# Bhutan Landscape Anomaly: Possible Effect on Himalayan Economy (In View of Optimal Description of Elevation Profiles)

Thach Ngoc Nguyen[1], Laxman Bokati[2],
Aaron Velasco[3], and Vladik Kreinovich[2]
[1]Banking University of Ho Chi Minh City
56 Hoang Dieu 2, Quan Thu Duc, Thu Duc
Ho Chi Minh City, Vietnam, Thachnn@buh.edu.vn
[2]Computational Science Program
[3]Department of Geological Sciences
University of Texas at El Paso, El Paso, Texas 79968, USA,
lbokati@miners.utep.edu, aavelasco@utep.edu, vladik@utep.edu

**Abstract**

Economies of countries located in seismic zones are strongly effected by this seismicity. If we underestimate the seismic activity, then a reasonably routine earthquake can severely damage the existing structures and thus, lead to huge economic losses. On the other hand, if we overestimate the seismic activity, we waste a lot of resources on unnecessarily fortifying all the buildings – and this too harms the economies. From this viewpoint, it is desirable to have estimations of regional seismic activities which are as accurate as possible. Current predictions are mostly based on the standard geophysical understanding of earthquakes as being largely caused by the movement of tectonic plates and terranes. This understanding works in most areas, but in Bhutan area of the Himalayas region, there seems to be a landscape anomaly. As a result, for this region, we have less confidence in the accuracy of seismic predictions based on the standard understanding and thus, have to use higher seismic thresholds in construction. In this paper, we find the optimal description of landscape-describing elevation profiles, and we use this description to show that the seeming anomaly is actually in perfect agreement with the standard understanding of the seismic activity. Our conclusion is that it is safe to apply, in this region, estimates based on the standard understanding and thus, avoid unnecessary expenses caused by an increased threshold.

# 1 Formulation of the Problem

**Seismicity affects economy.** In highly seismic areas like the Himalayas, economy is affected by our knowledge of possible seismicity.

Protection against possible earthquakes is very costly. If we have only a vague idea about possible seismic events – i.e., if we can potentially expect high-energy earthquakes at all possible locations – then, every time we build a house or a factory, we need to spend a lot of money on making it protected against such events – with little money left for any other development project.

On the other hand, if we can reasonably accurately localize potential hazards, then we can concentrate our building efforts mostly in safer zones. This will require less investment in earthquake protection and thus, leave more money for other development projects.

Thus, the economy of a highly seismic zone is directly affected by our understanding of the corresponding seismic processes.

**Bhutan landscape anomaly.** In general, modern geophysics has a reasonably good understanding of seismic processes and seismic zones. Specifically, the current understanding is that seismicity is usually caused by mutual movement of tectonic plates and their parts (terranes), and it is mostly concentrated on the borderline between two or more such plates or terranes. In general, while we still cannot predict the exact timing of earthquakes, geoscientists can reasonably well predict the size of a future earthquake based on the corresponding geophysical models.

Researchers and practitioners are reasonably confident in these predictions – at least for locations whose geophysics is well understood by the traditional geophysical models.

However, there are locations where observed phenomena are different from what we usually expect. In such cases, there are reasonable doubts in seismicity estimates produced by the traditional techniques – and thus, it is reasonable to be cautious and use higher strengths of potential earthquakes when building in these locations, which invokes significant additional expenses. For such domains, it is therefore desirable to come up with a better understanding of the observed geophysical phenomena – thus hopefully allowing us to make more accurate predictions and hence, save money (which is now wasted on possibly too-heavy earthquake protection) for other important activities.

One such areas in the vicinity of the Himalayan country of Bhutan, where the landscape profile is drastically different from the profiles of other Himalayan areas such as areas of Nepal. In general, a landscape can be described in numerical terms if we take a line orthogonal to the prevailing rivers (which are usually the lowest points on the landscape) and plot the elevation as a function of the distance from the corresponding river. The shape of the landscape (elevation) profile in Bhutan is visually drastically different from the landscape profile in Nepal; see, e.g., [1]. Namely, in most of the Himalayas – and, in general, in the most of the world – the corresponding curve is first convex (corresponding to the river valley), and then becomes concave – which corresponds to the mountain peaks. In contrast, in Bhutan, the profile turns concave very fast, way before we reach the mountain peaks area.

