RESEARCH ARTICLE　　　　　　　　　　　　　　　　　　　**Open Access**

CrossMark

Hilbert-Schmidt and Sobol sensitivity indices for static and time series Wnt signaling measurements in colorectal cancer - part A

Shriprakash Sinha

Abstract

Background: Ever since the accidental discovery of Wingless [Sharma R.P., Drosophila information service, 1973, 50, p 134], research in the field of Wnt signaling pathway has taken significant strides in wet lab experiments and various cancer clinical trials, augmented by recent developments in advanced computational modeling of the pathway. Information rich gene expression profiles reveal various aspects of the signaling pathway and help in studying different issues simultaneously. Hitherto, not many computational studies exist which incorporate the simultaneous study of these issues.

Results: This manuscript ● explores the strength of contributing factors in the signaling pathway, ● analyzes the existing causal relations among the inter/extracellular factors effecting the pathway based on prior biological knowledge and ● investigates the deviations in fold changes in the recently found prevalence of psychophysical laws working in the pathway. To achieve this goal, local and global sensitivity analysis is conducted on the (non)linear responses between the factors obtained from static and time series expression profiles using the density (Hilbert-Schmidt Information Criterion) and variance (Sobol) based sensitivity indices.

Conclusion: The results show the advantage of using density based indices over variance based indices mainly due to the former's employment of distance measures & the kernel trick via Reproducing kernel Hilbert space (RKHS) that capture nonlinear relations among various intra/extracellular factors of the pathway in a higher dimensional space. In time series data, using these indices it is now possible to observe where in time, which factors get influenced & contribute to the pathway, as changes in concentration of the other factors are made. This synergy of prior biological knowledge, sensitivity analysis & representations in higher dimensional spaces can facilitate in time based administration of target therapeutic drugs & reveal hidden biological information within colorectal cancer samples.

Keywords: Wnt pathway, Sensitivity analysis, Kernels, Psychophysical law, Colorectal cancer, Systems biology, Time series data

Background

Significance

Recently observed psychophysical laws working downstream of the Wnt pathway rely on ratio of deviations in input & absolute value of input. These deviations are crucial for observation of a phenotypic behaviour during a time interval. This work explores the influences of fold changes and deviations in fold changes in time

using density based sensitivity indices which employ kernel methods to capture nonlinear relations among the involved intra/extracellular factors. On static gene expression toy example in normal and tumor cases & time series dataset, they outperformed the variance based sensitivity indices. Synergy of prior biological knowledge, sensitivity analysis and representations in higher dimensional spaces facilitates development of time based target specific interventions at molecular level within the pathway.

i compartmentalize the manuscript into three different parts ● a short review containing the systems wide analysis of the Wnt pathway divided into introduction, problem statement and a solution to address the same via latest

Correspondence: ssinha@rtc.bt;sinha.shriprakash@yandex.com
Faculty of Maths & IT, Royal Thimphu College, Nagbiphu, 1122 Thimphu, Bhutan

sensitivity analysis methods • an extensive description of the methodology, the description of the dataset and the design of the experiments and finally • the biological findings from the system wide study of the Wnt pathway using sensitivity analysis.

A short review

Sharma's [1] accidental discovery of the Wingless played a pioneering role in the emergence of a widely expanding research field of the Wnt signaling pathway. A majority of the work has focused on issues related to • the discovery of genetic and epigenetic factors affecting the pathway ([2] & [3]), • implications of mutations in the pathway and its dominant role on cancer and other diseases [4], • investigation into the pathway's contribution towards embryo development [5], homeostasis ([6, 7]) and apoptosis [8] and • safety and feasibility of drug design for the Wnt pathway ([9–13]). Approximately forty years after the discovery, important strides have been made in the research work involving several wet lab experiments and cancer clinical trials ([9, 13]) which have been augmented by the recent developments in the various advanced computational modeling techniques of the pathway. More recent informative reviews have touched on various issues related to the different types of the Wnt signaling pathway and have stressed not only the activation of the Wnt signaling pathway via the Wnt proteins [14] but also the on the secretion mechanism that plays a major role in the initiation of the Wnt activity as a prelude [15].

With the rapid development of methods in biotechnology and the availability of of vast amounts of datasets at molecular level, there has arisen a need to understand the mechanism of these signaling pathways at a greater level. Systems biology is a field where the idea is to understand the deeper aspects of biology via various models that translate the biological problem into a computational/mathematical framework. Latest opinion on current trends in systems biology can be found in [16] and [17]. One of the earliest efforts to translate a biological problem into a mathematical framework regarding the Wnt pathway was done by [18]. The quantitative study involved the analysis of interactions among the known components of the pathway using differential equations that incorporates information regarding kinetics, synthesis/degradation and phosphorylation/dephosphorylation. Further improvements and analysis on such models have followed through in later years; for example - [19, 20]. Apart from these methods, bayesian methods have also played a crucial role in understanding certain aspects of the pathway. Recent work on parameter free methods by [21] employs bayesian methods for parameter inference from data and later use algebraic methods like matroid theory to analyse the models that fit the data, while not depending specifically on the parameters. Mixture models

approach has also been employed recently to understand aspects of the Wnt pathway via bayesian parameter estimation [22]. Besides parameter estimation methods related to differential equations, bayesian network analysis methods have also been proposed in investigate the cause-influence hypotheses among various factors affecting the pathway, by integrating heterogenous data and using concepts of d-connectivity/separability [23]. The author had the chance to provide a pedagogical perspective on the insilico analysis of [23] in the form of computer code in [24]. Not only have the signaling pathways been studied but also the some of the phenomenon that have been prevalent in the pathway have also been studied using mathematical models. The prevalence of the Weber's law (described later) has been shown in [25] using differential equation models and sensitivity analysis. Though the work has used proposed the use of sensitivity analysis, it is no evident as to which approach has been employed to conduct the study. Sensitivity analysis has been employed to investigate the pathway in [26] also for indentification of parameters. In these above cases there is involvement of the parameters which need to be studied in order to gain an insight into some mechanism of the pathway. Recently, a system wide investigation on time course data was conducted by [27] using correlational analysis. The aim of this study was to understand at a systems level how the components of the pathway was behaving in time. Even though the effort lead to the understanding of a few areas of the pathway, it has not revealed the entire analysis of the influence of each of the components at different time points as well as during the different intermittent time periods in a coherent simultaneous manner. In this manuscript, the author takes an extensive analysis of the pathway on the same time course measurements generated by [27], using the existing variance based sensitivity indices as well as the latest density based sensitivity indices. The work exploits the deeper formulation of the deviations in input in the logarithmic Bernoulli's formulations from which the Weber's law has been derived (explained later). The work in this paper investigates a systems wide study of the Wnt pathway via sensitivity analysis while using static [28] and time series [27] gene expression data retrieved from colorectal cancer samples.

Canonical Wnt signaling pathway

Before delving into the problem statement, a brief introduction to the Wnt pathway is given here. From the recent work of [23], the canonical Wnt signaling pathway is a transduction mechanism that contributes to embryo development and controls homeostatic self renewal in several tissues [4]. Somatic mutations in the pathway are known to be associated with cancer in different parts of the human body. Prominent among them is the colorectal cancer case [29]. In a succinct overview, the Wnt signaling

pathway works when the Wnt ligand gets attached to the Frizzled(*FZD*)/*LRP* coreceptor complex. *FZD* may interact with the Dishevelled (*DVL*) causing phosphorylation. It is also thought that Wnts cause phosphorylation of the *LRP* via casein kinase 1 (*CK*1) and kinase *GSK*3. These developments further lead to attraction of Axin which causes inhibition of the formation of the degradation complex. The degradation complex constitutes of *AXIN*, the *β-catenin* transportation complex *APC*, *CK*1 and *GSK*3. When the pathway is active the dissolution of the degradation complex leads to stabilization in the concentration of *β-catenin* in the cytoplasm. As *β-catenin* enters into the nucleus it displaces the *GROUCHO* and binds with transcription cell factor *TCF* thus instigating transcription of Wnt target genes. *GROUCHO* acts as lock on *TCF* and prevents the transcription of target genes which may induce cancer. In cases when the Wnt ligands are not captured by the coreceptor at the cell membrane, *AXIN* helps in formation of the degradation complex. The degradation complex phosphorylates *β-catenin* which is then recognized by *FBOX*/*WD* repeat protein *β-TRCP*. *β-TRCP* is a component of ubiquitin ligase complex that helps in ubiquitination of *β-catenin* thus marking it for degradation via the proteasome. Cartoons depicting the phenomena of Wnt being inactive and active are shown in Fig. 1a and b, respectively.

Problem statement in short

Succinctly, the endeavour is to address the following issues - • explore the strength of contributing factors in the signaling pathway, • analyse the existing causal relations among the inter/extracellular factors effecting the pathway based on prior biological knowledge and • investigate the significance of deviations in fold changes in the recently found prevalence of psychophysical laws working in the pathway in a multi-parameter setting. The issues related to • inference of hidden biological relations among the factors, that are yet to be discovered and • discovery of new causal relations using hypothesis testing, will be addressed in a subsequent manuscript. The current manuscript analyses the sensitivity indices for fold changes and deviations in fold changes in 17 different genes from a set of 74 genes as presented by [27]. An immediate followup of the manuscript is the analysis of the remaining 57 genes which happens to the part B of this manuscript.

A solution to the problem
Sensitivity analysis

In order to address the above issues, sensitivity analysis (SA) is performed on either the datasets or results obtained from biologically inspired causal models. The reason for using these tools of sensitivity analysis is that they help in observing the behaviour of the output and the importance of the contributing input factors via a robust and an easy mathematical framework. In this manuscript both local and global SA methods are used. Where appropriate, a description of the biologically inspired causal models ensues before the analysis of results from these models.

