in S2 (u^3)	ηV_e to S1 (%)	ηV_e to S2 (%)
Sparse	92.73 ± 4.25	51.14 ± 3.40	11.42	−5.76
Data	89.42 ± 3.95	49.72 ± 3.40	14.58	−2.82
Vertex	119.72 ± 4.25	49.35 ± 3.40	−14.37	−2.02

Third experiment (Fig. 7)

	V_e in S1 (u^3)	V_e in S2 (u^3)	ηV_e to S1 (%)	ηV_e to S2 (%)
1	85.26 ± 4.27	50.69 ± 3.40	18.56	−4.83
2	78.81 ± 4.12	49.51 ± 3.40	24.72	−2.39
3	78.52 ± 4.11	49.51 ± 3.40	24.99	−2.39
4	33.34 ± 3.59	31.00 ± 3.56	68.15	35.90

Fourth experiment (Fig. 8)

- Frank, T., Tertois, A.-L., and Mallet, J.-L. (2007). 3d-reconstruction of complex geological interfaces from irregularly distributed and noisy point data. *Comput and Geosci*, 33(7):932–943.
- Frodeman, R. (1995). Geological reasoning: Geology as an interpretive and historical science. *Geol. Soc. Am. Bull.*, 107(8):960–968.
- Garcia, M., Cani, M.-P., Ronfard, R., Gout, C., and Perrenoud, C. (2018). Automatic generation of geological stories from a single sketch. In *Proceedings of the Joint Symposium on Computational Aesthetics and Sketch-Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering*, pages 1–15.
- Grose, L., Laurent, G., Aillères, L., Armit, R., Jessell, M., and Caumon, G. (2017). Structural data constraints for implicit modeling of folds. *J. Struct. Geol.*, 104:80–92.
- Henrion, V., Caumon, G., and Cherpeau, N. (2010). Odsim: an object-distance simulation method for conditioning complex natural structures. *Math. Geosci.*, 42(8):911–924.
- Hillier, M. J., Schetselaar, E. M., de Kemp, E. A., and Perron, G. (2014). Three-dimensional modelling of geological surfaces using generalized interpolation with radial basis functions. *Math. Geosci.*, 46(8):931–953.
- Houlding, S. W. (1994). 3d geoscience modeling: computer techniques for geological characterization. Springer Verlag, 46(3):85–90.
- Hudleston, P. J. and Treagus, S. H. (2010). Information from folds: a review. *J. Struct. Geol.*, 32(12):2042–2071.
- Lajaunie, C., Courrioux, G., and Manuel, L. (1997). Foliation fields and 3d cartography in geology: principles of a method based on potential interpolation. *Math. Geol.*, 29(4):571–584.
- Laurent, G., Aillères, L., Grose, L., Caumon, G., Jessell, M., and Armit, R. (2016). Implicit modeling of folds and overprinting deformation. *Earth Planet Sci Lett*, 456:26–38.
- Mallet, J.-L. (1992). Discrete smooth interpolation in geometric modelling. *Comput. Aided Des.*, 24(4):178–191.
- Manchuk, J. G. and Deutsch, C. V. (2019). Boundary modeling with moving least squares. *Comput and Geosci*, 126:96–106.
- Maxelon, M., Renard, P., Courrioux, G., Brändli, M., and Mancktelow, N. (2009). A workflow to facilitate three-dimensional geometrical modelling of complex poly-deformed geological units. *Comput and Geosci*, 35(3):644–658.
- McClay, K. R. (2013). *The mapping of geological structures*. John Wiley & Sons.
- Ramsay, J. G. and Huber, M. I. (1987). *The techniques of modern structural geology: Folds and fractures*, volume 2. Academic press.
- Renaudeau, J., Malvesin, E., Maerten, F., and Caumon, G. (2019). Implicit structural modeling by minimization of the bending energy with moving least squares functions. *Math. Geosci.*, 51(6):693–724.
- Turk, G. and O’Brien, J. F. (2005). Modelling with implicit surfaces that interpolate. In *ACM SIGGRAPH 2005 Courses*, page 21. ACM.
- Wackernagel, H. (2013). *Multivariate geostatistics: an introduction with applications*. Springer Science & Business Media.
- Wellmann, F. and Caumon, G. (2018). 3-d structural geological models: Concepts, methods, and uncertainties. *Adv. Geophys.*, 59:1–121.
- Wendland, H. (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Adv Comput Math*, 4(1):389–396.