As of now, there are no good well-accepted explanations for this phenomenon – which makes it an anomaly. To be more precise, we know that the geophysics of the Bhutan area is somewhat different: in Nepal (like in most areas in the world), the advancing tectonic plate in orthogonal to the border of the mountain range, while in Bhutan, the plate pushes the range at an angle. However, it is not clear how this can explain the above phenomenon. This leads us to the following questions.

**Questions.** The first question is: can we explain the Bhutan anomaly within the existing geophysical paradigm? If we can do, this would mean that this anomaly is not an obstacle to applying this paradigm, and thus, that the estimates of future seismic activity obtained within this paradigm can be safely applied – without the need to make expensive extra precautions.

A related question is related to the fact that while we use convexity and concavity to describe elevation profiles, the only reason for using these two properties is because these are the basic properties that we learn in math. Is there any geophysical meaning in convexity vs. concavity?

**What we do in this paper.** In this paper, we provide answers to both questions: we explain why convexity and concavity are adequate ways to describe elevation profiles, and we explain how the at-an-angle pressure in the Bhutan area leads to the observed convex-followed-by-concave phenomenon.

To answer these questions, we first formulate the problem of adequately describing elevation profiles as an optimization problem. Then, we solve this problem, and use the solution to answer the above two questions.

## 2 What Is the Optimal Description of Elevation Profiles: Precise Formulation of the Optimization Problem and the Main Result

**How can we describe elevation profiles?** An elevation profile results from the joint effect of many different physical processes, from movement of tectonic plates to erosion. These process are largely independent from each other: e.g., erosion works the same way whether we have the landscape on the sea level or the same landscape which the geological processes raised to some elevation. Because of this independence, the observed profile $f(x)$ can be reasonably well represented as the sum of profiles corresponding to different processes:

$$f(x) = f_1(x) + \ldots + f_n(x).$$

Different profile-changing processes may have different intensity. So, to describe the effect of the $i$-th process, instead of a fixed function $f_i(x)$, it is more appropriate to use the correspondingly re-scaled term $C_i \cdot f_i(x)$, where the coefficients $C_i$ describe the intensity of the $i$-th process, so that

$$f(x) = C_1 \cdot f_1(x) + \ldots + C_n \cdot f_n(x).$$

Due to erosion, discontinuities in the elevation profiles are usually smoothed out, so we can safely assume that the corresponding functions $f_i(x)$ are smooth (differentiable).

Thus, we arrive at the following definition.

**Definition 1.** *Let $n$ be a positive integer. By a* description of elevation profiles*, we mean a family of functions*

$$\{C_1 \cdot f_1(x) + \ldots + C_n \cdot f_n(x)\}_{C_1,\ldots,C_n}, \tag{1}$$

*where the functions $f_1(x), \ldots, f_n(x)$ are fixed differentiable functions, and $C_1$, $\ldots$, $C_n$ are arbitrary real numbers.*

From this viewpoint, selecting a description means selecting $n$ functions $f_1(x)$, $\ldots$, $f_n(x)$.

**Towards the optimal description.** Which description is the best? To answer this question, we need to be able to decide, for each two families of functions $F$ and $F'$, whether the first family is better (we will denote it by $F' < F$) or the second family is better ($F < F'$), or both families have the same quality (we will denote it by $F \sim F'$). Clearly, if $F$ is worse than $F'$ and $F'$ is worse than $F''$, then $F$ should be worse than $F''$. So, we arrive at the following definition.