Seminal work by Russian mathematician [30] lead to development as well as employment of SA methods to

WNT - A cartoon of canonical Wnt pathway.
a - Wnt not captured at the cell receptor, i.e Wnt off
b - Wnt captured at the cell receptor, i.e Wnt on

Fig. 1 A cartoon of Wnt signaling pathway. Part (**a**) represents the destruction of *β-catenin* leading to the inactivation of the Wnt target gene. Part (**b**) represents activation of Wnt target gene

study various complex systems where it was tough to measure the contribution of various input parameters in the behaviour of the output. A recent unpublished review on the global SA methods by [31] categorically delineates these methods with the following functionality • screening for sorting influential measures ([32] method, Group screening in [33, 34], Iterated factorial design in [35], Sequential bifurcation design in [36] and [37]), • quantitative indicies for measuring the importance of contributing input factors in linear models ([38–41]) and nonlinear models ([42–57] and [58]) and • exploring the model behaviour over a range on input values ([59] and [60–62]). Iooss and Lema [31] also provide various criteria in a flowchart for adapting a method or a combination of the methods for sensitivity analysis. Figure 2 shows the general flow of the mathematical formulation for computing the indices in the variance based Sobol method. The general idea is as follows - A model could be represented as a mathematical function with a multidimensional input vector where each element of a vector is an input factor. This function needs to be defined in a unit dimensional cube. Based on ANOVA decomposition, the function can then be broken down into f_0 and summands of different dimensions, if f_0 is a constant and integral of summands with respect to their own variables is 0. This implies that orthogonality follows in between two functions of different dimensions, if at least one of the variables is not repeated. By applying these properties, it is possible to show that the function can be written into a unique expansion. Next, assuming that the function is square integrable variances can be computed. The ratio of variance of a group of input factors to the variance of the total set of input factors constitute the sensitivity index of a particular group.

Besides the above [30]'s variance based indicies, more recent developments regarding new indicies based on density, derivative and goal-oriented can be found in [63–65], respectively. In a latest development, [66] propose new class of indicies based on density ratio estimation [63] that are special cases of dependence measures. This in turn helps in exploiting measures like distance correlation [67] and Hilbert-Schmidt independence criterion [68] as new sensitivity indicies. The framework of these indicies is based on use of [69] f-divergence, concept of dissimilarity measure and kernel trick [70]. Finally, [66] propose feature selection as an alternative to screening methods in sensitivity analysis. The main issue with variance based indicies [30] is that even though they capture importance information regarding the contribution of the input factors, they • do not handle multivariate random variables easily and • are only invariant under linear transformations. In comparison to these variance methods, the newly proposed indicies based on density estimations [63] and dependence measures are more robust. Figure 3

shows the general flow of the mathematical formulation for computing the indices in the density based HSIC method. The general idea is as follows - The sensitivity index is actually a distance correlation which incorporates the kernel based Hilbert-Schmidt Information Criterion between two input vectors in higher dimension. The criterion is nothing but the Hilbert-Schmidt norm of cross-covariance operator which generalizes the covariance matrix by representing higher order correlations between the input vectors through nonlinear kernels. For every operator and provided the sum converges, the Hilbert-Schmidt norm is the dot product of the orthonormal bases. For a finite dimensional input vectors, the Hilbert-Schmidt Information Criterion estimator is a trace of product of two kernel matrices (or the Gram matrices) with a centering matrix such that HSIC evaluates to a summation of different kernel values.
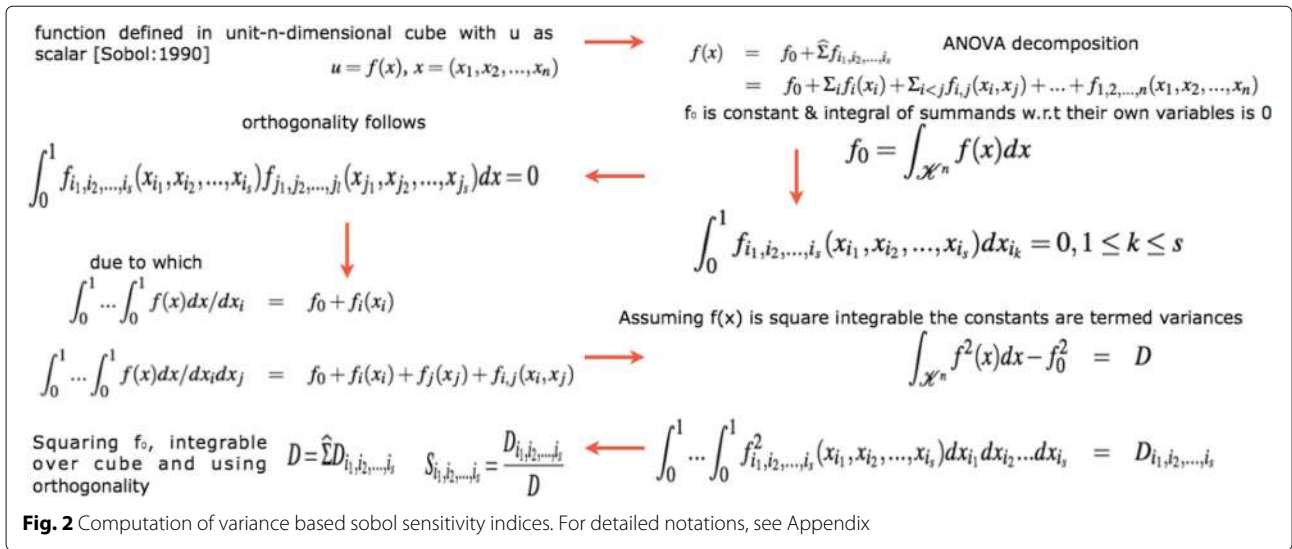
It is this strength of the kernel methods that HSIC is able to capture the deep nonlinearities in the biological data and provide reasonable information regarding the degree of influence of the involved factors within the pathway. Improvements in variance based methods also provide ways to cope with these nonlinearities but do not exploit the available strength of kernel methods. Results in the later sections provide experimental evidence for the same.

Application in systems biology

Recent efforts in systems biology to understand the importance of various factors apropos output behaviour has gained prominence. [71] compares the use of [30] variance based indices versus [32] screening method which uses a One-at-a-time (OAT) approach to analyse the sensitivity of $GSK3$ dynamics to uncertainty in an insulin signaling model. Similar efforts, but on different pathways can be found in [72] and [73].

SA provides a way of analyzing various factors taking part in a biological phenomena and deals with the effects of these factors on the output of the biological system under consideration. Usually, the model equations are differential in nature with a set of inputs and the associated set of parameters that guide the output. SA helps in observing how the variance in these parameters and inputs leads to changes in the output behaviour. The goal of this manuscript is not to analyse differential equations and the parameters associated with it. Rather, the aim is to observe which input genotypic factors have greater contribution to observed phenotypic behaviour like a sample being normal or cancerous in both static and time series data. In this process, the effect of fold changes and deviations in fold changes in time is also considered for analysis in the light of the recently observed psychophysical laws acting downstream of the Wnt pathway [25].

There are two approaches to sensitivity analysis. The first is the local sensitivity analysis in which if there is a

Fig. 2 Computation of variance based sobol sensitivity indices. For detailed notations, see Appendix

required solution, then the sensitivity of a function apropos a set of variables is estimated via a partial derivative for a fixed point in the input space. In global sensitivity, the input solution is not specified. This implies that the model function lies inside a cube and the sensitivity indices are regarded as tools for studying the model instead of the solution. The general form of g-function (as the model or output variable) is used to test the sensitivity of each of the input factor (i.e expression profile of each of the genes). This is mainly due to its non-linearity, non-monotonicity as well as the capacity to produce analytical sensitivity indices. The g-function takes the form -

$$f(x) = \Pi_{i=1}^{d} \frac{|4 * x_i - 2| + a_i}{1 + a_i} \tag{1}$$

were, d is the total number of dimensions and $a_i \geq 0$ are the indicators of importance of the input variable x_i. Note that lower values of a_i indicate higher importance of x_i. In our formulation, we randomly assign values of

$a_i \in [0, 1]$. For the static (time series) data $d = 18(71)$ (factors affecting the pathway). Thus the expression profiles of the various genetic factors in the pathway are considered as input factors and the global analysis conducted. Note that in the predefined dataset, the working of the signaling pathway is governed by a preselected set of genes that affect the pathway. For comparison purpose, the local sensitivity analysis method is also used to study how the individual factor is behaving with respect to the remaining factors while working of the pathway is observed in terms of expression profiles of the various factors.

Finally, in context of [25]'s work regarding the recent development of observation of Weber's law working downstream of the pathway, it has been found that the law is governed by the ratio of the deviation in the input and the absolute input value. More importantly, it is these deviations in input that are of significance in studying such a phenomena. The current manuscript explores the sensitivity of deviation in the fold changes between

Fig. 3 Computation of density based hsic sensitivity indices. For detailed notations, see Appendix
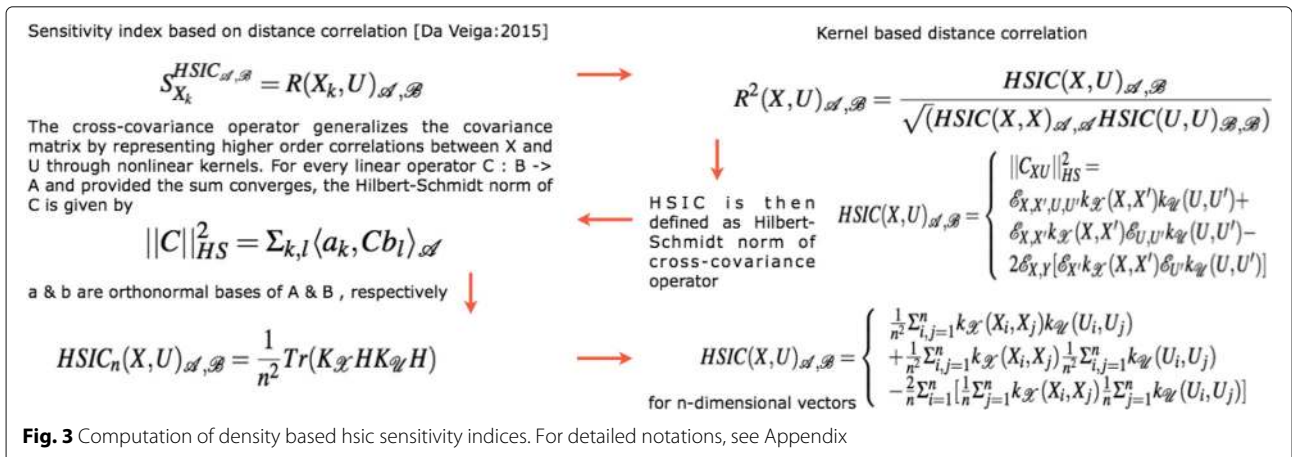
measurements of fold changes at consecutive time points to explore in what duration of time, a particular factor is affecting the pathway in a major way. This has deeper implications in the fact that one is now able to observe when in time an intervention can be made or a gene be perturbed to study the behaviour of the pathway in tumorous cases. Thus sensitivity analysis of deviations in the mathematical formulation of the psychophysical law can lead to insights into the time period based influence of the involved factors in the pathway. Thus, both global and local anaylsis methods are employed to observe the entire behaviour of the pathway as well as the local behaviour of the input factors with respect to the other factors, respectively, via analysis of fold changes and deviations in fold changes, in time.

