Trend Analysis in the Age of Machine Learning

Manel Baucells* Emanuele Borgonovo† Elmar Plischke‡ John Barr§
Herschel Rabitz¶

Version of May 21, 2021

Abstract

Trend indicators aim to visualize the effect of input assumptions and parameters into model output. Their use is increasingly important, especially when the output is generated by a black box algorithm. We investigate the properties of several trend indicators used in simulation and machine learning, and discover that not all give consistent results when applied to well understood problems. We show that some indicators (individual conditional expectations (ICE), gradients, and partial dependence functions) are convexity and monotonicity consistent even when inputs are not independent; while some others (ALE plots, regression lines, correlation coefficients) may not be. We advocate the combined use of ICE and partial dependence plots, discuss their connection with Tornado Diagrams and Spiderplots, and propose two discrepancy indices to ease their interpretation.

Keywords: Sensitivity Analysis, Decision Analysis, Grapical Visualization, Monotonicity, Convexity

1 Introduction

Determining how a quantity of interest responds to variations in exogenous variables is essential in management science modeling. Decision support models include machine learning tools fitted to available data as well as dedicated simulators. Examples of machine learning tools include classification trees (Bertsimas and O'Hair 2013), support vector machines (Cecchini et al. 2010) and artificial neural networks (Kim et al. 2005). Examples of dedicated simulators include discrete event models (Hong and Nelson 2006, Nelson 2013), integrated assessment models (Hu et al. 2012), SEIR (Susceptible, Exposed, Infectious, Recovered) models employed to support predictions during the COVID-19 pandemic (Currie et al. 2020), and agent based models (Rahmandad and Sterman 2008). The growth in computing power allows analysts to employ datasets of large dimensionality and to build simulators of increasing complexity. Decision-makers run the risk of remaining confused or overwhelmed if results are not properly visualized and understood.

Trend analysis is a collection of methods to glimpse the shape of the input-output mapping. In machine learning, there is a growing interest in studying the properties of trend indicators (Goldstein et al. 2015, Hooker and Mentch 2019, Apley and Zhu 2020). Trend analysis assists various phases of the decision-making process. In the model building phase, for instance, studying trend serves as a debugging tool. If a simulator is meant to model a relationship known to be by theory to be monotonic or convex, then a trend analysis should verify whether the simulator response is indeed monotonic or convex.

In decision analysis, popular trend indicators include spiderplots (Eschenbach 1992), one-way sensitivity functions (Clemen 1997, van der Gaag et al. 2007), and Tornado Diagrams (Howard

*Darden Business School
†Bocconi University
‡Clausthal University of Technology
§Princeton University
¶Princeton University

1988). These are local indicators, not engineered to take into account uncertainty or statistical dependence among the inputs (Wang et al. 2017). Alternatively, analysts can employ differential approaches (Rabitz 1989), correlation coefficients (Pearson 1895, Levy et al. 2001), regression functions (Chambers and Hastie 1992), partial dependence functions (Friedman et al. 2001), individual conditional expectation (ICE) plots (Goldstein et al. 2015) or the accumulated local effects (ALE) plots (Apley and Zhu 2020). While all these methods seek an answer to the same question, it is unclear the extent to which they do so consistently. Consider the business case discussed in Eschenbach (1992), where the present worth of an investment is expressed as a function of several exogenous variables (see Supplementary Appendix 10 for details). The decision-maker is investigat-

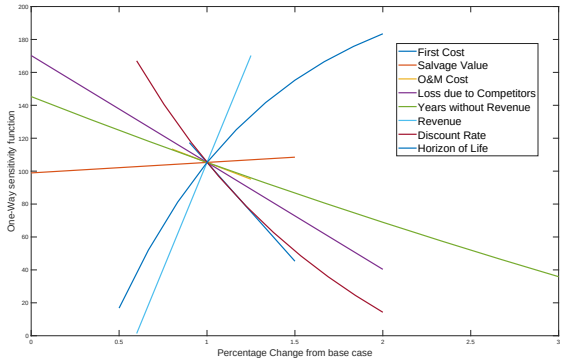


Figure 1: Spiderplot of Eschenbach (1992).

ing intervention possibilities to increase profit and current focus is on "horizon of life". The inputs are uncertain with the distributions in Appendix 10. Figure 1 reports the original spiderplot in Eschenbach (1992, Figure 2, p. 43). A spider plot is a collection of one-way sensitivity functions, one for each different input. The common point for all lines is a base case anchor. The graph suggests that if we increase horizon of life we increase the investment present worth. However, because the inputs are uncertain and the input-output mapping is approximately linear, the analyst considers the correlation coefficient between present worth and horizon of life obtaining a negative value. Should the analyst recommend the opposite course of action than Eschenbach's spiderplot? More broadly, if the insights coming from alternative trend indicators cannot be reconciled, which method or ensemble of methods is best?

Our goal is to examine two general properties of several trend indicators, namely monotonicity and convexity consistency. A trend indicator is monotonicity/convexity consistent if, given that the input-output mapping is monotonic/convex, it always reports such intrinsic monotonicity/convexity. We find that gradients (when defined) and one-way-sensitivity functions are always monotonicity consistent, hence so are ICE plots. At the same time, correlation coefficients and conditional regression functions are consistent when inputs are independent, but they may fail to be consistent when inputs are dependent. In contrast, partial dependence functions are monotonicity/convexity-consistent even when inputs are correlated. We also find that the sign of individual ALE plot functions is a monotonicity-consistent indicator of trend. We then study the connection between one-way sensitivity functions and partial dependence functions and introduce two new indices that aid the interpretation of the graphical results.

We examine consistency more in detail on the important family of multilinear mappings. We illustrate the findings through numerical experiments involving a SEIR model in the context of the COVID-19 pandemic.

2 Review of Existing Trend Indicators

We adopt the setup of Zhao and Hastie (2019, p. 2) where a quantity of interest, Y, depends on a set of inputs or parameters, $\mathbf{X}$. We presume that the input-output mapping is the result of a machine learning algorithm or a complex simulation, hence not transparently observable. Following standard notation, we write:

$$Y = g(\mathbf{X}) + \mathcal{E}(\mathbf{X}), \tag{1}$$

where $g : \mathcal{X} \to \mathbb{R}$, $\mathcal{X} \subseteq \mathbb{R}^n$ denote a multivariate input-output mapping, and $\mathcal{E}$ a stochastic error term with zero mean, hence $\mathbb{E}[Y|\mathbf{X} = \mathbf{x}^0] = g(\mathbf{x}^0)$. Specifically, we assume that $\mathbf{X}$, Y and $\mathcal{E}(X)$ are random vectors on a reference space $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$. Furthermore, let $F_{\mathbf{X}}$ and F_Y denote the cumulative distribution functions (cdf) of the input and output, respectively, while $f_{\mathbf{X}}$ and f_Y denote their respective densities (pdf), if available. If $\mathcal{E}$ is identically null, the simulator is said to be deterministic. Notationwise, it is useful to partition the domain $\mathcal{X}$ into $\mathcal{X}_i \subseteq \mathbb{R}$ and $\mathcal{X}_{-i} \subseteq \mathbb{R}^{n-1}$, with $\mathcal{X}_{-i} = \mathcal{X} \setminus \mathcal{X}_i$. We also write $\mathbf{x} \in \mathcal{X}$, $x_i \in \mathcal{X}_i$ and $\mathbf{x}_{-i} \in \mathcal{X}_{-i}$, and $\mathbf{x} = (x_i; \mathbf{x}_{-i})$.

Table 1: Trend Indicators.

Trend Indicator	**Symbol**	**Monotonicity Consistent**	**Convexity Consistent**	**Allows Discrete**	**Handles Uncertainty**
Gradient, Hessian	$g_i'(\mathbf{X})$	Always	Always	No	If randomized
One-way sensitivity / ICE plots	$g(x_i; \mathbf{x}^0_{-i})$	Always	Always	Yes	If randomized
Correlation coefficient	ρ_{YX_i}	Independence	N/A	Yes	Yes
Regression function	$r_i(x_i)$	Independence	Independence	Yes	Yes
Partial dependence	$h_i(x_i)$	Cartesian support	Cartesian support	Yes	Yes
ALE function	$ALE_i(x_i)$	Independence	Independence	Yes	Yes

Table 1 lists a number of trend indicators. If g is differentiable, a popular approach is represented by the calculation of gradients for monotonicity and Hessians for convexity, possibly using automatic differentiation (Griewank 2000). The signs of a first-order and second-order derivatives informs analysts about whether the output is, respectively, locally increasing/decreasing or convex/concave at $\mathbf{x}^0$.

In decision analysis, a popular class of trend indicators is represented by one-way sensitivity functions (Castillo et al. 1997, Clemen 1997). We fix a baseline point $\mathbf{x}^0 \in \mathcal{X}$ and, for each input, we consider the functions $g(x_i; \mathbf{x}^0_{-i})$ in which all inputs but x_i are fixed at their reference values $(\mathbf{x}^0_{-i})$, while x_i varies in $\mathcal{X}_i$. Thus, for each input there are as many one-way sensitivity functions as there are baseline locations $\mathbf{x}^0_{-i}$. The idea of plotting multiple one-way sensitivity functions at different baseline points is at the basis of ICE plots, introduced in machine learning by Goldstein et al. (2015). In the management sciences, one-way sensitivity functions have been intensively studied (van der Gaag et al. 2007), frequently implemented in commercial software (e.g., Treeage and @Risk) and visualized via break-even diagrams or spider-plots.

The indicators mentioned thus far provide local information and do not account for input uncertainty. A global approach requires the analyst to recognize that the inputs and outputs are uncertain. We assume that she has available a dataset of input-output realizations, possibly generated via a Monte Carlo uncertainty propagation. A first and popular way to gain indication about trend is to compute correlation coefficients. A positive (negative) sign of these coefficients indicates that g is increasing (decreasing) with respect to X_i in expectation. In fact, ρ_{Y,X_i} is proportional to the slope of the regression line $\mathbb{E}[Y|X_i = x_i] = a + bx_i$. Thus, it is an appropriate measure of trend when the relationship between Y and X_i is linear.

More generally, analysts rely on regression functions, by calculating and graphing the (not

necessarily linear) curve

$$r_i(x_i) = \mathbb{E}[g(\mathbf{X})|X_i = x_i] = \int_{\mathcal{X}_{-i}} g(x_i; \mathbf{x}_{-i}) dF_{\mathbf{X}_{-i}|X_i}(x_i; \mathbf{x}_{-i}), \tag{2}$$

This function yields information on the expected behavior of g conditional on X_i. Graphs of regression functions are part of data visualization tools such as conditioning plots (Chambers and Hastie 1992).

Partial dependence functions were introduced in machine learning by Friedman et al. (2001), and have been subsequently studied by Hooker (2007) and Goldstein et al. (2015). In their survey, Guidotti et al. (2018) include them among the tools recommended for increasing interpretability of machine learning algorithms (see also (Molnar 2018, Ch. 5)). Formally, a partial dependence function is defined as

$$h_i(x_i) = \int_{\mathcal{X}_{-i}} g(x_i; \mathbf{x}_{-i}) dF_{\mathbf{X}_{-i}}(\mathbf{x}_{-i}), \tag{3}$$

where $(x_i; \mathbf{x}_{-i})$ is a point in $\mathcal{X} = \mathcal{X}_i \times \mathcal{X}_{-i}$. A partial dependence function communicates the behavior of g as a function of X_i, averaging over the marginal distribution on the remaining inputs. We can think of $h_i(x_i)$ as the average of the one-way sensitivity function $g(x_i; x_{-i})$ over x_{-i}.

Recently, Hooker and Mentch (2019), Apley and Zhu (2020) and Molnar et al. (2020) highlight limitations of partial dependence functions when inputs are correlated. The common problem is that the computation of (3) may force the machine learning algorithm to generate forecasts in regions where no data are available. Extrapolation errors may result in misleading graphical representations of the input-output dependence.

Apley and Zhu (2020) propose a new graphical indicator called accumulated local effects (ALE) plot, defined as:

$$ALE_i(x_i) = \int_{x_{i,\min}}^{x_i} \mathbb{E}[g_i'(\mathbf{X})|X_i = z_i] dz_i, \tag{4}$$

where $g_i'(\mathbf{X})$ is the partial derivative of g with respect to X_i. This indicator is global, as it requires the evaluation of partial derivatives at randomized locations of the input space.