**Definition 2.** *Let $n$ be a positive number. By an* optimality criterion*, we mean the pair of relations $(<, \sim)$ on the set $S$ of all possible $n$-dimensional descriptions of elevation profiles that satisfies the following conditions:*

- *for every pair $F, F' \in S$, we have one and only one of the following options: either $F' < F$ or $F < F'$ or $F \sim F'$;*
- *for every $F$, $F'$, and $F''$, if $F < F'$ and $F' < F''$, then $F < F''$;*
- *for every $F$, $F'$, and $F''$, if $F < F'$ and $F' \sim F''$, then $F < F''$;*
- *for every $F$, $F'$, and $F''$, if $F \sim F'$ and $F' < F''$, then $F < F''$;*
- *for every $F$, $F'$, and $F''$, if $F \sim F'$ and $F' \sim F''$, then $F \sim F''$;*
- *for every $F$ and $F'$, if $F \sim F'$, then $F' \sim F$.*

**Definition 3.** *Let $(<, \sim)$ be an optimality criterion. We say that a family $F$ is* optimal *with respect to this optimality criterion if for every other family $F'$, we have either $F' < F$ or $F' \sim F$.*

We want to use an appropriate optimality criterion to select a family. If a criterion selected several different families as equally good, then we can use this non-uniqueness to optimize something else. For example, if we have several different families that provide an equally good approximation of the actual elevation profiles, then, from all these optimal families, we can select, e.g., the family which is the easiest to compute. This additional selection is, in effect, equivalent to replacing the original optimality criterion with the new one $<_{\rm new}$, according to which $F <_{\rm new} F'$ if:

- either $F < F'$ according to the original criterion,
- or $F \sim F'$ and $F'$ is easier to compute (in some formal sense, e.g., in terms of the computation time).

If the new criterion still selects several families as equally optimal, we can again modify it, etc. – until we end up with a *final* criterion for which there is exactly one optimal family.

**Definition 4.** *We say that an optimality criterion is* final *if it has exactly one optimal family.*

As a starting point for measuring $x$, we can take different locations. If we select a different location which is $x_0$ units before the current one, then each new location $x$ is identical to the old location $x' = x + x_0$. So, the same profile approximation that in the new units has the form $f(x)$ in the old units has the form $f(x + x_0)$. The relative quality of different profiles approximations should not change if we simply change the starting location. Thus, we arrive at the following definitions.

**Definition 5.** *For each family $F$ as described by the formula* (1) *and for each $x_0$, by a* shift $S_{x_0}(F)$, *we mean a family*

$$\{C_1 \cdot (S_{x_0} f_1)(x) + \ldots + C_n \cdot (S_{x_0} f_n)(x)\},$$

*where* $(S_{x_0} f_i)(x) \stackrel{\rm def}{=} f_i(x + x_0)$.

**Definition 6.** *We say that an optimality criterion is* shift-invariant *if for every $F$, $F'$, and $x_0$, the following two properties hold:*

- *if $F < F'$, then $S_{x_0}(F) < S_{x_0}(F')$;*
- *if $F \sim F'$, then $S_{x_0}(F) \sim S_{x_0}(F')$.*

Similarly, nothing should change if we simply change the measuring unit for $x$ – e.g., use miles instead of kilometers. If we replace the original measuring unit by a one which is $\lambda$ times larger, then the new value $x$ is identical to the old value $x' = \lambda \cdot x$. So, the same profile approximation that in the new units has the form $f(x)$ in the old units has the form $f(\lambda \cdot x)$. The relative quality of different profiles approximations should not change if we simply change the measuring unit. Thus, we arrive at the following definitions.

**Definition 7.** *For each family $F$ as described by the formula* (1) *and for each $\lambda > 0$, by a* rescaling $R_\lambda(F)$, *we mean a family*

$$\{C_1 \cdot (R_\lambda f_1)(x) + \ldots + C_n \cdot (R_\lambda f_n)(x)\},$$

*where* $(R_\lambda f_i)(x) \stackrel{\text{def}}{=} f_i(\lambda \cdot x)$.