Figures

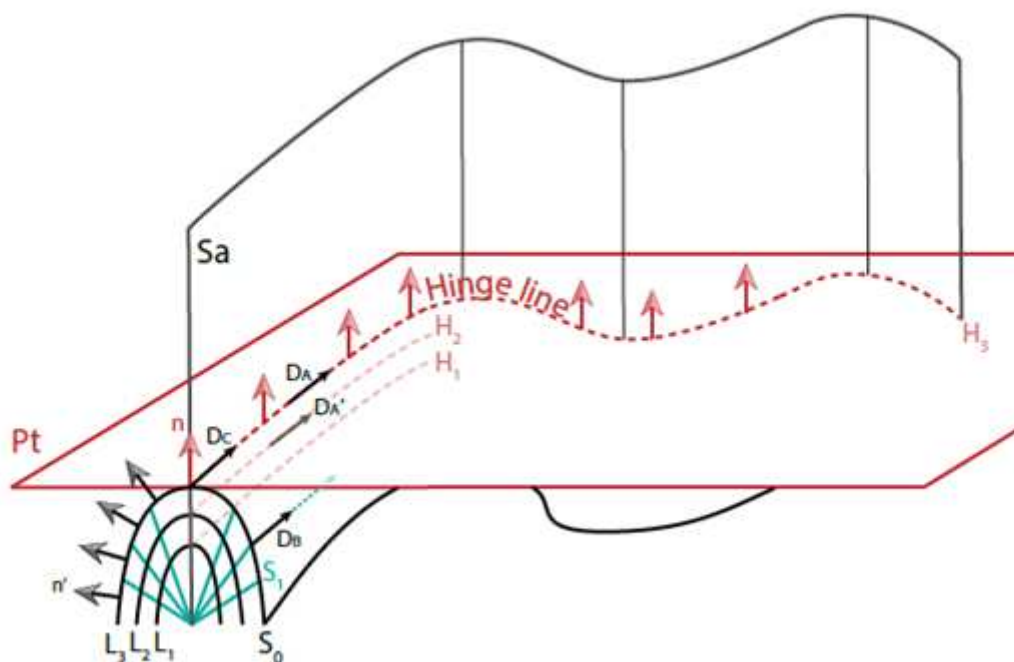


Figure 1

Fold geometry and planes of interest. Sa: Axial surface. n: Pole vectors belonging to the axial plane; n0: Other pole vectors of the same profile. Pt : Tangent plane. S0: Bedding. S1: Cleavage. H1, H2, H3: Hinge lines of beddings L1, L2, L3. The hinge line H3 and the crest line coincide in this case. DA and DA0 : Fold axis measurements on hinge lines H2 and H3. DB: Intersection lineation between S0 and S1 measurement. DC: Fold axis inference from stereographic projections using n0 pole measurements.