Pearson's linear correlation coefficient, regression functions, gradients and Hessians have been widely studied and their properties are well known. In contrast, partial dependence, ICE and ALE plots are of more recent introduction and their properties are less understood. Thus, we do not know under which conditions the different indicators provide univocal insights, or whether the indicators preserve properties such as monotonicity and convexity of the input-output mapping. Our goal is to bridge this gap.

3 An Analytical Viewpoint

To guide our analytical investigation, we introduce the principle of monotonicity/convexity consistency, respectively in the first and second part of this section.

3.1 Monotonicity-Consistency

For simplicity, we consider increasing input-output mappings, but the results trivially hold for decreasing ones.

Definition 1 (Monotonicity)**.** A function $g : \mathcal{X} \to \mathbb{R}$ is separately increasing in x_i if

$$g(x_i + t_i; \mathbf{x}_{-i}) \geq g(x_i; \mathbf{x}_{-i}) \tag{5}$$

for all $t_i > 0$, with $(x_i + t_i; \mathbf{x}_{-i})$ and $(x_i; \mathbf{x}_{-i}) \in \mathcal{X}$.

We call a trend indicator *monotonicity-consistent* if, given that g is separately increasing with respect to x_i, the indicator reports such monotonicity. The rationale is self-evident, for, if the analyst is aware that the simulator response is increasing in a certain input, she expects the trend indicator to report such behavior. A violation of consistency would surely question the reliability of the indicator.

Let us now analyze whether the indicators in Table 1 are monotonicity-consistent. All proofs are reported in Appendix 9.

Proposition 2. *Given any $(\mathcal{X}, \mathbb{B}(\mathcal{X}), \mathbb{P}_{\mathbf{X}})$, $g : \mathcal{X} \mapsto \mathbb{R}$, one-way sensitivity functions are monotonicity-consistent indicators. Moreover, if g has first order partial derivatives, gradients are monotonicity-consistent indicators in sign.*

Example 3. Let $g(X_1, X_2) = X_1 X_2$, with $X_1, X_2 > 0$. We focus on the dependence of g on X_1. We have $g(x_1; x_2^0) = x_1 x_2^0$ for all x_2^0. Because X_2 is positive, one-way sensitivity functions are increasing straight lines for all x_2^0. Similarly the partial derivative $g_1'(X_1, X_2) = X_2$ is always positive.

The following example shows that regression functions and correlation coefficients are not monotonicity consistent under input dependence.

Example 4 (Example 3 continued). Let $g(X_1, X_2)$ be as in Example 3, but with X_1 and X_2 having the joint probability distribution given in Table 2.

Table 2: Joint Probability Distribution $P(X_1 = x_1; X_2 = x_2)$ for Example 4.

	$X_1 = 1$	$X_1 = 2$	$X_1 = 3$	$P(X_1 = x_i)$	$r_1(x_1)$
$X_2 = 1$	0.25	0.10	0.01	0.36	0.85
$X_2 = 2$	0.30	0.01	0.01	0.32	0.48
$X_2 = 3$	0.30	0.01	0.01	0.32	0.27
$P(X_2 = x_i)$	0.85	0.12	0.03	1	

While $g(X_1, X_2)$ is separately increasing in X_1, we register a negative correlation coefficient $\rho_{Y,X_1} = -0.083$, and the regression function $r_1(x_1)$ is decreasing (Table 2, last column).

Regression functions $r_i(x_i)$ become monotonicity consistent under the following condition.

Proposition 5. *Let g be separately increasing in X_i. Then, a sufficient condition for $r_i(x_i)$ to be monotonicity consistent is that for all $(x_i + h, \mathbf{X}_{-i})$, $(x_i, \mathbf{X}_{-i}) \in \mathcal{X}$*

$$\frac{f_{\mathbf{X}_{-i}}(\mathbf{X}_{-i} | X_i = x_i + h)}{f_{\mathbf{X}_{-i}}(\mathbf{X}_{-i} | X_i = x_i)} \geq \frac{g(x_i, \mathbf{X}_{-i})}{g(x_i + h, \mathbf{X}_{-i})}. \tag{6}$$

Thus, the ratio between the conditional densities of $\mathbf{X}_i$ given that X_i is at $x_i + h$ vs x_i must exceed the ratio between the values of g at $(x_i, \mathbf{X}_{-i})$ and at $(x_i, \mathbf{X}_{-i})$. Previous literature has shown that regression functions are monotonicity consistent under independence (Beccacece and Borgonovo 2011), which guarantees (6). Indeed, if $\mathbf{X}_{-i}$ is independent of X_i, then the left hand side is equal to 1 and if g is separately increasing in X_i, then the ratio on the right hand side is less than or equal to 1. Note that the inequality in Equation 6 is reversed from "$\geq$" to "$\leq$" in the case g is separately decreasing in X_i, and that it holds also for discrete inputs, with densities replaced by the probabilities.

We now analyze partial dependence functions and show that are monotonicity-consistent even when inputs are dependent. The result requires that the support be the Cartesian product of the supports.

Proposition 6. *Consider $(\mathcal{X}, \mathbb{B}(\mathcal{X}), \mathbb{P}_{\mathbf{X}})$, with $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 ... \times \mathcal{X}_d$. If g is separately increasing in X_i, then $h_i(x_i)$ is increasing.*

The converse, however, is not always true: the fact $h_i(x_i)$ is increasing is not sufficient to guarantee that input-output mapping is monotonic, except when the input-output is additively or multiplicatively separable. Specifically, consider a mapping g additively separable in X_i, $g(\mathbf{x}) = a_i(x_i) + A_{-i}(\mathbf{x}_{-i})$, where $a_i : \mathcal{X}_i \mapsto \mathbb{R}$, and $A_{-i} : \mathcal{X}_{-i} \mapsto \mathbb{R}$. As Apley and Zhu (2020) state, partial dependence functions possess a convenient *additive recovery* property: $h_i(x_i)$ is equal to the true effect $a_i(x_i)$ plus a constant (Apley and Zhu 2020, p. 6). Thus, a partial dependence function reports the marginal behavior of g as a function of X_i exactly. Similarly, when g is multiplicatively separable in X_i, $g(\mathbf{x}) = m_i(x_i)M_{-i}(\mathbf{x}_{-i})$, the partial dependence function $h_i(x_i)$ is proportional to $m_i(x_i)$ (Hastie et al. 2009). Hence, it possesses an analogous multiplicative recovery property.

Corollary 7. *Let g be additively or multiplicatively separable in X_i. Then, g is monotonic in X_i if and only if $h(x_i)$ is monotonic in x_i.*

Let us now come to the recently introduced ALE plots.

Proposition 8. *Consider any $(\mathcal{X}, \mathbb{B}(\mathcal{X}), \mathbb{P}_{\mathbf{X}})$. If g is separately monotonic in X_i, then the function $ALE_i(x_i)$ is monotonicity-consistent in sign.*

Example 9 (Example 3 continued). Consider the input-output mapping in Example 3, now with X_1 and X_2 jointly exponentially distributed with the Gumbel (1960) cumulative distribution function $F_{\mathbf{X}}(x_1, x_2; \delta) = 1 - e^{-x_1} - e^{-x_2} + e^{-(x_1+x_2+\delta x_1 x_2)}$, $0 \leq \delta \leq 1$. The function $ALE_1(x_1)$ can be computed analytically:

$$ALE_1(x_1) = \begin{cases} \dfrac{(\delta x_1 + 1)(\ln\left(x + \frac{1}{\delta}\right) + \ln \delta) + \delta^2 x_1}{\delta^2 x_1 + \delta} & if \quad \delta \neq 0 \\ x_1 & if \quad \delta = 0 \end{cases} . \tag{7}$$

The top panel in Figure 2 displays the graphs of $ALE_1(x_1)$ for different values of δ. If $\delta = 0$, then $ALE_1(x_1) = h_1(x_1) = x_1$ (dotted line). Note that by the multiplicative recovery property $h_1(x_1) = x_1$ independently of δ. If $\delta > 0$ $ALE_1(x_1)$ is positive, consistently with the fact that g is increasing in X_1; however, the graph of $ALE_1(x_1)$ displays a non-linear dependence.

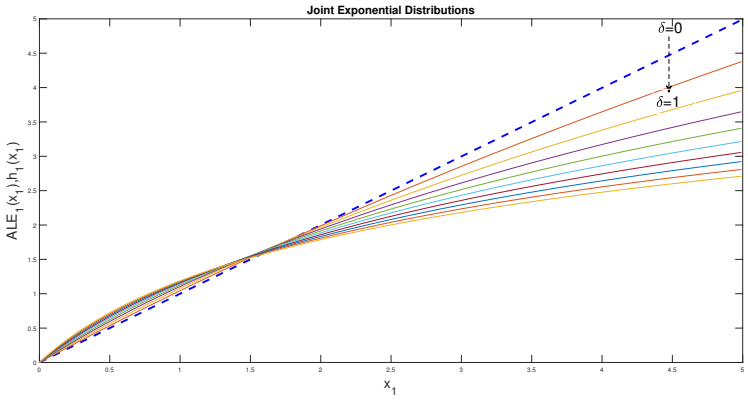


Figure 2: Plots of $ALE_1(x_1)$ [solid lines] and $h_1(x_1)$ [dotted lines] for the product model in Example 3.

Recall that ALE plots average the slope. Hence, the monotonicity consistency of first order ALE functions is to be understood in terms of their sign. The graph itself of an individual ALE function, however, does not necessarily reflect the original dependence of g on X_i. An exception is the case in which g is additive (Apley and Zhu 2020). The converse may not hold, as a constant sign of $ALE_i(x_i)$ is not a sufficient condition for the monotonicity of g.

Example 10. Consider the input-output mapping in Example 3 now with X_1 and X_2 jointly normal with correlation coefficient ρ and mean value and variance equal to unity. The function $ALE_1(x_1)$ in this case can be written as:

$$ALE_i(x_i) = (1-\rho)(x_1 - x_{min,1}) - \frac{\rho}{2}(x_1^2 - x_{min,1}^2).$$

The linear behavior is recovered only in the case $\rho = 0$. Note that if one sets $x_{min,1}$ at $-\infty$, $ALE_i(x_i)$ is ill-posed for any value of ρ, and otherwise $ALE_i(x_i)$ depends on an arbitrary $x_{min,1}$; this is a potential shortcoming of ALE plots when the support of X_i is unbounded. In the graph, we set $x_{min,1} = -5$. Note that for values of $\rho > 0$ we register $ALE_i(x_i) > 0$. However, the input-output mapping is non-monotonic in X_1, because X_2 can now assume negative values. In fact, for $\rho < 0$, $ALE_i(x_i)$ could be negative, evidencing the non-monotonicity of g.

Finally, we note that, except for gradients, the indicators $ALE_i(x_i)$, $h_i(x_i)$, $r_i(x_i)$ and ρ_{Y,X_i} can be defined also for discrete random variables and results remain the same.

3.2 Affine Multilinear Models

Numerous decision support models in the management sciences are characterized by a multilinear input-output mapping. Examples include Pseudo-Boolean functions (Foldes and Hammer 2005), Bayesian network polynomials (Darwiche (2003), see also Bhattacharjya and Shachter (2010, 2012)), reliability polynomials (Borgonovo and Smith 2011), linear regression response surfaces (Kleijnen 2017) and certain families of multicriteria utility functions (Bordley and Kirkwood 2004). We now examine the performance of different trend indicators when applied to this family of mappings.

Let $Z = \{1, 2, \ldots, n\}$ be the set of all model input indices and $\mathcal{Z} = 2^Z$ the corresponding power set. An affine multilinear function takes the form

$$g^{ML}(\mathbf{x}) = \sum_{z \in \mathcal{Z}} \beta_z \prod_{s \in z} x_s, \tag{8}$$

with the convention that sum and product on the empty set yields 0 and 1, respectively.