**Definition 8.** *We say that an optimality criterion is* scale-invariant *if for every $F$, $F'$, and $\lambda > 0$, the following two properties hold:*

- *if $F < F'$, then $R_\lambda(F) < R_\lambda(F')$;*
- *if $F \sim F'$, then $R_\lambda(F) \sim R_\lambda(F')$.*

**Proposition 1.** *For every $n$ and for every final shift- and scale-invariant optimality criterion, the optimal family $F_{\text{opt}}$ consists of polynomials of order $\le n-1$.*

*Comment.* This result is similar to results from [5].

**Proof.**

$1^\circ$. Let us first prove that the optimal family is shift- and scale-invariant, i.e., that $S_{x_0}(F_{\text{opt}}) = R_\lambda(F_{\text{opt}}) = F_{\text{opt}}$ for all $x_0$ and $\lambda$.

Let us first prove shift-invariance of $F_{\text{opt}}$. Since $F_{\text{opt}}$ is optimal, for every family $F$, we have $F < F_{\text{opt}}$ or $F \sim F_{\text{opt}}$. In particular, this is true for the family $S_{-x_0}(F)$, i.e., either $S_{-x_0}(F) < F_{\text{opt}}$ or $S_{-x_0}(F) \sim F_{\text{opt}}$. Since the optimality criterion is shift-invariant, this implies that either $S_{x_0}(S_{-x_0}(F)) < S_{x_0}(F_{\text{opt}})$ or $S_{x_0}(S_{-x_0}(F)) \sim S_{x_0}(F_{\text{opt}})$. However, as one can easily check, we have $S_{x_0}(S_{-x_0}(F)) = F$. Thus, for every family $F$, we have either $F < S_{x_0}(F_{\text{opt}})$ or $F \sim S_{x_0}(F_{\text{opt}})$. By definition of optimality, this means that the family $S_{x_0}(F_{\text{opt}})$ is also optimal.

Since the optimality criterion is final, there is only one optimal family, hence $S_{x_0}(F_{\text{opt}}) = F_{\text{opt}}$. Shift-invariance is proven.

Scale-invariance is proven similarly, by taking into account that for every $F$ and every $\lambda$, either $R_{1/\lambda}(F) < F_{\text{opt}}$ or $R_{1/\lambda}(F) \sim F_{\text{opt}}$. So, by applying the scaling $R_\lambda$ to both sides of these relations, we conclude that $R_\lambda(F_{\text{opt}})$ is also optimal and thus, $R_\lambda(F_{\text{opt}}) = F_{\text{opt}}$.

$2^\circ$. Shift-invariance means that every element of the family $S_{x_0}(F_{\text{opt}})$ also belongs to the same family $F_{\text{opt}}$. Let $f_i(x)$ denote the functions whose linear combinations (1) form the family $F_{\text{opt}}$. Then, in particular, invariance means that for every $i$, the shifted function $f_i(x + x_0)$ is a linear combination of functions $f_j(x)$:

$$\begin{gathered} f_1(x + x_0) = C_{11}(x_0) \cdot f_1(x) + \ldots + C_{1n}(x_0) \cdot f_n(x); \\ \ldots \\ f_n(x + x_0) = C_{n1}(x_0) \cdot f_1(x) + \ldots + C_{nn}(x_0) \cdot f_n(x), \end{gathered} \tag{2}$$

for some coefficients $C_{ij}$ depending on $x_0$.

For each $i$, we can take $n$ different values $x_1, \ldots, x_n$ of $x$ and get a system of $n$ linear equations with $n$ unknowns $C_{i1}(x_0), \ldots, C_{in}(x_0)$:

$$f_i(x_1 + x_0) = C_{i1}(x_0) \cdot f_1(x_1) + \ldots + C_{in}(x_0) \cdot f_n(x_1);$$

$$\ldots$$

$$f_i(x_n + x_0) = C_{i1}(x_0) \cdot f_1(x_n) + \ldots + C_{in}(x_0) \cdot f_n(x_n).$$

By Cramer's rule, the solutions $C_{ij}(x_0)$ to this system can be represented as a ratio of two polynomials in terms of $f_i(\cdot)$. Since the functions $f_i(x)$ are smooth, this implies that the functions $C_{ij}(x_0)$ are also differentiable functions of $x_0$.