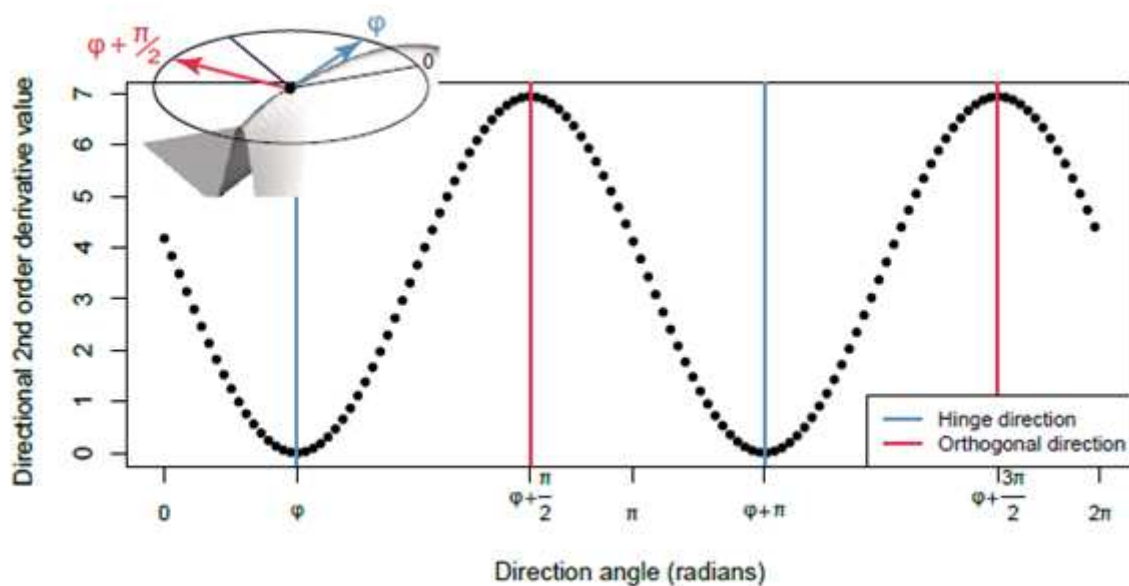


Figure 2

Please see the Manuscript PDF file for the complete figure caption.

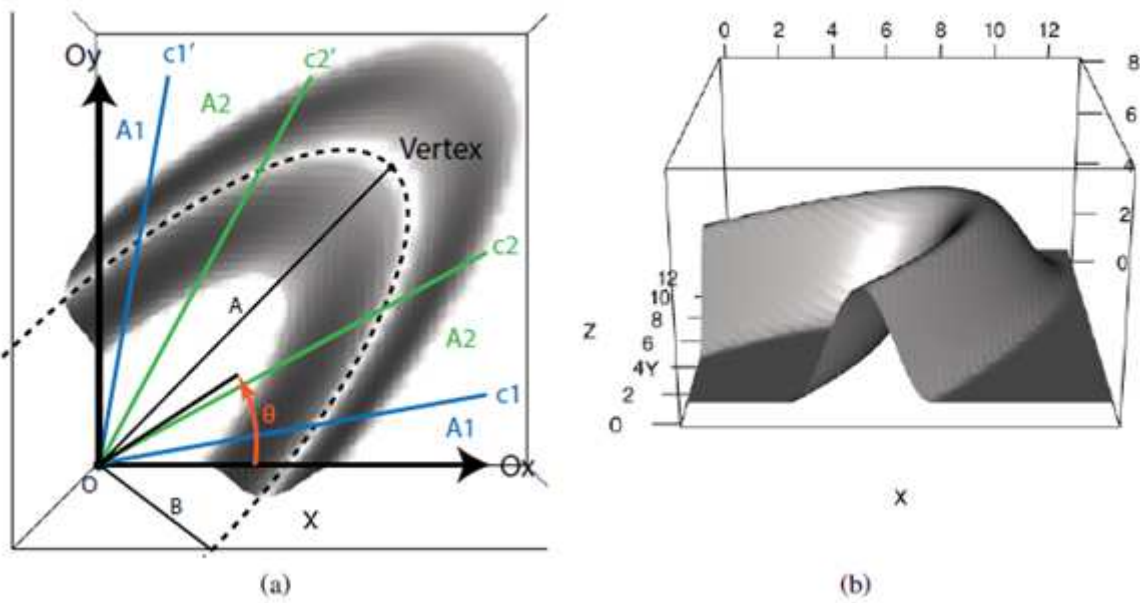


Figure 3

Please see the Manuscript PDF file for the complete figure caption.