Given the range of estimators available for testing the sensitivity, it might be useful to list a few which are going to be employed in this research study. These have been described in the Appendix.

The logarithmic psychophysical law
In a recent development, [25] point to two findings namely, • the robust fold changes of β-catenin and • the transcriptional machinery of the Wnt pathway depends on the fold changes in β-catenin instead of absolute levels of the same and some gene transcription networks must respond to fold changes in signals according to the Weber's law in sensory physiology. In an unpublished work by [74], preliminary analysis of results in [23] shows that the variation in predictive behaviour of β-catenin based transcription complex conditional on gene evidences follows power and logarithmic psychophysical law crudely, implying deviations in output are proportional to increasing function of deviations in input and showing constancy for higher values of input. This relates to the work of [75] on power and logarithmic law albeit at a coarse level. A description of these laws ensues before the analysis of the results.

Masin et al. [76] states the Weber's law as follows - Consider a sensation magnitude γ determined by a stimulus magnitude β. Fechner [77] (vol 2, p. 9) used the symbol $\Delta\gamma$ to denote a just noticeable sensation increment, from γ to $\gamma + \Delta\gamma$, and the symbol $\Delta\beta$ to denote the corresponding stimulus increment, from β to $\beta + \Delta\beta$. Fechner [77] (vol 1, p. 65) attributed to the German physiologist Ernst Heinrich Weber the empirical finding [78] that $\Delta\gamma$ remains constant when the relative stimulus increment $\frac{\Delta\beta}{\beta}$ remains constant, and named this finding Weber's law. [77] (vol 2, p. 10) underlined that Weber's law was empirical. ∎

It has been found that Bernoulli's principle [79] is different from Weber's law [78] in that it refers to $\Delta\gamma$ as any possible increment in γ, while the Weber's law refers only to just noticeable increment in γ. Masin et al. [76] shows

that Weber's law is a special case of Bernoulli's principle and can be derived as follows - Eq. 2 depicts the Bernoulli's principle and increment in sensation represented by $\Delta\gamma$ is proportional to change in stimulus represented by $\Delta\beta$.

$$\gamma = b \times \log \frac{\beta}{\alpha} \qquad (2)$$

were b is a constant and α is a threshold. To evaluate the increment, the following Eq. 3 and the ensuing simplification gives -

$$\begin{aligned} \Delta\gamma &= b \times \log \frac{\beta + \Delta\beta}{\alpha} - b \times \log \frac{\beta}{\alpha} \\ &= b \times \log(\frac{\beta + \Delta\beta}{\beta}) = b \times \log(1 + \frac{\Delta\beta}{\beta}) \end{aligned} \qquad (3)$$

Since b is a constant, Eq. 3 reduces to $\Delta\gamma \circ \frac{\Delta\beta}{\beta}$ were $\circ$ means "is constant when there is constancy of" from [76]. The final reduction is a formulation of Weber's laws in wordings and thus Bernoulli's principles imply Weber's law as a special case. Using [77] derivation, it is possible to show the relation between Bernoulli's principles and Weber's law. Starting from the last line of Eq. 3, the following steps yield the relation -

$$\begin{aligned} e^{\Delta\gamma} &= e^{b \times \log(1 + \frac{\Delta\beta}{\beta})} \implies \\ k_p &= e^{\log(1 + \frac{\Delta\beta}{\beta})^b}; \text{ were } k_p = e^{\Delta\gamma} \\ k_p &= (1 + \frac{\Delta\beta}{\beta})^b; \text{ since } e^{\log(x)} = x \\ \sqrt[b]{k_p} &= 1 + \frac{\Delta\beta}{\beta} \implies k_q - 1 = \frac{\Delta\beta}{\beta}; \text{ were } \sqrt[b]{k_p} = k_q \\ k_r &= \frac{\Delta\beta}{\beta}; \text{ the weber's law s.t. } k_r = \sqrt[b]{e^{\Delta\gamma}} - 1 \end{aligned}$$

$$(4)$$

The reduction $\Delta\gamma \circ \frac{\Delta\beta}{\beta}$ holds true given the last line of Eq. 4. By observation, it is important to note that the deviation Δ in the stimulus β plays a crucial role in the above depicted formulations. In the current study, instead of computing the sensitivity of the laws for each involved factor, the sensitivity of the deviations in the fold changes of each factor is taken into account. This is done in order to study the affect of deviations in fold changes in time as concentrations of *WNT3A* changes at a constant rate. Without loss of generality, it was observed over time that most involved factors had sensitivity indices or strength of contributions, parts or whole of whose graphs follow a convex or a concave curvature. These are usually represented by either an exponentially increasing or decreasing curve or nonlinear curves. This points towards the fact that with increasing changes in stimulated concentration of *WNT3A* the deviations in fold changes of an involved factor behave either in an increasing or decreasing fashion. Thus deviations in fold changes of various involved factors does affect the working of the signaling pathway

over time. Finally, these deviations approximately capture the difference in fold changes recorded between two time frames and are thus a measure of how much the involvement of a factor affects the pathway due to these differences. This measure of involvement is depicted via the estimated sensitivity indices. The study of deviations in fold changes might help in deciding when a therapeutic drug could be administered in time. Future wet lab tests can confirm the findings of the above solution.

Methods

Variance based sensitivity indices

The variance based indices as proposed by [30] prove a theorem that an integrable function can be decomposed into summands of different dimensions. Also, a Monte Carlo algorithm is used to estimate the sensitivity of a function apropos arbitrary group of variables. It is assumed that a model denoted by function $u = f(x)$, $x = (x_1, x_2, \ldots, x_n)$, is defined in a unit n-dimensional cube $\mathcal{K}^n$ with u as the scalar output. The requirement of the problem is to find the sensitivity of function $f(x)$ with respect to different variables. If $u^* = f(x^*)$ is the required solution, then the sensitivity of u^* apropos x_k is estimated via the partial derivative $(\partial u/\partial x_k)_{x=x^*}$. This approach is the local sensitivity. In global sensitivity, the input $x = x^*$ is not specified. This implies that the model $f(x)$ lies inside the cube and the sensitivity indices are regarded as tools for studying the model instead of the solution. Detailed technical aspects with examples can be found in [42] and [80].

Let a group of indices $i_1, i_2, \ldots, i_s$ exist, where $1 \leq i_1 < \cdots < i_s \leq n$ and $1 \leq s \leq n$. Then the notation for sum over all different groups of indices is -

$$\widehat{\Sigma} T_{i_1,i_2,\ldots,i_s} = \Sigma_{i=1}^n T_i + \Sigma_{s=1}^n \Sigma_{1 \leq i < j \leq n} T_{i,j} + \cdots + T_{1,2,\ldots,n} \tag{5}$$

Then the representation of $f(x)$ using Eq. 5 in the form -

$$f(x) = f_0 + \widehat{\Sigma} f_{i_1,i_2,\ldots,i_s} \tag{6}$$
$$= f_0 + \Sigma_i f_i(x_i) + \Sigma_{i<j} f_{i,j}(x_i, x_j) + \ldots$$
$$+ f_{1,2,\ldots,n}(x_1, x_2, \ldots, x_n) \tag{7}$$

is called ANOVA-decomposition from [55] or expansion into summands of different dimensions, if f_0 is a constant and integrals of the summands $f_{i_1,i_2,\ldots,i_s}$ with respect to their own variables are zero, i.e,

$$f_0 = \int_{\mathcal{K}^n} f(x) dx \tag{8}$$

$$\int_0^1 f_{i_1,i_2,\ldots,i_s}(x_{i_1}, x_{i_2}, \ldots, x_{i_s}) dx_{i_k} = 0, 1 \leq k \leq s \tag{9}$$

It follows from Eq. 7 that all summands on the right hand side are orthogonal, i.e if at least one of the indices in $i_1, i_2, \ldots, i_s$ and $j_1, j_2, \ldots, j_l$ is not repeated i.e

$$\int_0^1 f_{i_1,i_2,\ldots,i_s}(x_{i_1}, \ldots, x_{i_s}) f_{j_1,\ldots,j_l}(x_{j_1}, x_{j_2}, \ldots, x_{j_s}) dx = 0 \tag{10}$$

[30] proves a theorem stating that there is an existence of a unique expansion of Eq. 7 for any $f(x)$ integrable in $\mathcal{K}^n$. In brief, this implies that for each of the indices as well as a group of indices, integrating equation 7 yields the following -

$$\int_0^1 .. \int_0^1 f(x) dx/dx_i = f_0 + f_i(x_i) \tag{11}$$

$$\int_0^1 .. \int_0^1 f(x) dx/dx_i dx_j = f_0 + f_i(x_i) + f_j(x_j)$$
$$+ f_{i,j}(x_i, x_j) \tag{12}$$

were, dx/dx_i is $\prod_{\forall k \in \{1,..,n\}; i \notin k} dx_k$ and $dx/dx_i dx_j$ is $\prod_{\forall k \in \{1,..,n\}; i,j \notin k} dx_k$. For higher orders of grouped indices, similar computations follow. The computation of any summand $f_{i_1,i_2,\ldots,i_s}(x_{i_1}, x_{i_2}, \ldots, x_{i_s})$ is reduced to an integral in the cube $\mathcal{K}^n$. The last summand $f_{1,2,\ldots,n}(x_1, x_2, \ldots, x_n)$ is $f(x) - f_0$ from Eq. 7. Homma and Saltelli [42] stresses that use of Sobol sensitivity indices does not require evaluation of any $f_{i_1,i_2,\ldots,i_s}(x_{i_1}, x_{i_2}, \ldots, x_{i_s})$ nor the knowledge of the form of $f(x)$ which might well be represented by a computational model i.e a function whose value is only obtained as the output of a computer program.