Thus, we can differentiate both sides of (2) by $x_0$ and take $x_0 = 0$. As a result, we get a system of linear differential equations with constant coefficients:

$$f_1'(x) = c_{11} \cdot f_1(x) + \ldots + c_{1n} \cdot f_n(x);$$

$$\ldots \tag{3}$$

$$f_n'(x) = c_{n1} \cdot f_1(x) + \ldots + c_{nn} \cdot f_n(x),$$

where we denoted $c_{ij} \stackrel{\text{def}}{=} C_{ij}'(0)$.

The general solution to such a system is well-known (see, e.g., [2, 5]): it is a linear combination of terms of the type $\exp(\lambda_i \cdot x)$ and $x^k \cdot \exp(\lambda_i \cdot x)$, where $\lambda_i$ are eigenvalues of the matrix $(c_{ij})$, and $k \le n - 1$ is a positive integer corresponding to the case when we have a multiple eigenvalue.

3°. Similarly, scale-invariance means that every element of the family $R_\lambda(F_{\rm opt})$ also belongs to $F_{\rm opt}$. In particular, this means that for every $i$, the re-scaled function $f_i(\lambda \cdot x)$ is a linear combination of functions $f_j(x)$:

$$f_1(\lambda \cdot x) = C_{11}(\lambda) \cdot f_1(x) + \ldots + C_{1n}(\lambda) \cdot f_n(x);$$

$$\ldots \tag{4}$$

$$f_n(\lambda \cdot x) = C_{n1}(\lambda) \cdot f_1(x) + \ldots + C_{nn}(\lambda) \cdot f_n(x),$$

for some coefficients $C_{ij}$ depending on $\lambda$.

For each $i$, we can take $n$ different values $x_1, \ldots, x_n$ of $x$ and get a system of $n$ linear equations with $n$ unknowns $C_{i1}(\lambda), \ldots, C_{in}(\lambda)$:

$$f_i(\lambda \cdot x_1) = C_{i1}(\lambda) \cdot f_1(x_1) + \ldots + C_{in}(\lambda) \cdot f_n(x_1);$$

$$\ldots$$

$$f_i(\lambda \cdot x_n) = C_{i1}(\lambda) \cdot f_1(x_n) + \ldots + C_{in}(\lambda) \cdot f_n(x_n).$$

By Cramer's rule, the solutions $C_{ij}(\lambda)$ to this system can be represented as a ratio of two polynomials in terms of $f_i(\cdot)$. Since the functions $f_i(x)$ are smooth, this implies that the functions $C_{ij}(\lambda)$ are also differentiable functions of $\lambda$.

Thus, we can differentiate both sides of (4) by $\lambda$ and take $\lambda = 1$. As a result, we get the following system of linear differential equations:

$$x \cdot f_1'(x) = c_{11} \cdot f_1(x) + \ldots + c_{1n} \cdot f_n(x);$$
$$\ldots \tag{5}$$
$$x \cdot f_n'(x) = c_{n1} \cdot f_1(x) + \ldots + c_{nn} \cdot f_n(x),$$

where we denoted $c_{ij} \stackrel{\text{def}}{=} C'_{ij}(1)$.