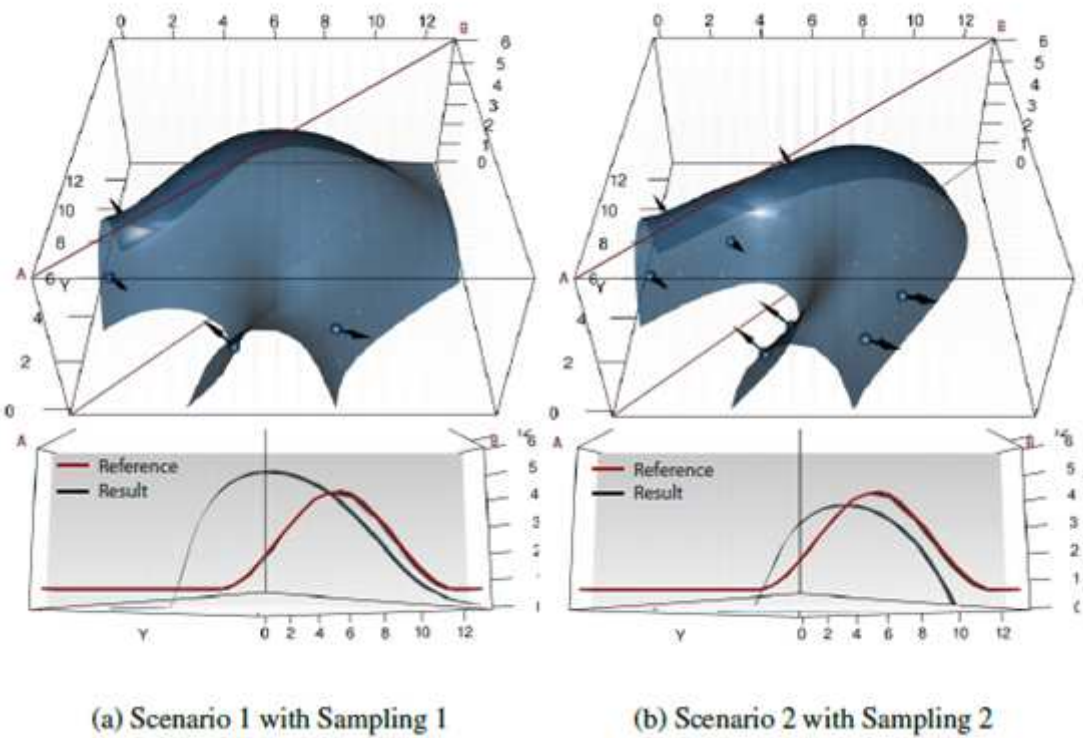


Figure 4

Initial results of the potential field method on two different initial samplings. Sampling 1 consists in 4 contact points (blue spheres) and 4 gradient data (black arrows). Sampling 2 consists in 8 contact points and 8 gradient data.