Proposition 11. *Given $(\mathcal{X}, \mathcal{B}(\mathcal{X}), \mathbb{P}_{\mathbf{X}})$, with $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_n$, let $g^{ML} : \mathcal{X} \mapsto \mathbb{R}$. Denote the standard deviations of X_i and $g^{ML}(\mathbf{X})$ by $\sigma_i = \sqrt{\mathbb{V}[X_i]}$ and $\sigma_Y = \sqrt{\mathbb{V}[g^{ML}(\mathbf{X})]}$. It holds:*

1. *Any one-way sensitivity function is linear and written as*

$$g^{ML}(x_i; \mathbf{x}_{-i}) = a_{-i}(\mathbf{x}_{-i})x_i + b_{-i}(\mathbf{x}_{-i}), \tag{9}$$

where $a_{-i}(\mathbf{x}_{-i}) = \sum_{z \in \mathcal{Z}_i} \beta_z \prod_{s \in z, s \neq i} x_s$ and $b_i(\mathbf{x}_{-i}) = \sum_{z \in \mathcal{Z}_{-i}} \beta_z \prod_{s \in z} x_s$.

2. *Correlation coefficients, regression, partial dependence and first order ALE functions are given, respectively, by:*

$$\rho_{YX_i} = \frac{\mathbb{E}\left[a_i(\mathbf{X}_{-i})X_i^2 + b_i(\mathbf{X}_{-i})X_i\right] - \overline{x}_i \;\; \mathbb{E}[a_i(\mathbf{X}_{-i})X_i + b_i(\mathbf{X}_{-i})]}{\sigma_Y \sigma_i} \tag{10}$$

$$r_i(x_i) = x_i \mathbb{E}[a_{-i}(\mathbf{X}_{-i})|X_i = x_i] + \mathbb{E}[b_{-i}(\mathbf{X}_{-i})|X_i = x_i], \tag{11}$$

$$h_i(x_i) = x_i \overline{a}_i + \overline{b}_i, \;\; \textit{and} \tag{12}$$

$$ALE_i(x_i) = \int_{x_{\min,i}}^{x_i} \mathbb{E}[a_{-i}(\mathbf{X}_{-i})|X_i = t]dt, \tag{13}$$

where $\overline{a}_{-i} = \mathbb{E}[a_{-i}(\mathbf{X}_{-i})]$, $\overline{b}_{-i} = \mathbb{E}[b_{-i}(\mathbf{X}_{-i})]$ and $\overline{x}_i = \mathbb{E}[X_i]$.

3. If X_i is independent of $\mathbf{X}_{-i}$, then:

$$\rho_{YX_i} = \frac{\overline{a}_i \sigma_i}{\sigma_Y}, \tag{14}$$

$$r_i(x_i) = h_i(x_i) = \overline{a}_{-i} x_i + \overline{b}_{-i}, \text{ and} \tag{15}$$

$$ALE_i(x_i) = \overline{a}_{-i}(x_i - x_{\min,i}). \tag{16}$$

Equations (9) and (12) reveal that one-way-sensitivity and partial dependence functions are linear in X_i, independently of whether the inputs are dependent. Note that the slope of a one-way sensitivity function, $a_i(\mathbf{x}^0_{-i})$, is the local slope of the multilinear map at $\mathbf{x}^0$. In contrast, the slope of a partial dependence function, $\overline{a}_i$, is the average slope of the input-output mapping. Specifically, $\overline{a}_i$ is equal to the mean value of the partial derivative of $g^{ML}(\mathbf{X})$ with respect to X_i. Thus, partial dependence functions deliver average information about the trend of Y as a function of X_i.

Regarding regression functions, note that the linear term x_i in Equation (11) is multiplied by the conditional expectation of the slope $a_{-i}(\mathbf{X}_{-i})$, which is a function of X_i. Similarly, the conditional expectation of the intercept $\mathbb{E}[b_{-i}(\mathbf{X}_{-i})|X_i = x_i]$ depends on X_i. Hence, the potential non-linearity/inconsistency of $r_i(x_i)$ is generated by the dependence of the conditional expectation on X_i. In contrast, the marginal expectation associated with partial dependence functions, makes the values of the slope $\mathbb{E}_{\mathbf{X}_{-i}}[a_{-i}(\mathbf{X}_{-i})]$ and of the intercept $\mathbb{E}_{\mathbf{X}_{-i}}[b_{-i}(\mathbf{X}_{-i})]$ independent of X_i, thus preserving the linear dependence (see Example 12).

Regarding individual ALE functions, Equations (13) and (16) show that these functions are linear in the independent input case, but may not be linear in the dependent case. The potential non-linearity is induced by the conditional expectation. Regarding correlation coefficients, Equation (14) shows that when inputs are independent ρ_{Y,X_i} is proportional to the slope of $h_i(x_i)$, hence also consistent with partial dependence functions. This is not true if inputs are dependent.

Example 12. Consider the multilinear input output mapping $g^{ML}(X_1, X_2, X_3) = X_1X_2 + X_1X_3$ with $\mathbf{X}$ multivariate normal with mean $\mu = [1, 1, 1]$ and correlation matrix Σ equal to the identity. Under independence, we have that $\rho_{Y,X_i} = [0.70, \quad 0.35, \quad 0.35]$. The partial dependence functions are: $h_1(x_1) = 2x_1$, $h_2(x_2) = x_2 + 1$, $h_3(x_3) = x_3 + 1$. We also have $r_i(x_i) = h_i(x_i)$, because inputs are independent. Note that in this example the input output mapping is symmetric in X_2 and X_3 and so are the indicators $\rho_{YX_2} = \rho_{YX_3}$, and $h_2(s) = h_3(s)$. Suppose that new data lead to a reassessment of the input distributions imposing a negative correlation $\rho_{1,2} = -0.75$ between X_1 and X_2, while the other input pairs remain independent. The correlation coefficients become $\rho_{YX_i} = [0.53, -0.212, 0.424]$. Thus, ρ_{YX_2} is no longer equal to ρ_{YX_3} and differs from ρ_{YX_1} in sign.

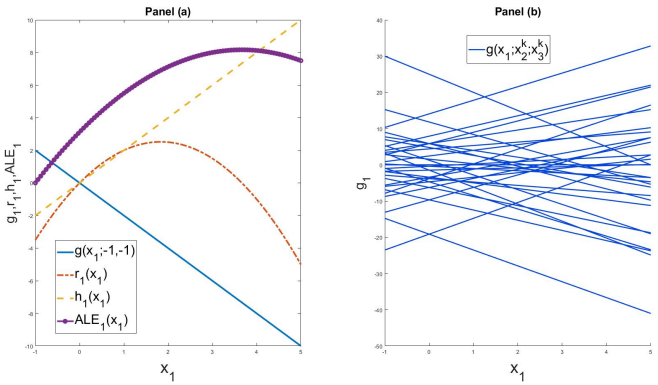


Figure 3: Panel a): Various trend indicators for Example 12 in the dependent input case. Panel b): ICE plots with one-way sensitivity functions at $k = 1, 2, \ldots, 30$ different baseline points (x_2^k, x_3^k).

Figure 3 reports the graphs of $g(x_1; -1, -1)$, $r_1(x_1)$, $h_1(x_1)$ and $ALE_1(x_1)$ in the dependent

input case. We observe that $r_1(x_1)$ is non-linear, which is due to the exponential dependence brought up by the normal density function.[1] Similarly, $ALE_1(x_1)$ is non-linear. Conversely, $h_1(x_1)$ and $g(x_i; -1, -1)$ are linear, consistently with the multilinear dependence of g on X_1. The fact that $g(x_1; -1, -1)$ is decreasing while $h_1(x_1)$ is increasing and that $ALE_1(x_1)$ changes sign indicate that the input-output mapping is non-monotonic. The one-way sensitivity functions in the ICE plot of Panel b) confirm this insight.

3.3 Convexity-consistency

Analysts frequently study the convexity of a response. This information is particularly relevant in optimization problems. To address convexity-consistency, we follow a logic similar to the one adopted in dealing with monotonicity consistency. We will therefore focus on convexity, but parallel results hold for concavity.

Let $g : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X}$ a convex set. Consider $\mathbf{x}^1$ and $\mathbf{x}^2 \in \mathcal{X}$ and the segment joining them, $\{\mathbf{z} : \mathbf{z} = \alpha\mathbf{x}^1 + (1-\alpha)\mathbf{x}^2\} \subset \mathcal{X}$, with $\alpha \in [0, 1]$. Then, a function $g : \mathcal{X}$ is convex in its domain if

$$g(\alpha\mathbf{x}^1 + (1-\alpha)\mathbf{x}^2) \leq \alpha g(\mathbf{x}^1) + (1-\alpha)g(\mathbf{x}^2), \tag{17}$$

for all $\mathbf{x}^1$ and $\mathbf{x}^2$ in $\mathcal{X}$ and all $\alpha \in [0, 1]$.

We say that a trend indicator is *convexity-consistent* if, given that g is convex with respect to X_i, the indicator exhibits such convexity.

Proposition 13. *Given any $(\mathcal{X}, \mathbb{B}(\mathcal{X}), \mathbb{P}_{\mathbf{X}})$, $g : \mathcal{X} \mapsto \mathbb{R}$, one-way sensitivity functions are convexity-consistent indicators. Moreover, if g is twice-differentiable in $\mathcal{X}$, then Hessians are convexity-consistent.*

Regression functions $r_i(x_i)$ are not convexity consistent in general, but they become convexity consistent under independence. Partial dependence functions are convexity-consistent even if inputs are dependent.

Proposition 14. *Under the assumptions of Proposition 6, if g is convex, then $h_i(x_i)$ is convex.*

Similarly to the case of monotonicity, the convexity of all partial dependence functions does not imply the overall convexity of g. However, the following holds.

Corollary 15. *Under the assumptions of Proposition 6, if g is additively separable in x_i then it is convex in x_i if and only if $h_i(x_i)$ is convex in x_i. If g is multiplicatively separable in x_i and $\int_{\mathcal{X}_{-i}} g_{-i}^{mul}(\mathbf{x}_{-i}) dF_{\mathbf{X}_{-i}}(\mathbf{x}_{-i}) > 0$ then $h_i(x_i)$ is convex in x_i if and only if g is convex in x_i.*

The plots in Figure 2 show that univariate ALE functions are not consistent with the univariate convexity of g. However, note that univariate ALE functions contain first order differential information (they deliver similar information as gradients); thus, they are not the appropriate tools to look at when convexity is concerned. Instead, convexity information is yielded by second order ALE functions given by:

$$ALE_{i,j}(x_i, x_j) = \int_{x_{\min,i}}^{x_i} \int_{x_{\min,j}}^{x_j} \mathbb{E}[g''_{i,j}(\mathbf{X}) | X_i = t, X_j = z] dt dz.$$

We refer to Apley and Zhu (2020) for additional details on second order ALE plots.

[1] Specifically,

$$r_1(x_1) = \mathbb{E}[g(\mathbf{X}_{-i}) | X_i = x_1] = \frac{x_1}{10.42} \exp\left(\frac{-2.285\,7x_1^2 + 8.0x_1}{2}\right) q(x_1),$$

where $q(x_1) = \int\int (x_2 + x_3) \exp(0.5(3.4x_1x_2 + 2.3x_2^2 - 8.0x_2 + x_3^2 - 2.0x_3 + 9.0)) dx_2 dx_3$.

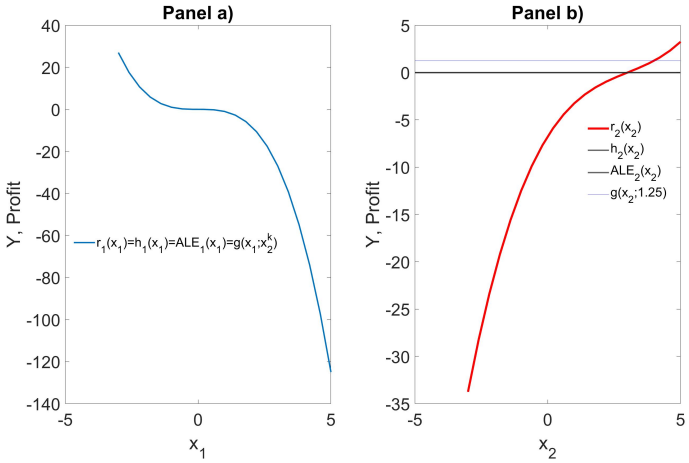


Figure 4: Trend analysis for the input-output mapping in Example 16.