Here, for each $i$, we have

$$x \cdot f_i'(x) = x \cdot \frac{df_i}{dx} = \frac{df_i}{dx/x}.$$

Since $dx/x = d(\ln(x))$, we thus conclude that for the new variable $X = \ln(x)$ (for which $x = \exp(X)$) and for the corresponding functions $F_i(X) = f_i(\exp(X))$, we get the system of linear differential equations with constant coefficients:

$$F_1'(X) = c_{11} \cdot F_1(X) + \ldots + c_{1n} \cdot F_n(X);$$
$$\ldots \tag{6}$$
$$F_n'(x) = c_{n1} \cdot F_1(X) + \ldots + c_{nn} \cdot F_n(X).$$

Hence, similarly to the previous subsection, we conclude that each solution of this system is a linear combination of terms of the type $\exp(\lambda_i \cdot X)$ and

$$X^k \cdot \exp(\lambda_i \cdot X).$$

Substituting $X = \ln(x)$ into this formula, we conclude that each function $f_i(x) = F_i(\ln(x))$ is a linear combination of functions $\exp(\lambda_i \cdot \ln(x))$ and

$$\ln^k(x) \cdot \exp(\lambda_i \cdot \ln(x)).$$

Here, $\exp(\lambda_i \cdot \ln(x)) = (\exp(\ln(x))^{\lambda_i} = x^{\lambda_i}$.

Thus, each function $f_i(x)$ is a linear combination of functions $x^{\lambda_i}$ and

$$\ln^k(x) \cdot x^{\lambda_i}.$$

$4^\circ$. Our functions $f_i(x)$ are *both* shift-invariant and scale-invariant. Thus, each of them has to be both of form described at the end of Part 2 of this proof and of the form described at the end of Part 3. So, out of terms from Part 2, we cannot have exponential terms with non-zero $\lambda_i$ – since these terms cannot be expressed in Part-3 form. Thus, the only possible terms are terms $x^k$ with $k \le n - 1$.

So, each function $f_i(x)$ is a linear combination of such terms – and is, thus, a polynomial of order $\le n - 1$. The proposition is proven.

# 3 Why Convexity and Concavity Are Important in Elevation Profiles: An Explanation Based on the Optimality Result

**Discussion.** The above result provides us, for different $n$, with families of approximations to the elevation profiles. Let us start with the simplest possible approximation.

For $n = 1$, we get the class of constant functions – no landscape at all. For $n = 2$, we get a class of linear functions – no mountains, no ravines, just a flat inclined surface. So, the only non-trivial description of a landscape starts with $n = 3$, i.e., with quadratic functions.

We want to provide a qualitative classification of all such possible elevation functions. It is reasonable to say that the two elevation functions are equivalent if they differ only by re-scaling and shift of $x$ and $y$:

**Definition 9.** *We say that two quadratic functions $f(x)$ and $g(x)$ are* equivalent *if for some values $\lambda_x > 0$, $\lambda_y > 0$, $x_0$, and $y_0$, we have*

$$g(x) = \lambda_y \cdot f(\lambda_x \cdot x + x_0) + y_0$$

*for all $x$.*

**Proposition 2.** *Every non-linear quadratic function is equivalent either to $x^2$ or to $-x^2$.*

**Discussion.** Thus, in this approximation, we have, in effect, two shapes: the shape corresponding to $x^2$ (convex) and the shape corresponding to $-x^2$ (concave). This result explains why our visual classification into convex and concave shapes makes perfect sense.

**Proof.** Every non-linear quadratic function $g(x)$ has the form

$$g(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2,$$

for some $a_2 \neq 0$.

If $a_2 > 0$, then this function can be represented as

$$a_2 \cdot \left(x + \frac{a_1}{2a_2}\right)^2 + \left(a_0 - \frac{a_1^2}{4a_2}\right),$$

i.e., can be represented in the desired form, with $f(x) = x^2$, $\lambda_x = 1$, $\lambda_y = a_2$, $x_0 = \dfrac{a_1}{2a_2}$, and $y_0 = a_0 - \dfrac{a_1^2}{4a_2}$.

If $a_2 < 0$, then this function can be represented as

$$|a_2| \cdot \left(-\left(x + \frac{a_1}{2a_2}\right)^2\right) + \left(a_0 - \frac{a_1^2}{4a_2}\right),$$

i.e., can be represented in the desired form, with $f(x) = -x^2$, $\lambda_x = 1$, $\lambda_y = |a_2|$, $x_0 = \dfrac{a_1}{2a_2}$, and $y_0 = a_0 - \dfrac{a_1^2}{4a_2}$.