Finally, assuming that $f(x)$ is square integrable, i.e $f(x) \in \mathcal{L}_2$, then all of $f_{i_1,i_2,\ldots,i_s}(x_{i_1}, x_{i_2}, \ldots, x_{i_s}) \in \mathcal{L}_2$. Then the following constants

$$\int_{\mathcal{K}^n} f^2(x) dx - f_0^2 = D \tag{13}$$

$$\int_0^1 .. \int_0^1 f_{i_1,..,i_s}^2(x_{i_1}, .., x_{i_s}) dx_{i_1}..dx_{i_s} = D_{i_1,..,i_s} \tag{14}$$

are termed as variances. Squaring Eq. 7, integrating over $\mathcal{K}^n$ and using the orthogonality property in Eq. 10, D evaluates to -

$$D = \widehat{\Sigma} D_{i_1,i_2,\ldots,i_s} \tag{15}$$

Then the global sensitivity estimates is defined as -

$$S_{i_1,i_2,\ldots,i_s} = \frac{D_{i_1,i_2,\ldots,i_s}}{D} \tag{16}$$

It follows from Eqs. 15 and 16 that

$$\widehat{\Sigma} S_{i_1,i_2,\ldots,i_s} = 1 \tag{17}$$

Clearly, all sensitivity indices are non-negative, i.e an index $S_{i_1,i_2,\ldots,i_s} = 0$ if and only if $f_{i_1,i_2,\ldots,i_s} \equiv 0$. The true potential of Sobol indices is observed when variables $x_1, x_2, \ldots, x_n$ are divided into m different groups with $y_1, y_2, \ldots, y_m$ such that $m < n$. Then $f(x) \equiv f(y_1, y_2, \ldots, y_m)$. All properties remain the same for the computation of sensitivity indices with the fact that integration with respect to y_k means integration with respect to all the x_i's in y_k. Details of these computations with

examples can be found in [30]. Variations and improvements over Sobol indices have already been stated in "Sensitivity analysis" section.

Density based sensitivity indices

As discussed before, the issue with variance based methods is the high computational cost incurred due to the number of interactions among the variables. This further requires the use of screening methods to filter out redundant or unwanted factors that might not have significant impact on the output. Recent work by [66] proposes a new class of sensitivity indicies which are a special case of density based indicies [63]. These indicies can handle multivariate variables easily and relies on density ratio estimation. Key points from [66] are mentioned below.

Considering the similar notation in previous section, $f : \mathcal{R}^n \rightarrow \mathcal{R}$ ($u = f(x)$) is assumed to be continuous. It is also assumed that X_k has a known distribution and are independent. Baucells and Borgonovo [81] state that a function which measures the similarity between the distribution of U and that of $U|X_k$ can define the impact of X_k on U. Thus the impact is defined as -

$$S_{X_k} = \mathcal{E}(d(U, U|X_k)) \tag{18}$$

were $d(\cdot, \cdot)$ is a dissimilarity measure between two random variables. Here d can take various forms as long as it satisfies the criteria of a dissimilarity measure. Csiszar [69]'s f-divergence between U and $U|X_k$ when all input random variables are considered to be absolutely continuous with respect to Lebesgue measure on $\mathcal{R}$ is formulated as -

$$d_F(U||U|X_k) = \int_{\mathcal{R}} F(\frac{p_U(u)}{p_{U|X_k}(u)}) p_{U|X_k}(u) du \tag{19}$$

were F is a convex function such that $F(1) = 0$ and p_U and $p_{U|X_k}$ are the probability distribution functions of U and $U|X_k$. Standard choices of F include Kullback-Leibler divergence $F(t) = -\log_e(t)$, Hellinger distance $(\sqrt{t} - 1)^2$, Total variation distance $F(t) = |t - 1|$, Pearson χ^2 divergence $F(t) = t^2 - 1$ and Neyman χ^2 divergence $F(t) = (1 - t^2)/t$. Substituting Eq. 19 in equation 18, gives the following sensitivity index -

$$
\begin{aligned}
S_{X_k}^F &= \int_{\mathcal{R}} d_F(U||U|X_k) p_{X_k}(x) dx \\
&= \int_{\mathcal{R}} \int_{\mathcal{R}} F(\frac{p_U(u)}{p_{U|X_k}(u)}) p_{U|X_k}(u) p_{X_k}(x) dx du \\
&= \int_{\mathcal{R}^2} F(\frac{p_U(u) p_{X_k}(x)}{p_{U|X_k}(u) p_{X_k}(x)}) p_{U|X_k}(u) p_{X_k}(x) dx du \\
&= \int_{\mathcal{R}^2} F(\frac{p_U(u) p_{X_k}(x)}{p_{X_k,U}(x,u)}) p_{X_k,U}(x,u) dx du
\end{aligned}
\tag{20}
$$

were p_{X_k} and $p_{X_k,Y}$ are the probability distribution functions of X_k and (X_k, U), respectively. Csiszar [69] f-divergences imply that these indices are positive and

equate to 0 when U and X_k are independent. Also, given the formulation of $S_{X_k}^F$, it is invariant under any smooth and uniquely invertible transformation of the variables X_k and U [82]. This has an advantage over Sobol sensitivity indices which are invariant under linear transformations.

By substituting the different formulations of F in Eqs 20, '[66]'s work claims to be the first in establishing the link that previously proposed sensitivity indices are actually special cases of more general indices defined through [69]'s f-divergence. Then Eq. 20 changes to estimation of ratio between the joint density of (X_k, U) and the marginals, i.e -

$$
\begin{aligned}
S_{X_k}^F &= \int_{\mathcal{R}^2} F(\frac{1}{r(x,u)}) p_{X_k,U}(x,u) dx du \\
&= \mathcal{E}_{(X_k,U)} F(\frac{1}{r(X_k, U)})
\end{aligned}
\tag{21}
$$

were, $r(x,y) = (p_{X_k,U}(x,u))/(p_U(u) p_{X_k}(x))$. Multivariate extensions of the same are also possible under the same formulation.

Finally, given two random vectors $X \in \mathcal{R}^p$ and $Y \in \mathcal{R}^q$, the dependence measure quantifies the dependence between X and Y with the property that the measure equates to 0 if and only if X and Y are independent. These measures carry deep links [83] with distances between embeddings of distributions to reproducing kernel Hilbert spaces (RHKS) and here the related Hilbert-Schmidt independence criterion (HSIC by [68]) is explained.

In a very brief manner from an extremely simple introduction by [84] - "We first defined a field, which is a space that supports the usual operations of addition, subtraction, multiplication and division. We imposed an ordering on the field and described what it means for a field to be complete. We then defined vector spaces over fields, which are spaces that interact in a friendly way with their associated fields. We defined complete vector spaces and extended them to Banach spaces by adding a norm. Banach spaces were then extended to Hilbert spaces with the addition of a dot product." Mathematically, a Hilbert space $\mathcal{H}$ with elements $r, s \in \mathcal{H}$ has dot product $\langle r, s \rangle_{\mathcal{H}}$ and $r \cdot s$. When $\mathcal{H}$ is a vector space over a field $\mathcal{F}$, then the dot product is an element in $\mathcal{F}$. The product $\langle r, s \rangle_{\mathcal{H}}$ follows the below mentioned properties when $r, s, t \in \mathcal{H}$ and for all $a \in \mathcal{F}$ -

- Associative : $(ar) \cdot s = a(r \cdot s)$
- Commutative : $r \cdot s = s \cdot r$
- Distributive : $r \cdot (s + t) = r \cdot s + r \cdot t$

Given a complete vector space $\mathcal{V}$ with a dot product $\langle \cdot, \cdot \rangle$, the norm on $\mathcal{V}$ defined by $||r||_{\mathcal{V}} = \sqrt{\langle r, r \rangle}$ makes this space into a Banach space and therefore into a full Hilbert space.

A reproducing kernel Hilbert space (RKHS) builds on a Hilbert space $\mathcal{H}$ and requires all Dirac evaluation functionals in $\mathcal{H}$ are bounded and continuous (on implies the other). Assuming $\mathcal{H}$ is the $\mathcal{L}_2$ space of functions from X to $\mathcal{R}$ for some measurable X. For an element $x \in X$, a Dirac evaluation functional at x is a functional $\delta_x \in \mathcal{H}$ such that $\delta_x(g) = g(x)$. For the case of real numbers, x is a vector and g a function which maps from this vector space to $\mathcal{R}$. Then δ_x is simply a function which maps g to the value g has at x. Thus, δ_x is a function from $(\mathcal{R}^n \mapsto \mathcal{R})$ into $\mathcal{R}$.

The requirement of Dirac evaluation functions basically means (via the [85] representation theorem) if ϕ is a bounded linear functional (conditions satisfied by the Dirac evaluation functionals) on a Hilbert space $\mathcal{H}$, then there is a unique vector ℓ in $\mathcal{H}$ such that $\phi g = \langle g, \ell \rangle_{\mathcal{H}}$ for all $\ell \in \mathcal{H}$. Translating this theorem back into Dirac evaluation functionals, for each δ_x there is a unique vector k_x in $\mathcal{H}$ such that $\delta_x g = g(x) = \langle g, k_x \rangle_{\mathcal{H}}$. The reproducing kernel K for $\mathcal{H}$ is then defined as : $K(x, x') = \langle k_x, k_{x'} \rangle$, were k_x and $k_{x'}$ are unique representatives of δ_x and $\delta_{x'}$. The main property of interest is $\langle g, K(x, x') \rangle_{\mathcal{H}} = g(x')$. Furthermore, k_x is defined to be a function $y \mapsto K(x, y)$ and thus the reproducibility is given by $\langle K(x, \cdot), K(y, \cdot) \rangle_{\mathcal{H}} = K(x, y)$.

Basically, the distance measures between two vectors represent the degree of closeness among them. This degree of closeness is computed on the basis of the discriminative patterns inherent in the vectors. Since these patterns are used implicitly in the distance metric, a question that arises is, how to use these distance metric for decoding purposes?