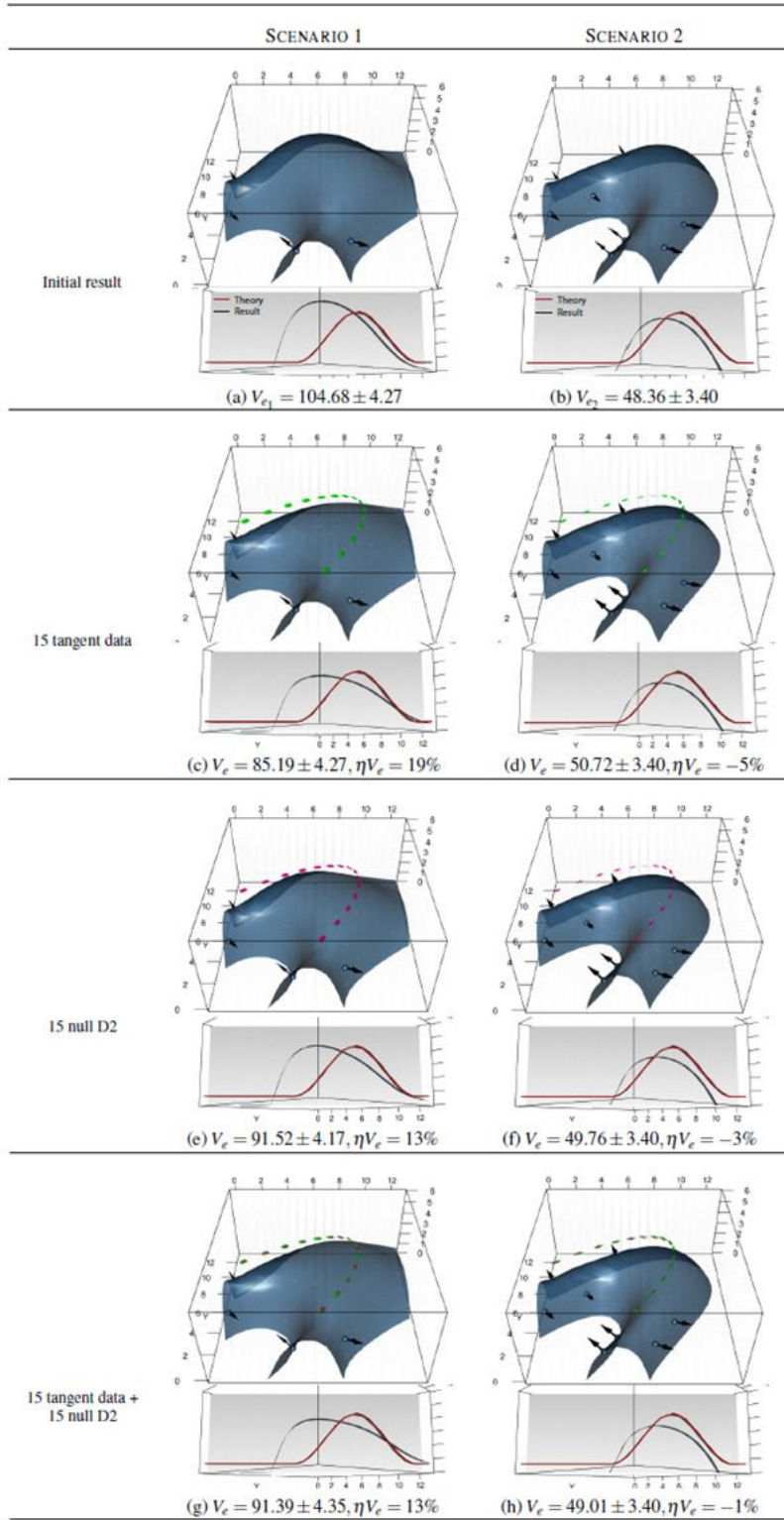


Figure 5

Impact of (c)(d) 15 tangent data (green arrows), (e)(f) 15 null second-order derivative data (D2, pink arrows), (g)(h) both data on the resulting surface compared to (a)(b) the initial result with Sampling 1 or