4 Consistency and Intervention

It is important to distinguish whether a trend analysis is performed to support an intervention or a forecast. In an intervention context an analyst answers questions of the type: if we can act on a variable, would profit increase or decrease? In a forecasting context, the analyst is interested in predicting what the profit will be after she receives new information about the variables. Consistency is relevant in the intervention context, because it is important that the trend indicator preserves the original behavior of the simulator.

Example 16. In a given decision problem, X_1 and X_2 are jointly normal, with $\mu_1 = \mu_2 = \sigma_1 = \sigma_2 = 1$ and $\rho_{1,2} = -0.5$. The profit depends only on X_1 through $Y = g(\mathbf{X}) = -X_1^3$, and suppose that the model is not known analytically but encoded in a black box. The analyst is asked whether it is possible to act on X_2 to increase profit, as acting on X_1 comes at a high cost. The analyst chooses a conditional regression function to display the behavior of Y as a function of X_1 and X_2 obtaining the graph reported in Panel b) of Figure 4. By the regression function the analyst would recommend to increase X_2, because in expectation such an increase leads to an increase in profit (this is the interpretation of the graph of $r_2(x_2)$). However, because g depends functionally on X_1 only, acting on X_2 will not result in any increase in profit. This is informed by the ICE [2], partial dependence, or ALE plots in Panel b), which are flat lines. Note the difference with respect to getting to know that X_2 is going to increase tomorrow. In that case, because we know that X_2 is negatively correlated with X_1, we would expect that X_1 would decrease. Then, the analyst would have to ask what happens to the profit if X_1 decreases performing a trend analysis on X_1. Panel a) in Figure 4 shows that $r_1(x_1)$, $h_1(x_1)$, $ALE_1(x_1)$ and any one-way sensitivity function informs us that the profit is decreasing in X_1 (recall that these indicators are monotonicity consistent). Thus, we can expect an increase in profit, which is however caused by the decrease in X_1 and not to the increase in X_2.

5 A Discrepancy and a Flatness Index

We have seen that one-way sensitivity functions and partial dependence functions are monotonicity as well as convexity-consistent indicators, even when inputs are correlated. Goldstein et al. (2015) propose to overlay partial dependence plots with the graphs of multiple one-way sensitivity functions in a graphical representation called individual conditional expectation (ICE) plot.[3] This

[2] We displayed one representative one-way sensitivity function, as they are all constant.
[3] One-way sensitivity functions $g(x_i; \mathbf{x}_i^s)$ are called individual conditional relationships in Goldstein et al. (2015).

allows analysts to visualize local information (one-way sensitivity plots) and global information on trend (partial dependence plots). Example 12 showcases an instance in which these two trend indicators provide very different indications. To quantify this difference, we propose a discrepancy index. The ideal index would take value zero when indications coming from one-way and partial dependence functions coincide and would take value 1 when information is always opposite. To introduce such index, consider the random ratio

$$W_i = \frac{g(X_i + \Delta; \mathbf{X}_{-i}) - g(\mathbf{X})}{h_i(X_i + \Delta) - h_i(X_i)}, \tag{18}$$

with $X_i + \Delta, \mathbf{X}_{-i}, \mathbf{X}$ belonging to $\mathcal{X}$ and with $h_i(X_i + \Delta) - h_i(X_i) \neq 0$ (The case $h_i(X_i + \Delta) - h_i(X_i) = 0$ is investigated later on). The ratio W_i in turn equals the ratio between $\frac{g(X_i+\Delta;\mathbf{X}_{-i})-g(\mathbf{X})}{\Delta}$, the Newton quotient centered at $\mathbf{X}$ and computed along the one-way sensitivity function $g(X_i; \mathbf{X}_{-i})$, and $\frac{h_i(X_i+\Delta)-h_i(X_i)}{\Delta}$, the Newton quotient computed along the partial dependence function $h_i(X_i)$. If these quotients have the same sign, then the one-way sensitivity function $g(X_i+\Delta; \mathbf{X}_{-i})$ and the partial dependence function $h_i(X_i)$ provide equivalent indications. That is, if the sign of W_i is positive, then the indication is concordant, if the sign is negative, the indication is discordant. We can then define the following *discrepancy index*,

$$I_i^{\text{Discr}} = \frac{1 - \mathbb{E}[sgn(W_i)]}{2}. \tag{19}$$

Proposition 17. *We have that* $0 \leq I_i^{Discr} \leq 1$*, for every* g*,* $(\mathcal{X},\ \mathcal{B}(\mathcal{X}), \mathbb{P}_{\mathbf{X}})$ *and* $i = 1, 2, \ldots, n$*. Moreover, if* g *is additive or separately monotonic in* X_i*, then* $I_i = 0$*.*

Thus, if the indications of partial dependence functions and one-way sensitivity functions are always consistent, W_i is always positive and $I_i^{\text{Discr}} = 0$ (no discrepancy). Small values of I_i^{Discr} indicate high agreement between one-way and partial dependence functions. Conversely, a value $I_i^{\text{Discr}} = 1$ suggests that the indications are different with probability 1. The second part of Proposition 17 addresses two sufficient conditions under which $g(x_i; \mathbf{x}_{-i})$ always agree: the case in which g is additive or separately monotonic. As a consequence, values of I_i^{Discr} different from zero imply that g cannot be additive or monotonic.

Example 18. Consider the mappings $a(\mathbf{X}) = X_1^2 + X_2^2 + X_3^2$, and $m(\mathbf{X}) = X_1X_2X_3$, with X_1, X_2, and X_3 uniformly and independently distributed on $\mathcal{X} = [0, 1]^3$. Then, both mappings are separately increasing in X_i. The partial dependence functions are:

$$h_i^a(x_i) = x_i^2 + \frac{2}{3} \quad \text{and} \quad h_i^m(x_i) = \frac{x_i}{4}, \quad \text{for } i = 1, 2, 3,$$

and report exactly the marginal behavior of the two input-output mappings. We have:

1. For the additive mapping,

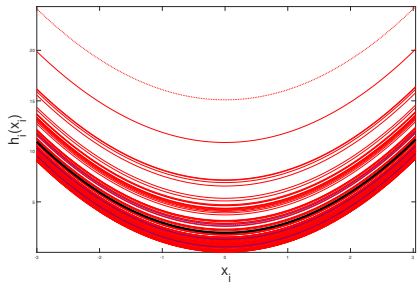
$$W_i = \frac{a(x_1 + \Delta, X_2, X_3) - a(\mathbf{x})}{h_1(x_1 + \Delta) - h_1(x_1)} = \frac{\Delta\,(2x_1 + \Delta)}{\Delta\,(2x_1 + \Delta)} = 1.$$

   Thus, one-way sensitivity functions and partial dependence functions deliver consistent information at any X_1 and $X_1 + \Delta$ in the domain, leading to $I_i^{\text{Discr}} = 0$.

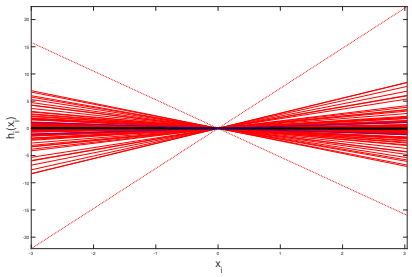
2. For the multiplicative mapping $m(\mathbf{X})$,

$$W_1 = \frac{m(x_1 + \Delta, X_2, X_3) - m(\mathbf{X})}{m_1(x_1 + \Delta) - m_1(x_1)} = \frac{X_2X_3}{\mathbb{E}[X_2X_3]}.$$

   On $[0, 1]^3$, $m(\mathbf{X})$ is separately monotonic in any of its variables and thus $I_i^{\text{Discr}} = 0$. Note also that on this domain $\mathbb{E}[X_2X_3]$ is positive and X_2 and X_3 are positive with probability 1.



(a) ICE plots for $a(\mathbf{X})$ in Example 19. In this case, $I_i^{\text{Discr}} = 0$, for $i = 1, 2, 3$.



(b) ICE plots (red) for $m(\mathbf{X})$ in Example 19. In this case, $I_i^{\text{Discr}} = 1$, for $i = 1, 2, 3$.

Figure 5: ICE (red lines) and Partial dependence plots (black bolded) for the functions in Example 19.

3. For the mapping in Example 12, we have that

$$W_1 = \frac{((x_1 + \Delta)X_2 + (x_1 + \Delta)X_3) - (X_1X_2 + X_1X_3)}{2(x_1 + \Delta) - 2x_1} = \frac{1}{2}(X_2 + X_3). \tag{20}$$

The input-output mapping is neither monotonic nor additive. The sign of W_1 depends on the point at which we are calculating the one-way sensitivity functions. Thus, $g(x_i; \mathbf{X}_{-i})$ and $h_i(x_i)$ do not always deliver concordant information for this mapping.

Note that for a separately multiplicative mapping, we have $W_i = \frac{\prod_{j=1, j\neq i}^{n} X_j}{\mathbb{E}[\prod_{j=1, j\neq i}^{n} X_j]}$. Thus, W_i is always positive if at any point $\mathbf{X}_{-i}$, the product $\prod_{j=1, j\neq i}^{n} X_j$ is concordant with $\mathbb{E}[\prod_{j=1, j\neq i}^{n} X_j]$. This is the case when the mapping is monotonic.

Example 19. Consider the additive and multiplicative functions in Example 18, but now let X_1, X_2 and X_3 independently and standard normally distributed. The respective ICE plots are given in Figures 5a and 8b. Each plot reports 100 one way sensitivity functions. For the additive function the estimated I_i^{Discr} is null for all variables. Note that we have concordance even if the mapping is not separately monotonic. In contrast, the multiplicative function $m(\mathbf{x})$ in Example 19 registers $I_i^{\text{Discr}} = 1$. In this case, the input-output mapping is not monotonic, one-way sensitivity functions are linear, with both a positive and negative slope, and the resulting partial dependence functions are identically null.

Example 19 showcases two extreme situations. In Figure 5a, the average information provided by partial dependence functions perfectly summarizes the local information of one-way sensitivity functions. In Figure 8b, partial dependence functions are not informative, and one needs to resort to one-way sensitivity functions to understand the dependence of g on X_2. This particular case is the result of a well known effect registered when the conditional or marginal expectation of g does not change as X_i varies, while other statistical properties do vary with X_i. Let us call it the *null expectation effect*. Note that the nullity of $h_i(x_i)$ makes the denominator in Equation (18) vanish so that I_i^{Discr} is not defined. We excluded this case in the definition of I_i^{Discr}.

In order to account for the null conditional expectation effect, we introduce a flatness index defined as follows. Consider an ordered sequence of realizations of X_i, $x_i^1, x_i^2 ..., x_i^k$ (this sequence can be obtained by sorting all realizations of X_i or it can just be an ordered subset), with $k = 1, 2, \ldots, R$ and $R \leq N$. Also, introduce a counter initialized at 0.

Then, consider the difference $|h_i(x_i^{k+1}) - h_i(x_i^k)|$. If this difference is zero, we add 1 to the counter; conversely, we proceed to the next k. The ratio between the number of times at which

the flatness condition is encountered and R (the total number of times it is computed) is then an indicator of flatness. In fact, if a function is constant the final value of the counter would be equal to R and the index would be equal to unity (complete flatness). Conversely, if a function is strictly monotonic, the counter would be null and the flatness index would equal zero.[4] However, this intuition needs to be slightly modified to account for numerical stability. That is, we may register a very small but still not null value of $|h_i(x_i^{k+1}) - h_i(x_i^k)|$ because of numerical noise (approximations and rounding). To overcome this problem, one can establish a threshold (ϵ_{flat}), such that $|h_i(x_i^{k+1}) - h_i(x_i^k)|$ is deemed null if $|h_i(x_i^{k+1}) - h_i(x_i^k)| < \epsilon_{flat}|h_i(x_i^k)|$. Then, we define a *flatness index* as:

$$I_i^{\text{Flat}} = \frac{\#\{|h_i(x_i^{k+1}) - h_i(x_i^k)| < \epsilon_{flat}|h_i(x_i^k)|\}}{R}, \tag{21}$$

where $\#\{\cdot\}$ is the counting function that returns the number of times the argument is met. This is an empirical index that complements I_i^{Discr}. In fact, if $I_i^{\text{Flat}} \simeq 1$, then the analyst has an indication that the behavior of the partial dependence function tends to be flat over the interval of interest. A high value of I_i^{Flat} together with a high value of I_i^{Discr} is indicative of the null expectation effect. That is, one-way sensitivity functions indicate a positive and negative dependence, while their average (the partial dependence function) is null and does not provide meaningful information. Conversely, the case $I_i^{\text{Discr}} = 0$ and $I_i^{\text{Flat}} > 0$ indicates an additive or monotonic input-output mapping. Finally, the case $I_i^{\text{Discr}} = 0$ and $I_i^{\text{Flat}} = 0$ indicates that g has a weak (if not absent) dependence on X_i.