The kernel formulation as proposed by [70], is a solution to our problem mentioned above. For simplicity, we consider the labels of examples as binary in nature. Let $\mathbf{x}_i \in \mathcal{R}^n$, be the set of n feature values with corresponding category of the example label (y_i) in data set $\mathcal{D}$. Then the data points can be mapped to a higher dimensional space $\mathcal{H}$ by the transformation ϕ:

$$\phi : \mathbf{x}_i \in \mathcal{R}^n \mapsto \phi(\mathbf{x}_i) \in \mathcal{H} \qquad (22)$$

This $\mathcal{H}$ is the *Hilbert Space* which is a strict inner product space, along with the property of completeness as well as separability. The inner product formulation of a space helps in discriminating the location of a data point w.r.t a separating hyperplane in $\mathcal{H}$. This is achieved by the evaluation of the inner product between the normal vector representing the hyperplane along with the vectorial representation of a data point in $\mathcal{H}$ (Fig. 4 represents the geometrical interpretation). Thus, the idea behind Eq. (22) is that even if the data points are nonlinearly clustered in space $\mathcal{R}^n$, the transformation spreads the data points into $\mathcal{H}$, such that they can be linearly separated in its range in $\mathcal{H}$.

Often, the evaluation of dot product in higher dimensional spaces is computationally expensive. To avoid
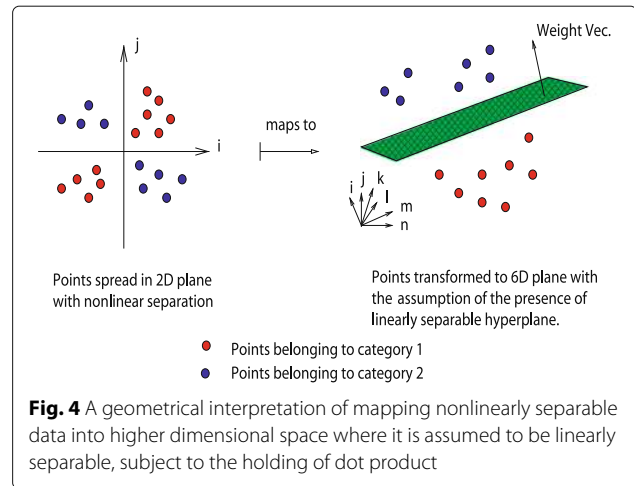
Fig. 4 A geometrical interpretation of mapping nonlinearly separable data into higher dimensional space where it is assumed to be linearly separable, subject to the holding of dot product

incurring this cost, the concept of kernels in employed. The trick is to formulate kernel functions that depend on a pair of data points in the space $\mathcal{R}^n$, under the assumption that its evaluation is equivalent to a dot product in the higher dimensional space. This is given as:

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle \qquad (23)$$

Two advantages become immediately apparent. First, the evaluation of such kernel functions in lower dimensional space is computationally less expensive than evaluating the dot product in higher dimensional space. Secondly, it relieves the burden of searching an appropriate transformation that may map the data points in $\mathcal{R}^n$ to $\mathcal{H}$. Instead, all computations regarding discrimination of location of data points in higher dimensional space involves evaluation of the kernel functions in lower dimension. The matrix containing these kernel evaluations is referred to as the *kernel* matrix. With a cell in the kernel matrix containing a kernel evaluation between a pair of data points, the kernel matrix is square in nature.

As an example in practical applications, once the kernel has been computed, a pattern analysis algorithm uses the kernel function to evaluate and predict the nature of the new example using the general formula:

$$\begin{aligned} f(\mathbf{z}) &= \langle \mathbf{w}, \phi(\mathbf{z}) \rangle + b \\ &= \langle \sum_{i=1}^{N} \alpha_i \times y_i \times \phi(\mathbf{x}_i), \phi(\mathbf{z}) \rangle + b \\ &= \sum_{i=1}^{N} \alpha_i \times y_i \times \langle \phi(\mathbf{x}_i), \phi(\mathbf{z}) \rangle + b \\ &= \sum_{i=1}^{N} \alpha_i \times y_i \times \kappa(\mathbf{x}_i, \mathbf{z}) + b \end{aligned}$$

$$(24)$$

where **w** defines the hyperplane as some linear combination of training basis vectors, **z** is the test data point, y_i the class label for training point $\mathbf{x}_i$, α_i and b are the constants. Various transformations to the kernel function can be employed, based on the properties a kernel must satisfy. Interested readers are referred to [86] for description of these properties in detail.

The Hilbert-Schmidt independence criterion (HSIC) proposed by [68] is based on kernel approach for finding dependences and on cross-covariance operators in RKHS. Let $X \in \mathcal{X}$ have a distribution P_X and consider a RKHS $\mathcal{A}$ of functions $\mathcal{X} \rightarrow \mathcal{R}$ with kernel $k_{\mathcal{X}}$ and dot product $\langle \cdot, \cdot \rangle_{\mathcal{A}}$. Similarly, Let $U \in \mathcal{Y}$ have a distribution P_Y and consider a RKHS $\mathcal{B}$ of functions $\mathcal{U} \rightarrow \mathcal{R}$ with kernel $k_{\mathcal{B}}$ and dot product $\langle \cdot, \cdot \rangle_{\mathcal{B}}$. Then the cross-covariance operator $C_{X,U}$ associated with the joint distribution P_{XU} of (X, U) is the linear operator $\mathcal{B} \rightarrow \mathcal{A}$ defined for every $a \in \mathcal{A}$ and $b \in \mathcal{B}$ as -

$$\langle a, C_{XU} b \rangle_{\mathcal{A}} = \mathcal{E}_{XU}[a(X), b(U)] - \mathcal{E}_X a(X) \mathcal{E}_U b(U) \tag{25}$$

The cross-covariance operator generalizes the covariance matrix by representing higher order correlations between X and U through nonlinear kernels. For every linear operator $C : \mathcal{B} \rightarrow \mathcal{A}$ and provided the sum converges, the Hilbert-Schmidt norm of C is given by -

$$||C||_{HS}^2 = \Sigma_{k,l} \langle a_k, C b_l \rangle_{\mathcal{A}} \tag{26}$$

were a_k and b_l are orthonormal bases of $\mathcal{A}$ and $\mathcal{B}$, respectively. The HSIC criterion is then defined as the Hilbert-Schmidt norm of cross-covariance operator -

$$HSIC(X, U)_{\mathcal{A,B}} = \begin{cases} ||C_{XU}||_{HS}^2 = \\ \mathcal{E}_{X,X',U,U'} k_{\mathcal{X}}(X, X') k_{\mathcal{U}}(U, U') + \\ \mathcal{E}_{X,X'} k_{\mathcal{X}}(X, X') \mathcal{E}_{U,U'} k_{\mathcal{U}}(U, U') - \\ 2\mathcal{E}_{X,Y} \left[\mathcal{E}_{X'} k_{\mathcal{X}}(X, X') \mathcal{E}_{U'} k_{\mathcal{U}}(U, U') \right] \end{cases} \tag{27}$$

were the equality in terms of kernels is proved in [68]. Finally, assuming (X_i, U_i) $(i = 1, 2, ..., n)$ is a sample of the random vector (X, U) and $K_{\mathcal{X}}$ and $K_{\mathcal{U}}$ denote the Gram matrices with entries $K_{\mathcal{X}}(i, j) = k_{\mathcal{X}}(X_i, X_j)$ and $K_{\mathcal{U}}(i, j) = k_{\mathcal{U}}(U_i, U_j)$. [68] proposes the following estimator for $HSIC_n(X, U)_{\mathcal{A,B}}$ -

$$HSIC_n(X, U)_{\mathcal{A,B}} = \frac{1}{n^2} Tr(K_{\mathcal{X}} H K_{\mathcal{U}} H) \tag{28}$$

were H is the centering matrix such that $H(i, j) = \delta_{i,j} - \frac{1}{n}$. Then $HSIC_n(X, U)_{\mathcal{A,B}}$ can be expressed as -

$$HSIC(X, U)_{\mathcal{A,B}} = \begin{cases} \frac{1}{n^2} \Sigma_{i,j=1}^n k_{\mathcal{X}}(X_i, X_j) k_{\mathcal{U}}(U_i, U_j) \\ + \frac{1}{n^2} \Sigma_{i,j=1}^n k_{\mathcal{X}}(X_i, X_j) \times \\ \frac{1}{n^2} \Sigma_{i,j=1}^n k_{\mathcal{U}}(U_i, U_j) \\ - \frac{2}{n} \Sigma_{i=1}^n [\frac{1}{n} \Sigma_{j=1}^n k_{\mathcal{X}}(X_i, X_j) \times \\ \frac{1}{n} \Sigma_{j=1}^n k_{\mathcal{U}}(U_i, U_j)] \end{cases} \tag{29}$$

Finally, [66] proposes the sensitivity index based on distance correlation as -

$$S_{X_k}^{HSIC_{\mathcal{A,B}}} = R(X_k, U)_{\mathcal{A,B}} \tag{30}$$

were the kernel based distance correlation is given by -

$$R^2(X, U)_{\mathcal{A,B}} = \frac{HSIC(X, U)_{\mathcal{A,B}}}{\sqrt{(HSIC(X, X)_{\mathcal{A,A}} HSIC(U, U)_{\mathcal{B,B}})}} \tag{31}$$

were kernels inducing $\mathcal{A}$ and $\mathcal{B}$ are to be chosen within a universal class of kernels. Similar multivariate formulation for Eq. 28 are possible.

Description of the dataset & design of experiments

STATIC DATA - A simple static dataset containing expression values measured for a few genes known to have important role in human colorectal cancer cases has been taken from [28]. Most of the expression values recorded are for genes that play a role in Wnt signaling pathway at an extracellular level and are known to have inhibitory affect on the Wnt pathway due to epigenetic factors. For each of the 24 normal mucosa and 24 human colorectal tumor cases, gene expression values were recorded for 14 genes belonging to the family of *SFRP*, *DKK*, *WIF*1 and *DACT*. Also, expression values of established Wnt pathway target genes like *LEF*1, *MYC*, *CD*44 and *CCND*1 were recorded per sample.

TIME SERIES DATA - Contrary to the static data described above, [27] presents a bigger set of 71 Wnt-related gene expression values for 6 different times points over a range of 24-hour period using qPCR. The changes represent the fold-change in the expression levels of genes in 200 ng/mL *WNT3A*-stimulated HEK 293 cells in time relative to their levels in unstimulated, serum-starved cells at 0-hour. The data are the means of three biological replicates. Only genes whose mean transcript levels changed by more than two-fold at one or more time points were considered significant. Positive (negative) numbers represent up (down) -regulation.