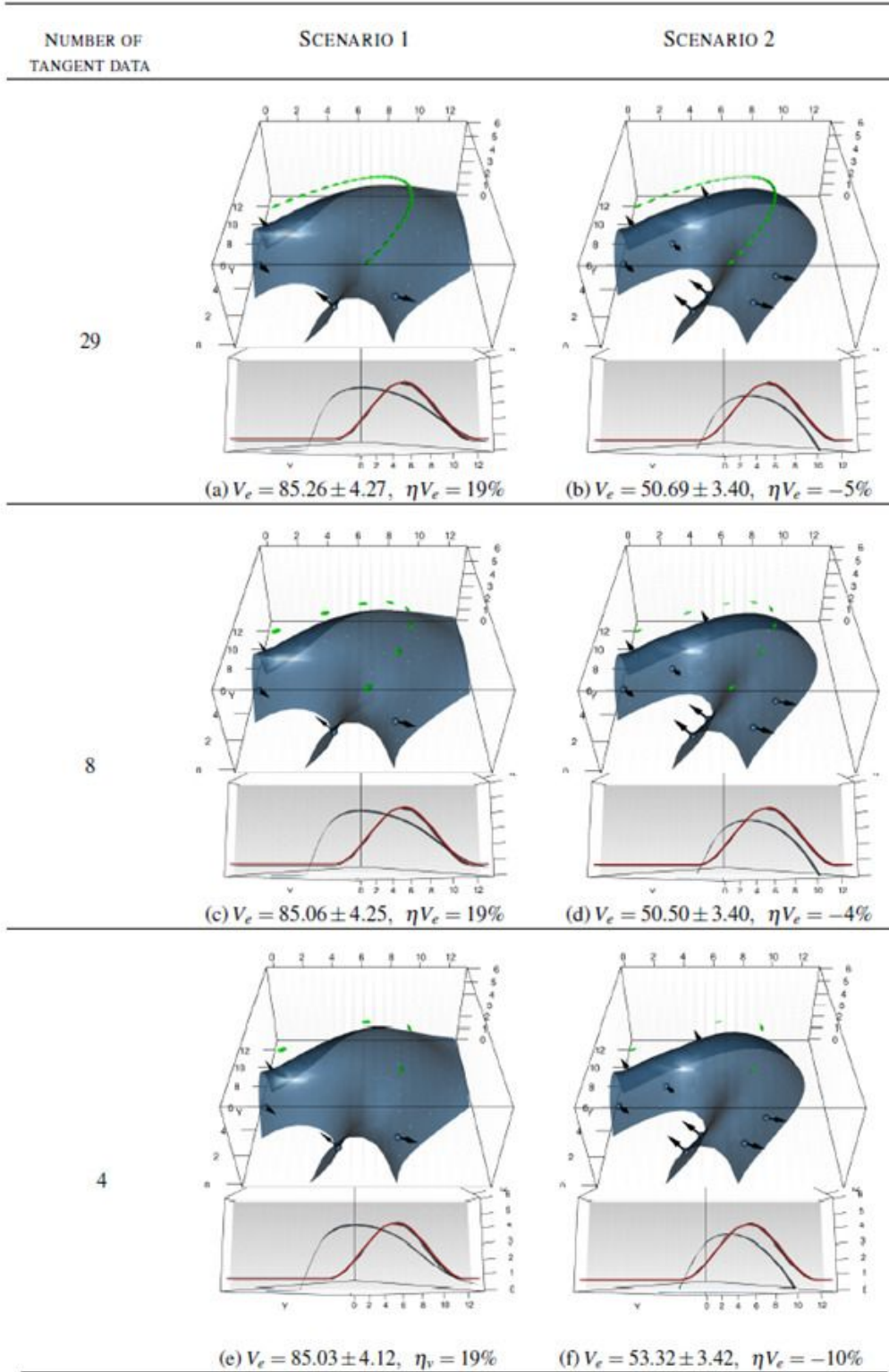


Figure 6

Impact of the number of tangent data along the hinge line in the two different scenarios

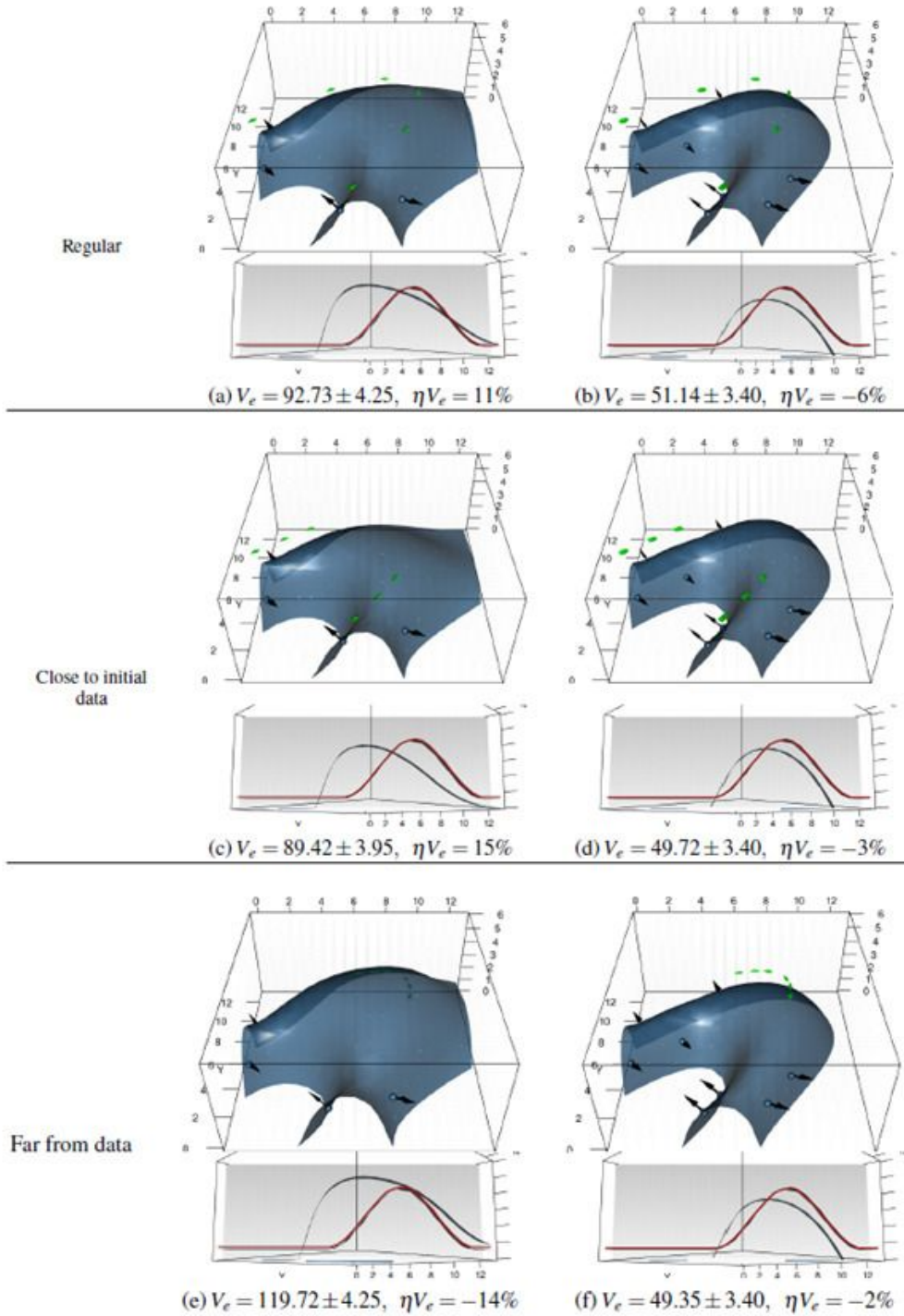
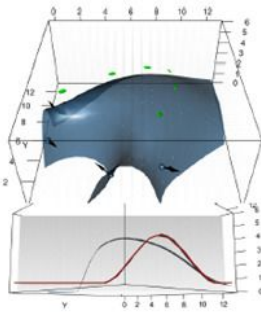