6 Partial Dependence Spiderplots and Tornado Diagrams

Tornado diagrams are frequently used in decision analysis (Howard 1988, Bhattacharjya and Shachter 2008, Wang et al. 2017). For each input of interest, we consider the finite increments $x_i^0 + \Delta_i \in \mathcal{X}$. The sensitivity measures of a Tornado diagram are the finite differences:

$$\Delta g_i^0 = g(x_i^0 + \Delta_i; \mathbf{x}_{-i}^0) - g(x_i^0; \mathbf{x}_{-i}^0). \tag{22}$$

It is easy to see that Δg_i^0 is the difference between the values of the one-way sensitivity function $g(x_i; \mathbf{x}_{-i}^0)$ evaluated at $x_i = x_i^0 + \Delta_i$ at x_i and then at x_i^0. Thus, the sensitivity measure of a Tornado diagram can be visualized by a vertical section of a spiderplot at $x_i^0 + \Delta_i$. Tornado diagrams are, by construction, deterministic tools.

In an ICE plot, we have N one-way sensitivity functions centered at $\mathbf{x}^k$, $k = 0, 2, ..., N-1$, potentially corresponding to N randomized Tornado diagrams. At each of these centres, the sensitivity measures of a Tornado diagram can then be written as

$$\Delta g_i^k = g(x_i^k + \Delta; \mathbf{x}_{-i}^k) - g(\mathbf{x}^k),\ k = 0, 2, ..., N-1.$$

To relate Tornado diagrams with partial dependence functions, we take the average and obtain:

$$\Delta \overline{g}_i^k = \int g(x_i^k + \Delta; \mathbf{x}_{-i}) dF_{\mathbf{X}_{-i}}(\mathbf{x}_{-i}) - \int g(x_i^k; \mathbf{x}_{-i}) dF_{\mathbf{X}_{-i}}(\mathbf{x}_{-i}) = h_i(x_i^k + \Delta) - h_i(x_i^k). \tag{23}$$

Thus, $\Delta \overline{g}_i^k$ is the difference of the partial dependence function $h_i(x_i)$ at $x_i^k + \Delta$ v.s. x_i^k. Note that such measure is not sensitive to input dependence. In this respect, it complements the research efforts of works such as Wang et al. (2017), in which Tornado Diagrams are extended to the case of dependent inputs using a copula approach. By plotting the partial dependence functions in a spiderplot-like graph or displaying its vertical cuts in a Tornado diagram, one obtains average information about the relative magnitude of the influence of X_i. Note that, in an ordinary spiderplot the curves always meet at the base case value of the inputs; this is not necessarily the case for a partial dependence spiderplot.

[4] A more refined version of the index could consist in taking the difference $|h_i(x_i^{k+1}) - h_i(x_i^k)|$ for all pairs of points in $(x_i^1, x_i^2 ..., x_i^k)$.

Example 20. As an example, consider the input-output mapping $g(\mathbf{x}) = 3x_1^2 + x_2^2/x_3^2$ on $[1,2]^3$. The partial dependence spiderplot is reported in Figure 6 along with two partial dependence tornado diagrams obtained with vertical snapshots at $\pm 10\%$ and $\pm 30\%$ input variations.

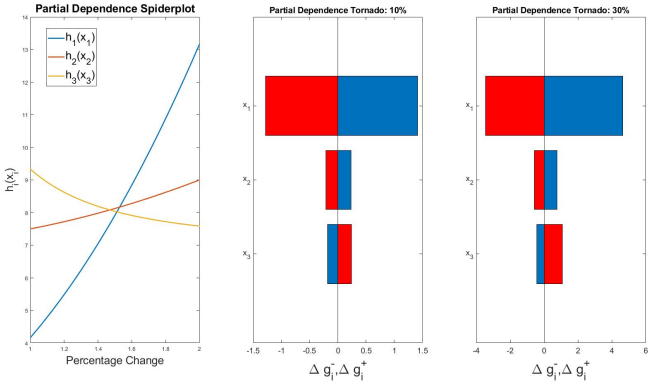


Figure 6: Partial Dependence Spiderplot (Left) and two Tornado Diagrams, obtained for parametric changes equal to $\pm 10\%$ and $\pm 30\%$ input variations.

The graphs in Figure 6 provide insights on a) direction of change: g increases in X_1 and X_2, and decreases in X_3, and, b) on the relevance of the inputs as well: X_1 is, on average, more influential than X_2 and X_3 for $\pm 10\%$ and $\pm 30\%$ variations in the inputs. However, to make insights on relative importance more robust we would recommend the complementary use of single number global indicators (Baucells and Borgonovo 2013), if computational burden allows their quantification.

7 A Case Study

This section conducts a trend analysis for a mathematical model for the COVID-19 pandemic. The simulator is a member of Susceptible Exposed Infected Recovered (SEIR) family of models (Kermack and McKendrick 1927), widely used during the COVID pandemic (Currie et al. 2020). The SEIR model we employ follows the works of Peng et al. (2020) in China and **?** in Italy, respectively. The code and data are available from `https://github.com/ECheynet/SEIR`. The model equations are

$$\dot{S} = -\alpha S - \beta \frac{S \cdot I}{N_{pop}}, \qquad \dot{E} = \beta \frac{S \cdot I}{N_{pop}} - \gamma E, \qquad \dot{I} = \gamma E - \delta I,$$
$$\dot{Q} = \delta I - \lambda Q - \kappa \cdot Q, \qquad \dot{R} = \lambda Q, \qquad \dot{D} = \kappa Q,$$
$$\dot{P} = \alpha S,$$

where $S(t)$, $E(t)$, $I(t)$, $R(t)$, $Q(t)$, $D(t)$, $P(t)$ are, respectively, the number of individuals susceptible, exposed, infected, recovered, quarantined, deceased and protected at time t. The model input names, symbols and probability measures are reported in Table 3. The output of interest is the total number of infected individuals on April 17, 2020, from the beginning of the data series, on February 24 2020.

Table 3: Input data for the SEIR model.

Input	$X_1 = \alpha$, protection rate	$X_2 = \beta$, infection rate	$X_3 = \gamma$, average latent time
Distribution	Uniform	Uniform	Uniform
Support	$\alpha^0 \pm 10\%$	$\beta^0 \pm 10\%$	$\gamma^{-1} \pm 10\%$
Input	$X_4 = \delta$, quarantine rate	$X_5 = I_0$, initially infected	X_6, intervention time
Distribution	Uniform	Discrete uniform	Discrete uniform
Support	$\delta^0 \pm 30\%$	$I_0^0 \pm 20\%$	{08-Mar-20+z}, with $z = 0, 1, ..., 7$

In **?** the model is calibrated to Lombardy data of the first months of the pandemic, in order to assess the effect of the lockdown time. In Lombardy, the intervention date was March 8, 2020. Intervention time is then assumed as a discrete random variable with 7-days as possible realizations to account for the effect of potential delays in the effectiveness of intervention measures. For the remaining inputs, the reference values α^0, β^0,$\gamma^{-1,0}$, δ^0, I_0^0 in Table 3 are set by calibrating the model to data after the intervention time in Lombardy. We assume a -0.50 correlation between intervention time and I_0: the larger the number of initially infected individuals, the sooner intervention measures will be applied. We also assume a 0.75 correlation between infection rate and number of initially infected individuals.

We propagate uncertainty in the model by generating a Monte Carlo input set from the distributions in Table 3. Correlations are incorporated using the method by Iman and Conover (1982). The simulator is then evaluated in correspondence of the input values; for each realization, the code projects one possible trajectory of the pandemic and the cumulative number of infected individuals by April 17 is registered. Partial dependence and one-way sensitivity functions are computed directly with the simulator in the loop. We present results for $N = 1,000$ simulations. The overall analysis takes about 20 minutes on a pc with 64GBRAM and 1.80GHz CPU.

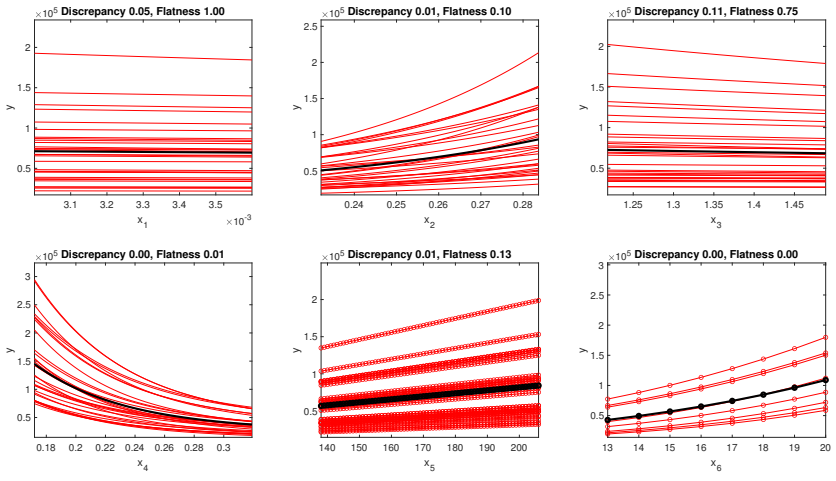


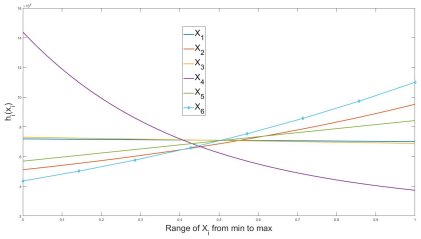
Figure 7: Trend analysis for the six inputs of the SEIR model fitted to Lombardy data in the period February-April 2020. Red lines: one-way sensitivity functions. Bolded black lines, partial dependence functions $h_i(x_i)$. Results are for $N = 1,000$ Monte Carlo simulations.

Figure 7 displays six ICE/PD plots. The first panel shows the dependence of the cumulative number of infected individuals on the protection rate, X_1. We register a low discrepancy index ($I_1^{\text{Discr}} = 0.05$) and a flatness index close to 1, indicating that the output is weakly influenced by this input. In contrast, the infection rate, X_2, has a positive and non-negligible effect on the

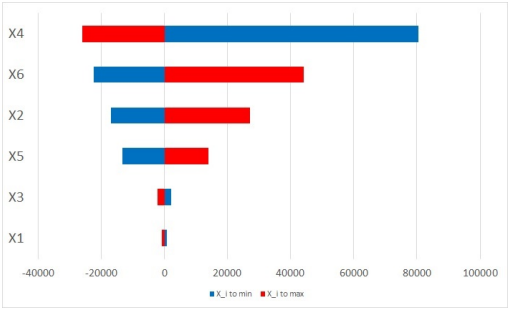
simulation output, as one should expect. The average latent time, X_3, has a negative effect on the simulation output, however the high flatness index indicates a possibly weak effect. The quarantine rate, X_4, has a negative effect on the simulation output. Indeed, the higher the efficacy provided by the quarantine, the lower the total number of infected individuals. For this input, we register a null effect of the discrepancy index, as all one-way sensitivity functions are decreasing in X_4. This is a strong indication that the simulator output might be monotonically decreasing in X_4. The initial number of infected individuals, $X_5 = I_0$, has a non-negligible and positive effect on the total number of infected individuals, in accordance with intuition. The practically null value of the discrepancy index shows a monotonic behavior of the output as a function of I_0. Also, we register no crossings among one-way sensitivity functions, suggesting that the output depends monotonically and additively on this input.

The last panel shows the dependence on the total number of infected individuals on the intervention time, X_6. We register a strong and increasing dependence. Both the discrepancy and the flatness indices are equal to zero, indicating a monotonic behavior of the total number of effected individuals as a function of the intervention time; also, we register no crossings in the one-way sensitivity functions, which is an indication of a potentially additive behavior of the code as a function of this input. This panel also suggests that an increase in waiting time before intervention would cause the total number of infected individuals to increase rapidly: in expectation a 7 days delay amounts at an expected 250% increase in the number of infected individuals.