Note that green (red) represents activation (repression) in the heat maps of data in [28] and [27].

GENERAL ISSUES - • Here the input factors are the gene expression values for both normal and tumor cases in static data. For the case of time series data, the input factors are the fold change (deviations in fold change) expression values of genes at different time points (periods). Also, for the time series data, in the first experiment the analysis of a pair of the fold changes recorded at to different consecutive time points i.e t_i & t_{i+1} is done. In the second experiment, the analysis of a pair of deviations in fold changes recorded at t_i & t_{i+1} and t_{i+1} & t_{i+2}. In this work, in both the static and the time series datasets, the analysis is done to study the entire model/pathway rather

than find a particular solution to the model/pathway. Thus global sensitivity analysis is employed. But the local sensitivity methods are used to observe and compare the affect of individual factors via 1st order analysis w.r.t total order analysis (i.e global analysis). In such an experiment, the output is the sensitivity indices of the individual factors participating in the model. This is different from the general trend of observing the sensitivity of parameter values that affect the pathway based on differential equations that model a reaction. Thus the model/pathway is studied as a whole by observing the sensitivities of the individual factors.

• Static data - Note that the 24 normal and tumor cases are all different from each other. The 18 genes that are used to study in [28] are the input factors and it is unlikely that there will be correlations between different patients. The phenotypic behaviour might be similar at a grander scale. Also, since the sampling number is very small for a network of this scale, large standard deviations can be observed in many results, especially when the Sobol method is used. But this is not the issue with the sampling number. By that analysis, large deviations are not observed in kernel based density methods. The deviations are more because of the fact that the nonlinearities are not captured in an efficient way in the variance based Sobol methods. Due to this, the resulting indicies have high variance in numerical value. For the same number of samplings, the kernel based methods don't show high variance.

• Time series data - All the measurement data at each time point are generated by a normal distribution with fixed standard deviation of 0.005 plus a noise term. One might enquire as to how does this data generation match with the real experimental data? The kernel based density methods requires a distribution of data. The original experimental data of fold change was taken from each of the genes per time point. Gujral and MacBeath [27] states that to determine the fold-change in gene expression induced by stimulation with Wnt3a, the normalized expression of each gene in the Wnt3a-stimulated sample was divided by the normalized expression of the same gene in the unstimulated sample. The qPCR data presented are mean of three biological replicates. By using a stringent margin of 0.005 and a noise term, the distribution of the data near the mean value is kept constricted. How much it deviates from the reality beyond the errors of measurement is not known to the author! Finally, 74 gene expression values are taken as input per time point for evaluting the sensitivity of each of the genetic factor that affect the model/pathway. Again, one is not looking for a solution to the model in terms of good value for parameters but studying the degree of influence of each of the input factors that constitute the model/pathway.

DESIGN OF EXPERIMENTS - The reported results will be based on scaled as well as unscaled datasets. For the static data, only the scaled results are reported. This is mainly due to the fact that the measurements vary in a wide range and due to this there is often an error in the computed estimated of these indices. The data for time series does not vary in a wide range and thus the results are reported for both the scaled and the non scaled versions. Total sensitivity indices and 1st order indices will be used for sensitivity analysis. For addressing a biological question with known prior knowledge, the order of indices might be increased. While studying the interaction among the various genetic factors using static data, tumor samples are considered separated from normal samples. Bootstrapping without replicates on a smaller sample number is employed to generate estimates of indices which are then averaged. This takes into account the variance in the data and generates confidence bands for the indices. For the case of time series data, interactions among the contributing factors are studied by comparing (1) pairs of fold-changes at single time points and (2) pairs of deviations in fold changes between pairs of time points. Generation of distribution around measurements at single time points with added noise is done to estimate the indices.

Results and discussion
Static data
To measure the strength of the contributing factors in the static dataset by [28], 1st order and total sensitivity indices were generated. For each of the expression values of the genes recorded in the normal and tumor cases, the computation of the indices was done using bootstrapped samples in three different experiments each with a sample size of 8, 16 and 24, respectively. With only 24 samples in total, 20 bootstraps were generated for each set and the results were generated. From these replicates, the mean of the indices is reported along with the 95% confidence bands. Figure 5 represents the cartoon of the experimental setup followed to acheive the desired results. Note that plots of sensitivity indices have been relegated to Appendix.

Using the sensiFdiv, all indices are computed as positive and those nearing to zero indicate the contribution of a factor as independent from the behaviour under consideration. Here, while comparing the indices of the gene expression values for normal and tumor cases, it was found that most of the involved intra/extracellular factors had some degree of contribution in the normal case and almost negligible contribution in the tumor case (see Figs. 6, 7 and 8). Apart from the negative reading for the KL divergence Fig. 9 the interpretations remain the same. This implies that the basic [69] f-divergence based indices might not capture the intrinsic genotypic effects for the normal and the tumorous cases. From the biological perspective, these graphs do not help in interpreting the strength of the contributions in normal and tumor
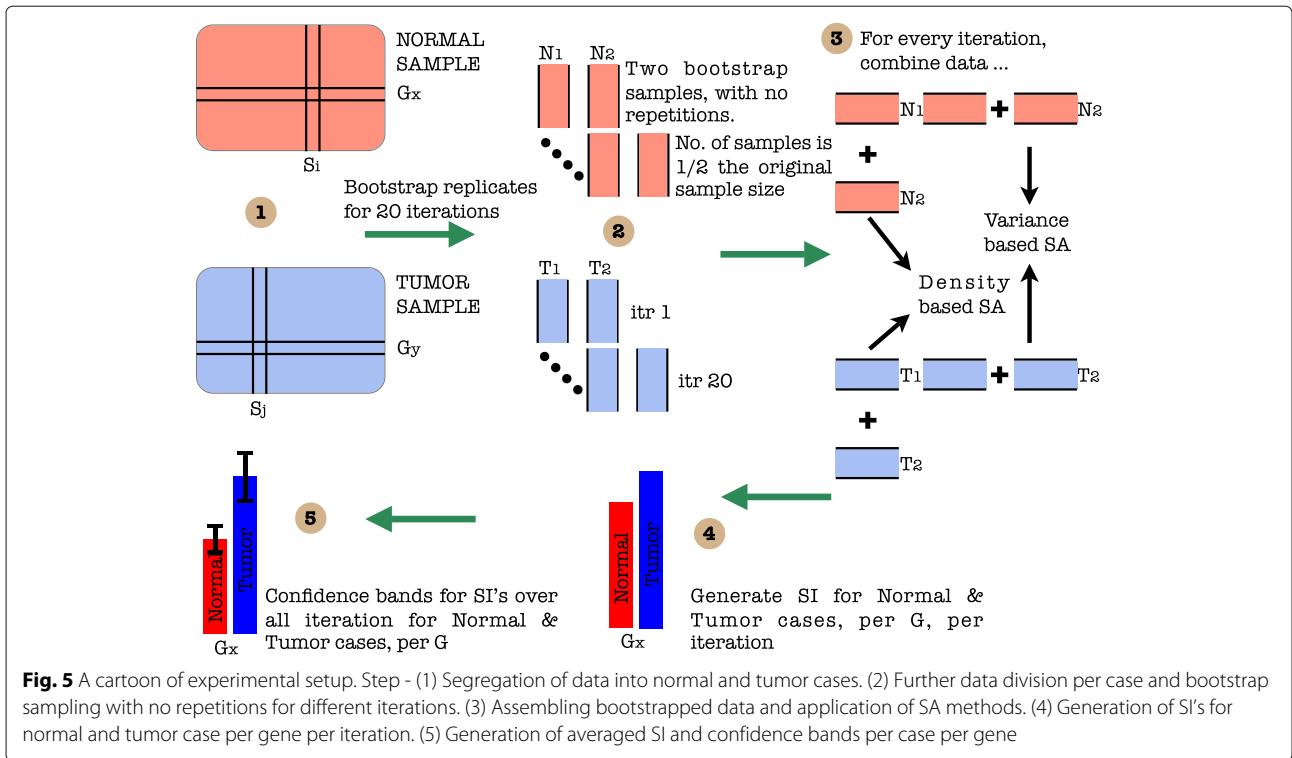
Fig. 5 A cartoon of experimental setup. Step - (1) Segregation of data into normal and tumor cases. (2) Further data division per case and bootstrap sampling with no repetitions for different iterations. (3) Assembling bootstrapped data and application of SA methods. (4) Generation of SI's for normal and tumor case per gene per iteration. (5) Generation of averaged SI and confidence bands per case per gene
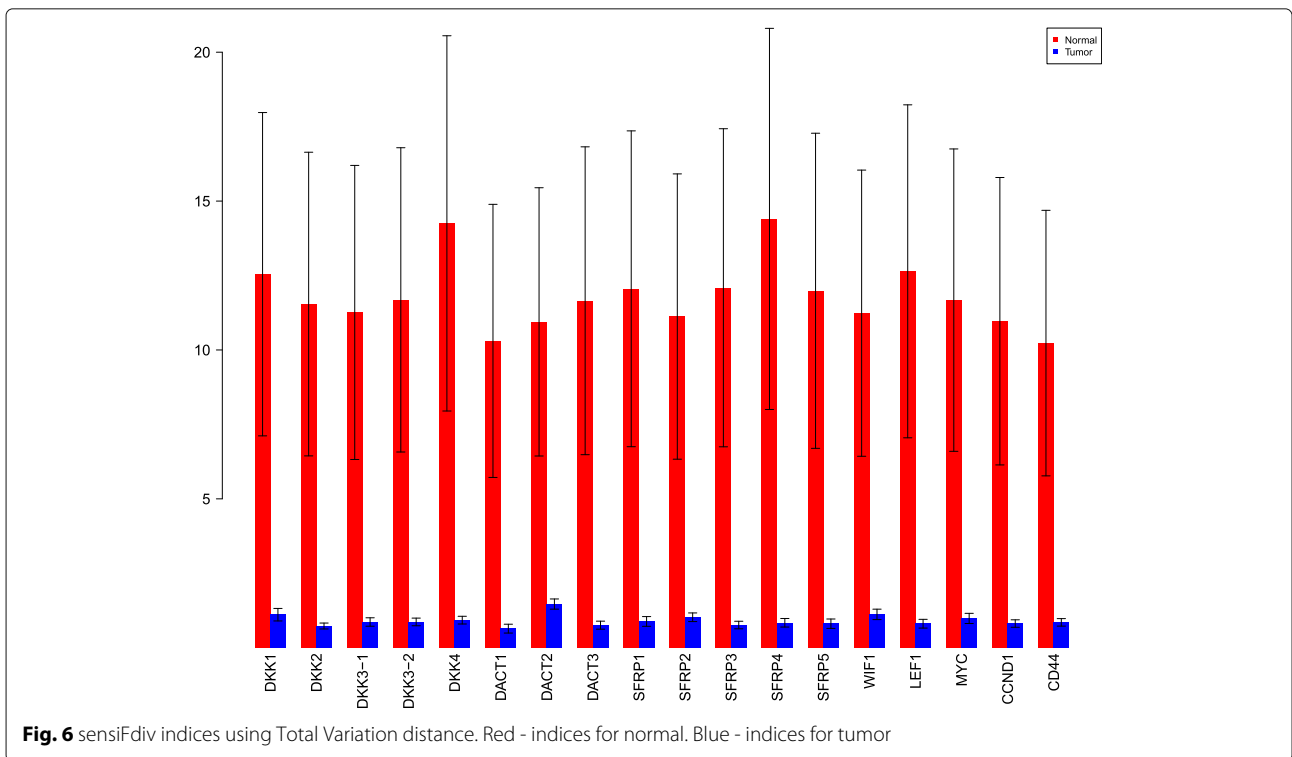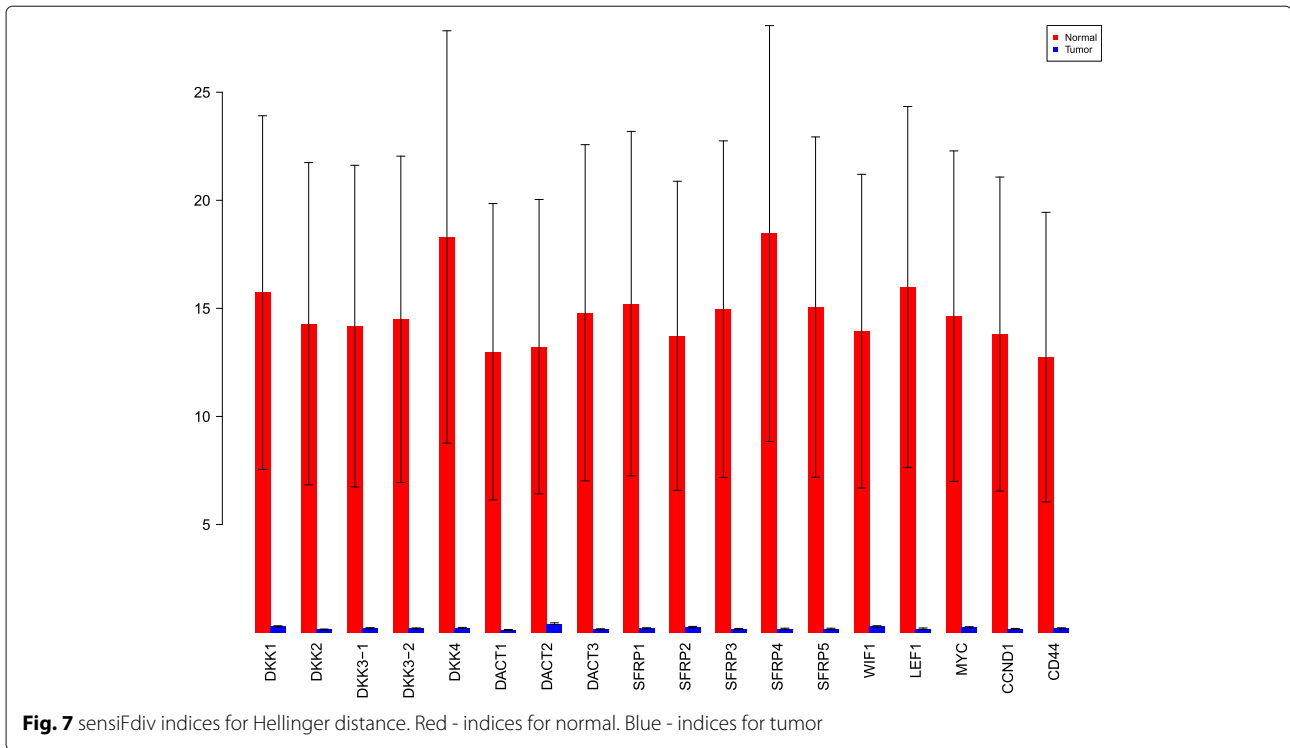