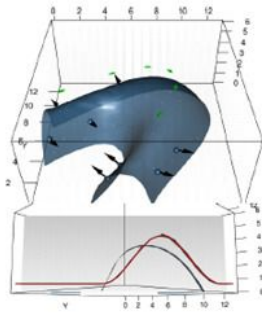
Figure 7

Impact of the spatial distribution of the tangent data along the hinge line in the two different scenarios. 6 tangent data are poised in different ways: (a)(b) regularly spaced along the hinge line, (c)(d) close to the initial data, (e)(f) far from the initial data

1

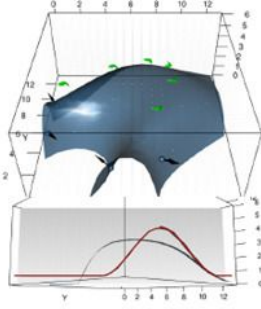


(a) $V_e = 85.26 \pm 4.27$, $\eta V_e = 19\%$

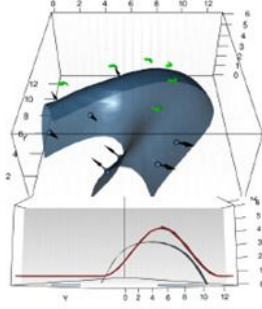


(b) $V_e = 50.69 \pm 3.40$, $\eta V_e = -5\%$

2

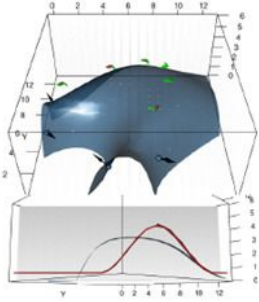


(c) $V_e = 78.81 \pm 4.12$, $\eta V_e = 25\%$

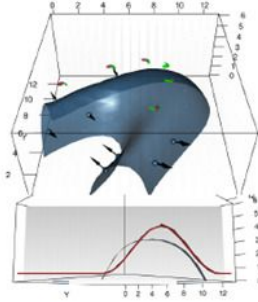


(d) $V_e = 49.51 \pm 3.40$, $\eta V_e = -2\%$

3

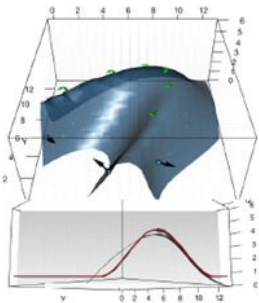


(e) $V_e = 78.52 \pm 4.11$, $\eta V_e = 25\%$

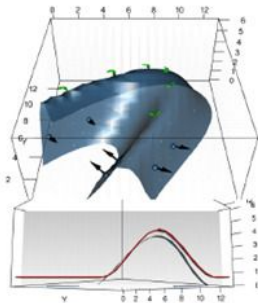


(f) $V_e = 49.51 \pm 3.40$, $\eta V_e = -2\%$

4



(g) $V_e = 33.34 \pm 3.59$, $\eta V_e = 68\%$



(h) $V_e = 31.00 \pm 3.56$, $\eta V_e = 36\%$

Figure 8

Comparison of the impact of (first row) 1, (second row) 2, (third row) 3 and (last row) 4 constraints per point at the same 6 locations on the fold axis of the theoretical surface