Figure 8 reports the partial dependence spiderplot and Tornado diagram for this case study.



(a) Partial Dependence Spiderplot.



(b) Partial Dependence Tornado Diagram.

Figure 8: Partial Dependence Spiderplot and Tornado Diagram for the SEIR model with correlated inputs.

Overlaying the partial dependence functions in a spiderplot allows a visual comparison of the relative magnitude of the impact caused by each input on the output. Panel *a)* evidences that the quarantine rate, X_4, has a stronger impact than the intervention time X_6 and than the infection rate X_2, with the remaining inputs playing a minor role. This indication is confirmed by the partial dependence Tornado diagram in Panel *b)*, in which the variation of the output is considered setting the inputs at their extreme values. We checked these results calculating the Kuiper-based global sensitivity measure of Baucells and Borgonovo (2013). The ranking of the inputs coincides with the one in the partial dependence Tornado diagram in Panel b).

Overall, these results confirm that for Lombardy a weak or delayed applications of the quarantine and lockdown measures would have lead to an even greater increase in the number of infected individuals. It is not unreasonable to expect that such increased number would have caused additional difficulties in what already was a dramatic situation for the region during the spring of 2020.

8 Conclusions

In several applications decision-makers are concerned with the trend of a quantity of interest as a function of some independent variable. The selection of the appropriate indicator among the several available is not innocuous. The notion of monotonicity/convexity consistency plays a key role in guiding this choice. Our investigation shows the following (Table 1, fifth column): one-way and partial dependence functions are monotonicity consistent. Gradients and ALE plots are consistent in sign, but with a less straightforward graphical interpretation. The same drawback applies to regression plots, as well as to correlation coefficients (correlation coefficients are also well known to hide a lot of information for non-linear and complex input-output mappings). Thus, ICE plots, which are formed by consistent one-way indicators, become natural candidates for visualizing results to stakeholders. Partial dependence functions, which are the average of ICE plots, provide a valuable summary, with the advantage of not depending on some arbitrary anchor.

There is a growing debate regarding trend indicators in use (Hooker and Mentch 2019, Apley and Zhu 2020, Molnar et al. 2020). We have introduced an analytical perspective for the study of the properties of these indicators. Our results show that partial dependence functions are monotonicity and convexity-consistent even under input correlations. Then, potential graphical distortions are not due to the definition of partial dependence functions themselves, but to the fact that the machine learning tool fails in extrapolation (Apley and Zhu 2020, Molnar et al. 2020). These inconsistencies should not appear when the analyst has built a simulator, as there are no issues of extrapolation because g is either known directly or through the corresponding computer program. Conversely, detecting these inconsistencies could have a debugging effect if such distortions were observed for simulations where monotonicity/convexity are known.

At the same time, we have seen that regression lines are not monotonicity consistent under dependence. The missed consistency is due to the additional dependence on the variable in question generated by the conditional expectation operation. This result is in line with the observations in Apley and Zhu (2020) and suggests that the missed consistency is a further manifestation of the omitted-variable bias. Discrepancies between indicators are caused by input-input correlations and it matters whether pairwise correlation is positive or negative. When trend indicators disagree the analyst should really stop and investigate further such input dependences in the problem at hand.

Based on our results, we recommend a combination of ICE and partial dependence plots reinforced by the discrepancy and flatness indices. Computational burden can play a role in the selection and we report additional discussion in Section 11. These graphs unite a popular tool in the management sciences (one-way sensitivity functions) with a popular tool in machine learning (partial dependence functions). Partial dependence functions can also be displayed in spiderplot or Tornado forms. The monotonicity/convexity consistency of these graphs reassures interpretation and the indices of discrepancy and flatness allow the analyst to gain further insights.

References

Apley DW, Zhu J (2020) Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 82(4):1059–1086.

Baucells M, Borgonovo E (2013) Invariant Probabilistic Sensitivity Analysis. *Management Science* 59(11):2536–2549.

Beccacece F, Borgonovo E (2011) Functional ANOVA, Ultramodularity and Monotonicity: Applications in Multiattribute Utility Theory. *European Journal of Operational Research* 210(2):326–335, ISSN 0377-2217.

Bertsimas D, O'Hair A (2013) Learning preferences under noise and loss aversion: An optimization approach. *Operations Research* 61(5):1190–1199.

Bhattacharjya D, Shachter R (2010) Three new sensitivity analysis methods for influence diagrams. *Proceedings of the Twenty-Sixth Annual Conference on Uncertainty in Artificial Intelligence (UAI-10)*, 56–64 (Corvallis, OR: AUAI Press).

Bhattacharjya D, Shachter R (2012) Formulating Asymmetric Decision Problems as Decision Circuits. *Decision Analysis* 9(2):138–145.

Bhattacharjya D, Shachter RD (2008) Sensitivity analysis in decision circuits. *In McAllester D., and Myllymaki P., editors, Proc. of 24th UAI, Finland, Helsinki AUAI Press* 32–42.

Bordley RF, Kirkwood CW (2004) Multiattribute Preference Analysis with Performance Targets. *Operations Research* 52(6):823–835.

Borgonovo E, Smith C (2011) A Study of Interactions in the Risk Assessment of Complex Engineering Systems: An Application to Space PSA. *Operations Research* 59(6):1461–1476.

Castillo E, Gutierrez J, Hadi A (1997) Sensitivity analysis in discrete Bayesian networks. *IEEE Transactions on Systems, Man, and Cybernetics* 27:412–423.

Cecchini M, Aytug H, Koehler G, Pathak P (2010) Detecting management fraud in public companies. *Management Science* 56(7):1146–1160.

Chambers JM, Hastie TJ (1992) *Statistical Models in S* (Pacific Grove, CA: Wadsworth & Brooks/Cole).

Clemen RT (1997) *Making Hard Decisions: An Introduction to Decision Analysis* (Duxbury Press, CA, USA.), ISBN 978-0534260347.

Currie CSM, Fowler JW, Kotiadis K, Monks T, Onggo BSA, Robertson DA, Tako AA (2020) How simulation modelling can help reduce the impact of COVID-19. *Journal of Simulation* 14(2):83–97.

Darwiche A (2003) A differential approach to Inference in Bayesian Networks. *Journal of the ACM* 50(3):280–305.

Eschenbach TG (1992) Spiderplots versus Tornado Diagrams for Sensitivity Analysis. *Interfaces* 22(6):40–46.

Foldes S, Hammer PL (2005) Submodularity, Supermodularity, and Higher-Order Monotonicities of Pseudo-Boolean Functions. *Mathematics of Operations Research* 30(2):453–451.

Friedman J, Hastie T, Tibshirani R, Others (2001) *The elements of statistical learning*, volume 1 (Springer series in statistics New York).

Friedman JH, Popescu BE (2008) Predictive learning via rule ensembles. *The Annals of Applied Statistics* 2(3):916–954.

Goldstein A, Kapelner A, Bleich J, Pitkin E (2015) Peeking Inside the Black Box: Visualizing Statistical Learning With Plots of Individual Conditional Expectation. *Journal of Computational and Graphical Statistics* 24(1):44–65.

Griewank A (2000) *Evaluating Derivatives, Principles and Techniques of Algorithmic Differentiation*, volume 19 of *Frontiers in Appl. Math.* (Philadelphia: SIAM).

Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D (2018) A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys* 51(5):1–42.

Gumbel EJ (1960) Bivariate Exponential Distributions. *Journal of the American Statistical Association* 55(292):698–707.

Hastie TJ, Tibshirani R, Friedman JH (2009) *The Elements of Statistical Learning* (New York, NY, USA: Springer-Verlag), second edi edition, ISBN 978-0-387-84857-0.

Hong LJ, Nelson BL (2006) Discrete Optimization via Simulation Using COMPASS. *Operations Research* 54(1):115–129.

Hooker G (2007) Generalized Functional ANOVA Diagnostics for High Dimensional Functions of Dependent Variables. *Journal of Computational and Graphical Statistics* 16(3):709–732.

Hooker G, Mentch L (2019) Please Stop Permuting Features An Explanation and Alternatives. *ArXiv* 1905(03151):1–15.

Howard RA (1988) Decision Analysis: Practice and Promise. *Management Science* 34(6):679–695.

Hu Z, Cao J, Hong LJ (2012) Robust Simulation of Global Warming Policies Using the DICE Model. *Management Science* 58(12):2190–2206.

Iman RL, Conover WJ (1982) A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics - Simulation and Computation* 11(3):311–334.

Kermack W, McKendrick A (1927) A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society A* 115(772):700–721.

Kim Y, Street WN, Russell GJ, Menczer F (2005) Customer targeting: A neural network approach guided by genetic algorithms. *Management Science* 51(2):264–276.

Kleijnen JPC (2017) Regression and Kriging metamodels with their experimental designs in simulation: A review. *European Journal of Operational Research* 256(1):1–16.

Levy H, Guttman I, Tkatch I (2001) Regression, correlation, and the time interval: Additive-multiplicative framework. *Management Science* 47(8):1150–1159.

Molnar C (2018) *Interpretable Machine Learning - A Guide for Making Black Box Models Explainable.*

Molnar C, Koenig G, Bischl B, Casalicchio G (2020) Model-agnostic Feature Importance and Effects with Dependent Features: A Conditional Subgroup Approach. *ArXiv* :2006.0462(v1):1–20.

Nelson B (2013) *Foundations and Methods of Stochastic Simulation* (Springer US).

Pearson K (1895) Notes on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London* 58:240–242.

Peng L, Yang W, Zhang D, Zhuge C, Hong L (2020) Epidemic analysis of COVID-19 in China by dynamical modeling. *MedArXiv* URL `http://arxiv.org/abs/2002.06563`.

Rabitz H (1989) Systems analysis at the molecular scale. *Science* 246:221–226.

Rahmandad H, Sterman J (2008) Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science* 54(5):998–1014.

van der Gaag L, Renooij S, Coupe V (2007) Sensitivity Analysis of Probabilistic Networks. *Advances in Probabilistic Graphical Models, Studies in Fuzziness and Soft Computing Volume 214, 2007, pp 103-124* 214:103–124.

Wang T, Dyer JS, Hahn W (2017) Sensitivity Analysis of Decision Making under Dependent Uncertainties Using Copulas. *EURO Journal on Decision Processes* 5(1):117–139.

Zhao Q, Hastie T (2019) Causal Interpretations of Black-Box Models. *Journal of Business & Economic Statistics* 0(0):1–10.