The proposition is proven.

# 4 Bhutan Anomaly Explained

**Discussion.** In the previous text, we have shown that the optimal description of an elevation profiles is by polynomials of a fixed degree.

In the first approximation, a landscape profile can be described by a quadratic function. To get a more accurate description, let us also consider cubic terms, i.e., let us consider profiles of the type

$$f(x) = a_0 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3. \tag{7}$$

As a starting point $x = 0$ for the elevation profile, it makes sense to select the lowest (or the highest) point. In both cases, according to calculus, the first derivative of the elevation profile is equal to 0 at this point: $f'(0) = 0$. Substituting the above expression for $f(x)$ into this formula, we conclude that $a_1 = 0$ and thus,

$$f(x) = a_0 + a_2 \cdot x^2 + a_3 \cdot x^3. \tag{8}$$

Let us analyze how this approximation works for the above two cases: the case of Nepal and the case of Bhutan.

**Case of Nepal.** In the case of Nepal, the forces compressing the upper plate are orthogonal to the line of contact. This means that in this case, the forces do not change if we change left to right and right to left.

Since the whole mountain range was created by this force, it is reasonable to conclude that the corresponding elevation profile is also invariant with respect to swapping left and right, i.e., with respect to the transformation $x \to -x$:

$$f(x) = f(-x). \tag{9}$$

Substituting the cubic expression (8) for the profile $f(x)$ into this formula, we conclude that $a_3 = 0$. Thus, in this case, the elevation profile is quadratic even in this next approximation – and is, therefore, either convex or concave.

**Case of Bhutan.** In the case of Bhutan, the force is applied at an angle. Here, there is no symmetry with respect to $x \to -x$, so, in general, we have $a_3 \neq 0$. Thus, the second derivative – that describes whether a function is locally convex (when this second derivative is positive) or locally concave (when the derivative is negative) – becomes a linear function $6a_3 \cdot x + 2a_2$, with $a_3 \neq 0$.

A non-constant linear function always changes signs – this explains why in the case of Bhutan, convexity follows by concavity.

## 5 Auxiliary Question: How to Best Locate an Inflection Point

**Practical problem.** Many geophysical ideas are applicable only to valley-type convex domains or only to mountain-type concave domains. So, to apply these ideas to a real-life landscape, it is necessary to divide the whole landscape into convex and concave zones. What is the best way to do it? In other words, what is the best way to locate an *inflection point*, i.e., the point at which local convexity changes to local concavity?

**First idea: a straightforward least squares approach.** The first natural idea – motivated by the above analysis – is to approximate the actual elevation profile by a cubic function (7). The corresponding coefficients $c_0$, $c_1$, $c_2$, and $c_3$ can be obtained, e.g., by applying the least squares method to the corresponding system of linear equations

$$y_i \approx c_0 + c_1 \cdot x_i + c_2 \cdot x_i^2 + c_3 \cdot x_i^3,$$

where $x_i$ is the $i$-th location and $y_i$ is the $i$-th elevation.

The least squares method minimizes the sum

$$\sum_i (y_i - (c_0 + c_1 \cdot x_i + c_2 \cdot x_i^2 + c_3 \cdot x_i^3))^2.$$

Differentiating this expression with respect to each of the unknowns $c_j$ and equating all four derivatives to 0, we get an easy-to-solve system of four linear equations with four unknowns.

Once we find the characteristics, we then estimate the location of the inflection point as the value at which the second derivative is equal to 0, i.e., the value $x_{\text{infl}} = -\dfrac{c_2}{3c_3}$.