Fig. 6 sensiFdiv indices using Total Variation distance. Red - indices for normal. Blue - indices for tumor

Fig. 7 sensiFdiv indices for Hellinger distance. Red - indices for normal. Blue - indices for tumor

cases. One might rank the indices for relative contributions, but this might not shed enough light on the how the factors are behaving in normal and tumor cases.

A more powerful way to analyse the contributions is the newly proposed HSIC based measures by [66]. These distances use the kernel trick which can capture intrinsic properties inherent in the recorded measurements by transforming the data into a higher dimensional space. Using these distances in sensiHSIC, it was found that the contributions of the various factors in the normal and the

Fig. 8 sensiFdiv indices for Pearson χ^2 distance. Red - indices for normal. Blue - indices for tumor
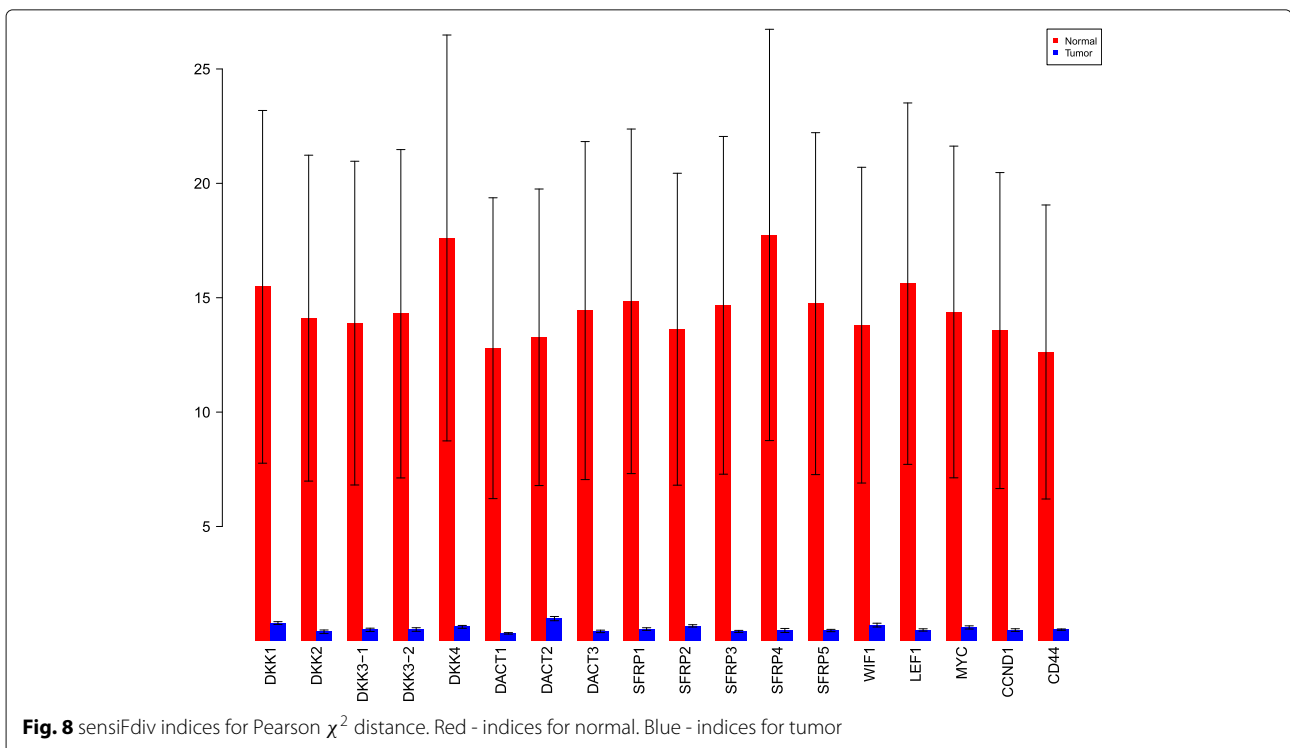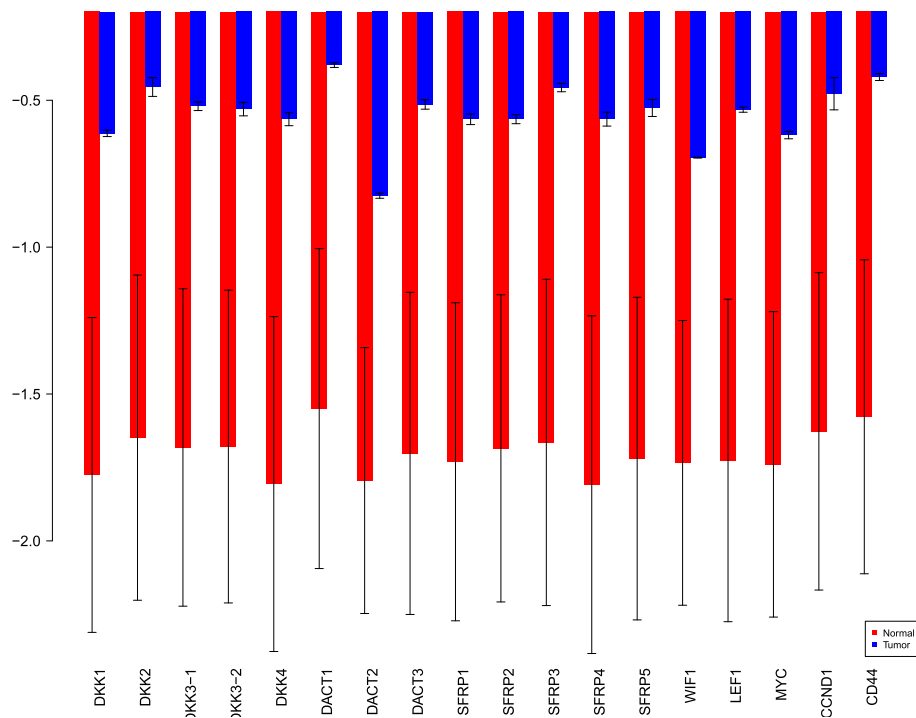
Fig. 9 sensiFdiv indices for Kullback-Leibler divergence. Red - indices for normal. Blue - indices for tumor

tumor cases vary drastically. This is shown in Figs. 10, 11 and 12. The laplace and the rbf kernels give more reliable sensitivity estimates for the involved factors than the linear kernel. Studying the results in figures 6 and 7 of [23] based on prior biological knowledge encoded in the