9 Proofs

Proof. [Proof of Proposition 2] Regarding one-way sensitivity functions, item 1 in Definition 1 implies that if g is monotonic in X_i then $g(x_i)$ is monotonic, irrespectively of where $\mathbf{x}_{-i}$ is fixed and of the probability distribution assigned on $(\mathcal{X}, \mathbb{B}(\mathcal{X}))$. Thus, one-way sensitivity functions are trend-preserving. In fact, if g is differentiable and increasing (decreasing) in x_i, then the sign of the derivatives is always positive (negative). Thus, if the analyst considers the sign of derivatives as an indicator of trend, this sign correctly informs about the separate monotonicity of g, if any. □

Proof. [Proof of Proposition 5] Given the above assumptions, the monotonicity of $r_i(x_i)$ can be written as $r_i(x_i + h) - r_i(x_i) \geq 0$ for every allowed x_i and $h > 0$. This condition is then equivalent to $\mathbb{E}[Y|X = x_i + h] - \mathbb{E}[Y|X = x_i] \geq 0$. The left hand side of the above expression is:

$$\int ... \int g(x_i + h, \mathbf{X}_{-i}) p(\mathbf{X}_{-i}|X_i = x_i + h) d\mathbf{X}_i - \int ... \int g(x_i, \mathbf{X}_{-i}) p(\mathbf{X}_{-i}|X_i = x_i) d\mathbf{X}_i$$

which can be equivalently rewritten as

$$\int ... \int [g(x_i + h, \mathbf{X}_{-i}) p(\mathbf{X}_{-i}|X_i = x_i + h) - g(x_i, \mathbf{X}_{-i}) p(\mathbf{X}_{-i}|X_i = x_i)] d\mathbf{X}_i$$

By the monotonicity of the integral, a sufficient condition for this quantity to be positive is

$$g(x_i + h, \mathbf{X}_{-i}) p(\mathbf{X}_{-i}|X_i = x_i + h) - g(x_i, \mathbf{X}_{-i}) p(\mathbf{X}_{-i}|X_i = x_i) \geq 0 \tag{24}$$

for all values of $(x_i + h, \mathbf{X}_{-i}), (x_i, \mathbf{X}_{-i})$ in $\mathcal{X}$. Note that if g is non-decreasing in X_i, then $g(x_i + h, \mathbf{X}_{-i}) > g(x_i, \mathbf{X}_{-i})$. Thus, supposing $g(x_i, \mathbf{X}_{-i}) \neq 0$, we have

$$g(x_i + h, \mathbf{X}_{-i}) p(\mathbf{X}_{-i}|X_i = x_i + h) \geq g(x_i, \mathbf{X}_{-i}) p(\mathbf{X}_{-i}|X_i = x_i) \tag{25}$$

if

$$\frac{g(x_i+h,\mathbf{X}_{-i})}{g(x_i,\mathbf{X}_{-i})} \geq \frac{p(\mathbf{X}_{-i}|X_i=x_i)}{p(\mathbf{X}_{-i}|X_i=x_i+h)}. \tag{26}$$

Thus, the monotonicity of $r_i(x_i)$ is ensured if the conditional probability distribution respects the ratio

$$\frac{p(\mathbf{X}_{-i}|X_i=x_i+h)}{p(\mathbf{X}_{-i}|X_i=x_i)} \leq \frac{g(x_i,\mathbf{X}_{-i})}{g(x_i+h,\mathbf{X}_{-i})}. \tag{27}$$

□

Proof. [Proof of Proposition 6] Let us consider the inequality in (5). Then, by multiplying both sides by $f_{\mathbf{X}_{-i}}(\mathbf{x}_{-i})$ we can write

$$g(x_i+t_i;\mathbf{x}_{-i})f_{\mathbf{X}_{-i}}(\mathbf{x}_{-i}) \overset{(\leq)}{\geq} g(\mathbf{x})f_{\mathbf{X}_{-i}}(\mathbf{x}_{-i})$$

for all $\mathbf{x}$ and $t_i > 0$. Then, because $g(x_i;\mathbf{x}_{-i})$ is integrable for all $x_i \in \mathcal{X}_i$, by the monotonicity of the integral, we find

$$\int_{\mathcal{X}_{-i}} g(x_i+t_i;\mathbf{x}_{-i})f_{\mathbf{X}_{-i}}(\mathbf{x}_{-i})d\mathbf{x}_{-i} \overset{(\leq)}{\geq} \int_{\mathcal{X}_{-i}} g(x_i;\mathbf{x}_{-i})f_{\mathbf{X}_{-i}}(\mathbf{x}_{-i})d\mathbf{x}_{-i},$$

whence

$$h_i(x_i+t_i) \overset{(\leq)}{\geq} h_i(x_i),$$

which is the definition of univariate monotonicity. □

Proof. [Proof of Corollary 7] If $g = a_i(x_i) + a_{-i}(\mathbf{x}_{-i})$ or $g = m_i(x_i)m_{-i}(\mathbf{x}_{-i})$, then the partial dependence function $h_i(x_i)$ differs from $g(x_i)$ by an additive or a multiplicative constant. In fact, in the additively separable case, we have:

$$h_i(x_i) = \int_{\mathcal{X}_{-i}} (a_i(x_i) + a_{-i}(x_{-i}))f_{\mathbf{X}_{-i}}(\mathbf{x}_{-i})d\mathbf{x}_{-i} = a_i(x_i) + A_i, \tag{28}$$

where $A_i = \int_{\mathcal{X}_{-i}} a_{-i}(\mathbf{x}_{-i})f_{\mathbf{X}_{-i}}(\mathbf{x}_{-i})d\mathbf{x}_{-i}$. Then, in case X_i is increasing or decreasing, then $h_i(x_i)$ is increasing or decreasing. In case g is separately multiplicative, we have

$$h_i(x_i) = m_i(x_i)\int_{\mathcal{X}_{-i}} m_{-i}(\mathbf{x}_{-i})f_{\mathbf{X}_{-i}}(\mathbf{x}_{-i})d\mathbf{x}_{-i} = m_i(x_i)M_i, \tag{29}$$

where $M_i = \int_{\mathcal{X}_{-i}} m_{-i}(\mathbf{x}_{-i})f_{\mathbf{X}_{-i}}(\mathbf{x}_{-i})d\mathbf{x}_{-i}$. Then, $h_i(x_i)$ is proportional to $m_i(x_i)$. Thus, if M_i is positive, then $h_i(x_i)$ is monotonic with respect to X_i with the same monotonicity of g with respect to X_i (if g is increasing(decreasing) in X_i, then h_i is increasing(decreasing)). Conversely, if M_i is negative, if g is monotonic, $h_i(x_i)$ remains monotonic (increasing or decreasing), but the direction is reversed; that is if g is increasing, then $h_i(x_i)$ is decreasing and the converse. □

Proof. [Proof of Proposition 8] In Proposition 2, we have seen that partial derivatives are consistent indicators. Then, suppose g is increasing in X_i. Then, consider first order ALE functions. The definition is

$$ALE_i(x_i) = \int_{x_{\min,i}}^{x_i} \mathbb{E}[g_i'(\mathbf{X})|X_i=z_i]dz_i. \tag{30}$$

Then, for every value of X_i, $ALE_i(x_i)$ is the integral of the average of partial derivatives $g_i'(\mathbf{X})$. Then, if $g(\mathbf{X})$ is partially monotonic in X_i, $g_i'(\mathbf{X})$ will consistently have the same sign (positive or negative). Then, the conditional expectation $\mathbb{E}[g_i'(\mathbf{X})|X_i=z_i]$ is positive(negative) and, by the monotonicity of the integral, also $ALE_i(x_i)$ is positive/negative. □

Proof. [Proof of Proposition 13] Regarding one-way sensitivity functions, by equations (**??**) and (17), for all $\mathbf{x}_{-i}$ the following holds:

$$g(\alpha x_i^1 + (1-\alpha)x_i^2; \mathbf{x}_{-i}) \leq \alpha g(x_i^1; \mathbf{x}_{-i}) + (1-\alpha) g(x_i^2; \mathbf{x}_{-i}). \tag{31}$$

That is, if g is convex, all one-way sensitivity functions are convex. Therefore, one-way sensitivity functions are non-distorted measures of the convexity (if any) of g. Regarding differential sensitivity measures, the properties of convex and twice-differentiable functions reassure us that the sign of second order partial derivatives is always positive(negative) for convex functions and, thus, the sign of second order derivatives is a non-distorted measure of convexity. This holds for every assignment of $(\mathcal{X}, \mathbb{B}(\mathcal{X}), \mathbb{P}_{\mathbf{X}})$. □

Proof. [Proof of Proposition 14] In this proof it is useful to write $\mathbf{x}^2 = \mathbf{x}^1 + \mathbf{s}$. Now, let us recall the definition of $h_i(x_i)$. Letting $v_{-_i}(\mathbf{x}_{-i})$ a positive function, and provided that $\int \cdots \int_{\mathcal{X}\setminus\mathcal{X}_i} g(\mathbf{x}) v_{-_i}(\mathbf{x}_{-i}) d\mathbf{x}_{-i}$ exists for all values of x_i, we can define

$$h_i(x_i) = \int \cdots \int_{\mathcal{X}\setminus\mathcal{X}_i} g(\mathbf{x}) v_{-_i}(\mathbf{x}_{-i}) d\mathbf{x}_{-i} = \int \cdots \int_{\mathcal{X}\setminus\mathcal{X}_i} g(x_i; \mathbf{x}_{-i}) v_{-i}(\mathbf{x}_{-i}) d\mathbf{x}_{-i}, \tag{32}$$

where $v_{-_i}(\mathbf{x}_{-i}) \geq 0$ for all $\mathbf{x}_{-i} \in \mathcal{X}\setminus\mathcal{X}_i$. Then, we have

$$h_i(x_i^1) = \int \cdots \int_{\mathcal{X}\setminus\mathcal{X}_i} g(x_i^1; \mathbf{x}_{-i}) v_{-i}(\mathbf{x}_{-i}) d\mathbf{x}_{-i} \tag{33}$$

and

$$h_i(x_i^2) = \int \cdots \int_{\mathcal{X}\setminus\mathcal{X}_i} g(x_i^2; \mathbf{x}_{-i}) v_{-i}(\mathbf{x}_{-i}) d\mathbf{x}_{-i}. \tag{34}$$

Therefore, $h_i(x_i)$ is convex if for $\alpha + \beta = 1$

$$h_i(\alpha x_i^1 + \beta x_i^2) \leq \alpha h_i(x_i^1) + \beta h_i(x^2). \tag{35}$$

Then, let us go back to g. We have:

$$\begin{aligned} g(\alpha\mathbf{x}^1 + \beta\mathbf{x}^2) &= g(\alpha x_i^1 + \beta(x_i^1 + s_i); \alpha\mathbf{x}_{-i}^1 + \beta(\mathbf{x}_{-i}^1 + \mathbf{s}_{-i})) & (36)\\ &= g(\alpha x_i^1 + \beta x_i^1 + \beta s_i); \alpha\mathbf{x}_{-i}^1 + \beta\mathbf{x}_{-i}^1 + \beta\mathbf{s}_{-i}). & (37)\end{aligned}$$

Then, recalling that $\alpha + \beta = 1$, we can rewrite the above expression as

$$g(\alpha\mathbf{x}^1 + \beta\mathbf{x}^2) = g(x_i^1 + \beta s_i; \mathbf{x}_{-i}^1 + \beta\mathbf{s}_{-i}). \tag{38}$$

Now, letting $\mathbf{t} = \beta\mathbf{s}$, we have

$$g(\alpha\mathbf{x}^1 + \beta\mathbf{x}^2) = g(x_i^1 + t_i; \mathbf{x}_{-i}^1 + \mathbf{t}_{-i}). \tag{39}$$

Let us then go to u_i^v. Multiplying by $v_{-i}(\mathbf{x}_{-i})$ and integrating both sides, we have

$$\int \cdots \int_{\mathcal{X}\setminus\mathcal{X}_i} g(\alpha\mathbf{x}^1 + \beta\mathbf{x}^2) v_{-i}(\mathbf{x}_{-i}) d\mathbf{x}_{-i} = \int \cdots \int_{\mathcal{X}\setminus\mathcal{X}_i} g(x_i^1 + t_i; \mathbf{x}_{-i}^1 + \mathbf{t}_{-i}) v_{-i}(\mathbf{x}_{-i}) d\mathbf{x}_{-i}. \tag{40}$$

So, the above equality holds for all $\mathbf{t}$. Then, in particular, it holds for $\mathbf{t} = \begin{bmatrix} 0 & \dots & t_i & \dots & 0 \end{bmatrix}$. So that, we have

$$h_i(\alpha x_i^1 + \beta x_i^2) = \int \cdots \int_{\mathcal{X}\setminus\mathcal{X}_i} g(\alpha x_i^1 + \beta x_i^2; \mathbf{x}_{-i}) v_{-i}(\mathbf{x}_{-i}) d\mathbf{x}_{-i}. \tag{41}$$

Then, we have:

$$g(\alpha x_i^1 + \beta x_i^2; \mathbf{x}_{-i}) \leq \alpha g(x_i^1; \mathbf{x}_{-i}^1) + \beta g(x_i^2; \mathbf{x}_{-i}^2). \tag{42}$$

Multiplying both sides by $v_{-i}(\mathbf{x}_{-i})d\mathbf{x}_{-i}$, and integrating, by the monotonicity of the integral, we obtain

$$\begin{aligned}\int\cdots\int_{\mathcal{X}\setminus\mathcal{X}_i} g(\alpha x_i^1+\beta x_i^2;\mathbf{x}_{-i})v_{-i}(\mathbf{x}_{-i})d\mathbf{x}_{-i} \quad &\leq \quad \alpha\int\cdots\int_{\mathcal{X}\setminus\mathcal{X}_i} g(x_i^1;\mathbf{x}_{-i}^1)v_{-i}(\mathbf{x}_{-i})d\mathbf{x}_{-i}\\ &\quad +\beta\int\cdots\int_{\mathcal{X}\setminus\mathcal{X}_i} g(x_i^2;\mathbf{x}_{-i}^2)v_{-i}(\mathbf{x}_{-i})d\mathbf{x}_{-i}.\end{aligned}$$