**Second idea: a model-free least squares approach.** Instead of restricting ourselves to a cubic approximation, we can consider general convex functions. For a function $f(x)$ defined by its values $y_1 = f(x_1)$, $y_2 = f(x_2)$, $\ldots$, on a equally spaced grid

$$x_1, x_2 = x_1 + \Delta x, x_3 = x_1 + 2\Delta x, \ldots, x_N,$$

convexity is equivalent to the sequence of inequalities

$$y_i \le \frac{y_{i-1} + y_{i+1}}{2}. \tag{10}$$

For each set of actual profile points $\widetilde{y}_i$, we can therefore find the closest convex profile by looking for the values $y_i$ that minimize the mean square error (MSE)

$$\frac{1}{N} \cdot \sum_i (\widetilde{y}_i - y_i)^2$$

under the constraints (10). The minimized expression is a convex function of the unknowns $y_i$, and each constraint – and thus, their intersection – defines a convex set. Thus, we can find the corresponding minimum by using a known algorithm for convex optimization (= minimizing a convex function on a convex domain); see, e.g., [6, 7, 8].

By applying this algorithm to actually convex profiles, we can find the largest and thus, the corresponding MSE. Let us denote the largest of such values by $M$. Then, to find an inflection point, we can consider larger and larger fragments of the original series $f(x_1)$, $f(x_2)$, $\ldots$, until we reach a point at which the corresponding MSE exceeds $M$. This is the desired inflection point.

We can speed up this algorithm if instead of slowly increasing the size of the still-convex fragment, we use bisection. Specifically, we always keep two values $\underline{p}$ and $\overline{p}$ such that the fragment until $\underline{p}$ is convex (within accuracy $M$), while the fragment up to the point $\overline{p}$ is not convex within the given accuracy.

In the beginning, we first apply our criterion to the whole list of $N$ values. If the result is $M$-close to convex, we consider the profile convex – no inflection point here. If the result is not $M$-convex, then we take $\underline{p} = 1$ and $\overline{p} = N$.

Once we have two values $\underline{p} < \overline{p}$, we then take a midpoint $m \stackrel{\text{def}}{=} \dfrac{\underline{p} + \overline{p}}{2}$. If the segment up to this midpoint is $M$-convex, then we replace $\underline{p}$ with $m$. If this segment is not $M$-convex, we replace $\overline{p}$ with $m$.

In both case, we get a new interval $[\underline{p}, \overline{p}]$ whose width decreased by a factor of two. We started with width $N$. Thus, in $\log_2(N)$ steps, this size decreases to $N/2^{\log_2(N)} = N/N = 1$, i.e., we get the exact location of the inflection point.

*Comment.* Other algorithms for detecting inflection points are described, e.g., in [3, 4].

# Acknowledgments

This work was supported in part by the US National Science Foundation via grant HRD-1242122 (Cyber-ShARE Center of Excellence).

The authors are greatly thankful to Byron Adams for valuable discussions.

# References

[1] B. A. Adams, K. X. Whipple, K. V. Hodges, and A. M. Heimsath, "In situ development of high-elevation, low-relief landscapes via duplex deformation in the Eastern Himalayan hinterland, Bhutan", *Journal of Geophysical Research: Earth Surface*, 2016, Vol. 121, pp. 294–319.

[2] K. B. Howell, *Ordinary Differential Equations: An Introduction to the Fundamentals*, CRC Press, Boca Raton, Florida, 2016.

[3] N. K. Kachouie and A. Schwartzman, "Non-parametric estimation of a single inflection point in noisy observed signal", *Journal of Electrical and Electronic Systems*, 2013, Vol. 2, No. 2, Paper 1000108.

[4] D. Manocha and J. F. Canny, "Detecting cusps and inflection points in curves", *Computer Aided Geometric Design*, 1992, Vol. 9, No. 1, pp. 1–24.

[5] H. T. Nguyen and V. Kreinovich, *Applications of Continuous Mathematics to Computer Science*, Kluwer, Dordrecht, 1997.

[6] J. Nocedal and S. Wright, *Numerical Optimization*, Springer, New York, 2006.

[7] P. Pardalos, *Complexity in Numerical Optimization*, World Scientific, Singapore, 1993.

[8] R. T. Rockafeller, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1997.