Then, by equation (41) and by the definition of h_i, we have

$$h_i(\alpha x_i^1+\beta x_i^2)\leq \alpha h_i(x_i^1)+\beta h_i(x_i^2).$$

This is, if the multivariate function g is convex on $\mathcal{X}$, the univariate functions $h_i(x_i)$ obtained by partial integration are convex. □

Proof. [Proof of Corollary 15] The proof is similar to the proof of Corollary 7, to which we refer. In a nutshell, if g is separable then $h(x_i)$ differs from $g_i(x_i)$ only by a multiplicative or additive constant and therefore it reproduces the partial behavior of g with respect to X_i. □

Proof. [Proof of Proposition 11] Letting $\mathcal{Z}_i$ and $\mathcal{Z}_{-i}$ be the subsets of $\mathcal{Z}$ that include index i and that exclude it, respectively, then equation (8) can be rewritten as

$$g^{ML}(\mathbf{x})=a_{-i}(\mathbf{x}_{-i})x_i+b_i(\mathbf{x}_{-i}), \tag{43}$$

where $a_{-i}(\mathbf{x}_{-i})=\sum_{z\in\mathcal{Z}_i}\beta_z\prod_{s\in z,s\neq i}x_s$ and $b_{-i}(\mathbf{x}_{-i})=\sum_{z\in\mathcal{Z}_{-i}}\beta_z\prod_{s\in z}x_s$. Items 1 and 2 follow immediately by (43). For item 2, we note that

$$\begin{gathered}\overline{a}_{-i}=\int_{\mathcal{X}_{-i}}a_{-i}(\mathbf{x}_{-i})f_{\mathbf{X}_{-i}}(\mathbf{x}_{-i})d\mathbf{x}_{-i}=\int_{\mathcal{X}_{-i}}a_{-i}(\mathbf{x}_{-i})\int_{\mathcal{X}_i}f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}_{-i}dx_i=\\ \int_{\mathcal{X}_{-i}}\int_{\mathcal{X}_i}a_{-i}(\mathbf{x}_{-i})f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}_{-i}dx_i=\int_{\mathcal{X}}a_{-i}(\mathbf{x}_{-i})f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}=\mathbb{E}[a_{-i}(\mathbf{X}_{-i})],\end{gathered} \tag{44}$$

where we have used the fact that

$$f_{\mathbf{X}_{-i}}(\mathbf{x}_{-i})=\int_{\mathcal{X}_i}f_{\mathbf{X}}(\mathbf{x})dx_i. \tag{45}$$

Because $a_{-i}(\mathbf{X}_{-i})=g_i'(\mathbf{X})$, we also have:

$$r_i(x_i)=\int\ldots\int[a_{-i}(\mathbf{X}_{-i})x_i+b_{-i}(\mathbf{X}_{-i})]dF_{\mathbf{X}_{-i}}(\mathbf{X}_{-i},x_i)=\mathbb{E}[x_ig_i'(\mathbf{X})|X_i=x_i]+\mathbb{E}[b_{-i}(\mathbf{X}_{-i})|X_i=x_i], \tag{46}$$

and

$$\begin{gathered}ALE_i(x_i)=\int_{x_{\min,i}}^{x_i}\int\ldots\int g_i'(\mathbf{X})dF_{\mathbf{X}_{-i}|X_i}(\mathbf{X}_{-i};t)dt=\\ \int_{x_{\min,i}}^{x_i}\int\ldots\int a_{-i}(\mathbf{X}_{-i})dF_{\mathbf{X}_{-i}|X_i}(\mathbf{X}_{-i},t)dt=\int_{x_{\min,i}}^{x_i}\mathbb{E}[a_{-i}(\mathbf{X}_{-i})|X_i=t]dt\end{gathered} \tag{47}$$

and

$$\overline{a}_{-i}=\mathbb{E}[g_i'(\mathbf{X})]. \tag{48}$$

For the correlation coefficient, consider the numerator of ρ_{YX_i}, $Cov(Y,X_i)$. By definition of covariance, we have

$$Cov(Y,X_i)=\mathbb{E}\left[(X_i)(A_iX_i+B_i)\right]-\overline{x}_i\overline{y}, \tag{49}$$

where we have set $\overline{y}=\mathbb{E}[Y]$, $\overline{x}_i=\mathbb{E}[X_i]$, $B_i=b_{-i}(\mathbf{X}_{-i})$, $A_i=a_{-i}(\mathbf{X}_{-i})$

$$\begin{aligned}\rho_{YX_i}&=\mathbb{E}\left[(X_i)(A_iX_i+B_i)\right]-\overline{x}_i\overline{y}=\mathbb{E}\left[(X_i)(A_iX_i+B_i)\right]-\overline{x}_i\mathbb{E}[A_iX_i+B_i]\\ &=\mathbb{E}\left[A_iX_i^2+B_iX_i\right]-\overline{x}_i\mathbb{E}[A_iX_i+B_i]\\ &=\mathbb{E}\left[a_{-i}(\mathbf{X}_{-i})X_i^2+b_i(\mathbf{X}_{-i})X_i\right]-\overline{x}_i\mathbb{E}[a_{-i}(\mathbf{X}_{-i})X_i+b_i(\mathbf{X}_{-i})].\end{aligned} \tag{50}$$

Table 4: Eschenbach's distribution assignment (lower and upper limits in percentage of base-case values).

Variable	Name	Lower Limit	Base-Case Value	Upper Limit
C_{ini}	first cost	90%	\$120,000	150%
S	salvage	0%	\$20,000	150%
N	horizon of life	50%	12 years	200%
i	discount rate	60%	10%	200%
C_{OM}	annual O&M cost	80%	\$6,000	125%
R	annual revenue	60%	\$55,000	125%
N_{NR}	years without revenue	0%	1	300%
R_{FC}	loss due to competitors	0%	0.2	200%

Then, if X_i is independent of $\mathbf{X}_{-i}$, we have

$$ALE_i(x_i) = \mathbb{E}[a_{-i}(\mathbf{X}_{-i})](x_i - x_{\min,i}) = \overline{a}_{-i}(x_i - x_{\min,i}), \tag{51}$$

and

$$r_i(x_i) = x_i\mathbb{E}[g_i'(\mathbf{X})] + \mathbb{E}[b_{-i}(\mathbf{X}_{-i})] = \overline{a}_{-i}x_i + \overline{b}_i = h_i(x_i), \tag{52}$$

and

$$\begin{aligned}\rho_{YX_i} &= \mathbb{E}[a_{-i}(\mathbf{X}_{-i})]\mathbb{E}[X_i^2] + \mathbb{E}[b_i(\mathbf{X}_{-i})]\overline{x}_i - \overline{x}_i^2\mathbb{E}[a_{-i}(\mathbf{X}_{-i})] - \overline{x}_i\mathbb{E}[b_i(\mathbf{X}_{-i})] \\ &= \mathbb{E}[a_{-i}(\mathbf{X}_{-i})]\left(\mathbb{E}[X_i^2] - \overline{x}_i^2\right) = \overline{a}_{-i}\sigma_i^2.\end{aligned} \tag{53}$$

□

Proof. [Proof of Proposition 17] Item 1: The quantity $sgn(W_i)$ is a random quantity equal to 1 or -1. Thus, the expectation of such quantity cannot exceed 1 or -1. In the case $sgn(W_i) = -1$ for any realization of X_i, we would have $I_i^{\text{Discr}} = \frac{1-(-1)}{2} = 1$. Conversely, if $sgn(W_i) = 1$ for all realizations of X_i , then $I_i^{\text{Discr}} = 0$.

Item 2a: if g is separately additive ($g = \sum_{i=1}^{n} t_i(x_i)$), we have:

$$W_i = \frac{t_i(X_i + \Delta) + K_w - t_i(X_i) - K_w}{t_i(X_i + \Delta) + K_h - t_i(X_i) - K_h} = \frac{t_i(X_i + \Delta) - t_i(X_i)}{t_i(X_i + \Delta) - t_i(X_i)} = 1$$

for all X_i and, therefore $sgn(W_i) = 1$ for all X_i. Hence, $I_i^{\text{Discr}} = \frac{1-1}{2} = 0$.

If g is separately monotonic, we have the following. We have proven that one-way sensitivity functions are monotonicity and convexity preserving. Therefore, if g is increasing (decreasing) then $g(X_i + \Delta; \mathbf{X}_{-i}) - g(\mathbf{X})/\Delta$ is positive(negative) for all X_i and Δ. We have also proven that $h_i(x_i)$ is monotonicity and convexity preserving on a Cartesian domain for all . if g is increasing (decreasing) then $g(X_i + \Delta) - g(\mathbf{X})/\Delta$ is positive(negative) for all X_i and Δ. Therefore, the Newton ratios $g(X_i + \Delta; \mathbf{X}_{-i}) - g(\mathbf{X})/\Delta$ and $h_i(X_i + \Delta) - h_i(X_i)/\Delta$ have the same sign for all $\mathbf{X}_{-i}$. Consequently, $sign(W_i) = 1$ for all X_i. □

10 Eschenbach's Business Plan Model

Eschenbach (1992) discusses a simple economic model of a prospective new product, with variables listed in Table 4. Using single payment present worth $F(i, N) = (1 + i)^{-N}$ and uniform series present worth $A(i, N) = i^{-1}(1 - (1 + i)^{-N})$ the present worth of the investment is given by

$$PW \;=\; -C_{\text{ini}} + S \cdot F(i, N) - C_{\text{OM}} \cdot A(i, N) + R(1 - R_{\text{FC}}) \cdot A(i, N - N_{\text{NR}}) \cdot F(i, N_{\text{NR}})$$

assuming end-of-year for all costs except the first cost. Now assume that Table 4 describes triangular distributions with the base-case values representing the modes of the distributions, the limiting values describe its support. Then a simulation yields $\rho_{Y,R_{\text{FC}}} = 0.37$. Now, if R and R_{FC} were to exhibit $\rho_{R,R_{\text{FC}}} = -0.9$, then $\rho_{Y,R_{\text{FC}}}$ turns to $\rho_{Y,R_{\text{FC}}} = -0.05$.

11 Some Observations on Computational Cost

The computational cost of an algorithmic procedure is traditionally measured as the number of times it requires to run the simulator/machine-learning-algorithm. Regarding the indicators in Table 1, we have the following. The cost of correlation coefficients, as well as regression functions is equal to the cost of generating the Monte Carlo sample, that is N. The cost of calculating a one one-way sensitivity function is $C^{\text{one-way}} = K$, where K is the number of points of the support of $\mathcal{X}_i$ at which g is evaluated and n is the number of inputs. To clarify, suppose that the support of X_i is an interval $(x_i^{\max}, x_i^{\max})$; is this interval is partitioned in K equally spaced sub-intervals, then the cost is $C^{\text{one-way}} = K + 1$. Clearly, the greater K the higher the precision in the graphical representation of the partial behavior of g, but the greater the computational cost. If we compute a one-way sensitivity function for each input, then the cost becomes $C^{\text{one-ways}} = n \cdot K + 1$. Also, we can compute a one-way sensitivity function for each input for each of the N points at which the inputs are sampled, with an the overall computational cost $C^h = n \cdot K \cdot N$. This is the nominal cost of computing a partial dependence functions. Instead, ALE functions have a cost of $n * N$ model runs, avoiding the dependence on K (Apley and Zhu 2020). Thus, correlation coefficient and regression functions come at the lowest computational cost, followed by ALE plots and partial dependence functions. However, the true discriminant for the computational cost is the execution time of the simulator/machine learning algorithm. In fact, if the simulator running time is long relative to the budget, analysts proceed by fitting an emulator (a gradient boosting machine, a regression tree, a Gaussian process regression) over the input output mapping (Friedman et al. (2001), Friedman and Popescu (2008)), and then use the emulator to execute additional calculations. In this way, the cost reduces to N realizations of the original simulator plus the evaluations of the (fast) emulator. Or analysts can proceed in two steps (Hastie et al. 2009): first, identify the most important variables through measures of importance, then perform the trend analysis on a restricted set of inputs (the most important ones), thus reducing computational burden.

Acknowledgments

The authors thank X. Lu for providing codes and data.