Bayesian network models along with the indices of aforementioned figures, it can be found that indices of the family of $DACT-1/2/3$ show higher (lower) sensitivity in the normal (tumor) case where the activation (repression) happens. Again, of the $DACT$ family, $DACT-1$ has
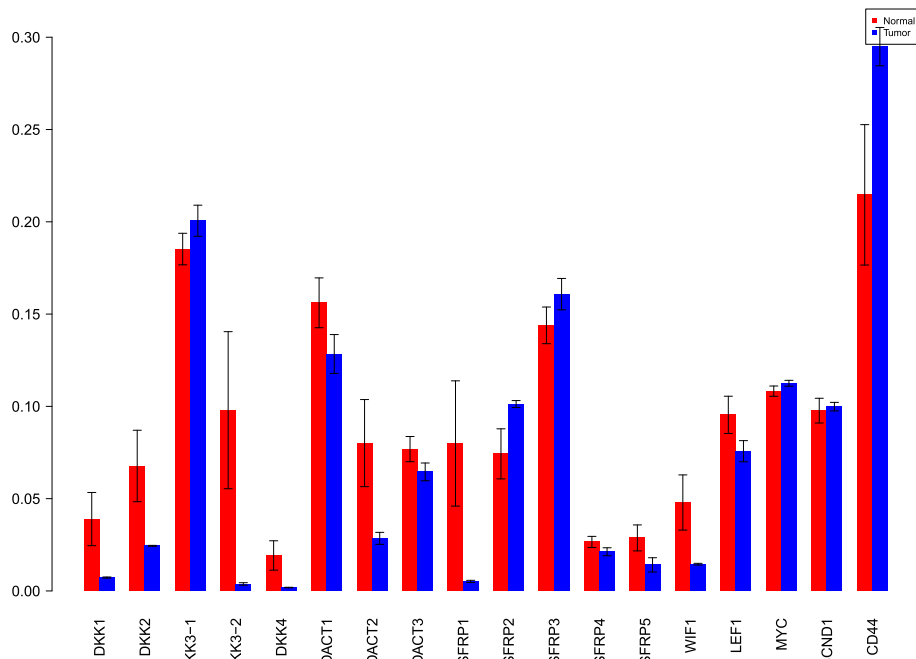
Fig. 10 sensiHSIC indices for linear kernel. Red - indices for normal. Blue - indices for tumor

Fig. 11 sensiHSIC indices for laplace kernel. Red - indices for normal. Blue - indices for tumor
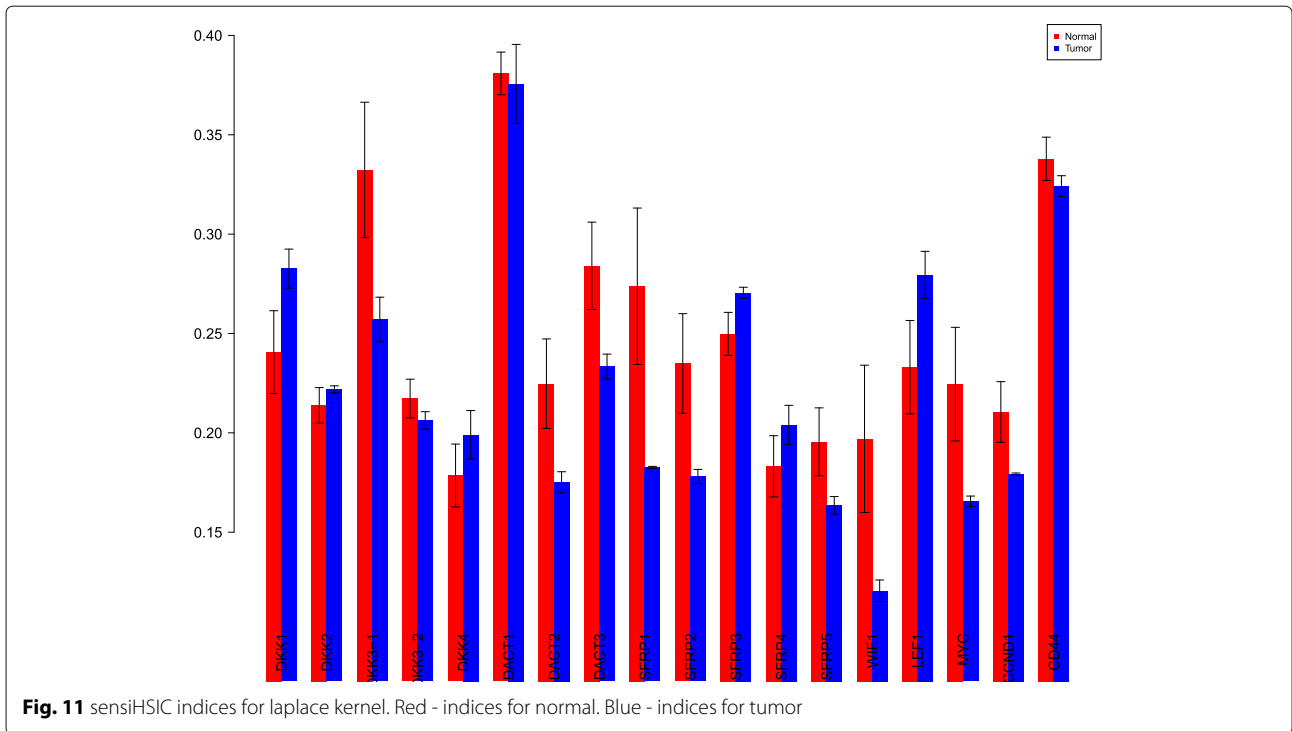
greater influence than $DACT-3$ (than $DACT-2$) based on the values of the sensitivity indices. These indices indicate the dependence of a factor on the output of the model characterized by the signaling being active (passive) in the normal (tumor) cases. 0(1) mean no (full) dependence

of the output on the input factor. The laplace and the rbf kernels were found to give more consistent results than the linear kernel and the following description discusses the results from these kernels. For the *SFRP* family $SFRP-1/2/5$ show higher (lower) sensitivity in normal

Fig. 12 sensiHSIC indices for rbf kernel. Red - indices for normal. Blue - indices for tumor
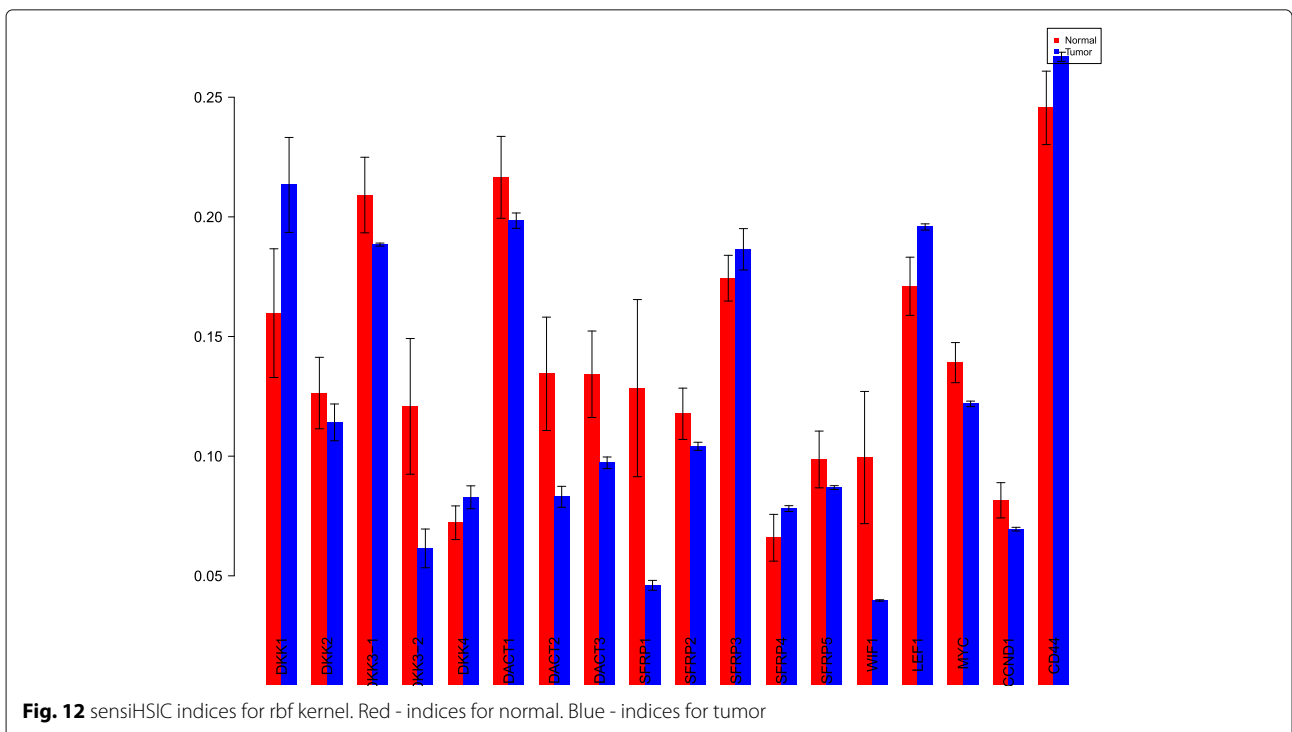
(tumor) case where the activation (repression) happens (see Figs. 11 and 12). For $SFRP-3/4$ the influence is higher (lower) in the tumor (normal) case. In all the three types of kernels, $WIF1$, MYC and $CCND1$ show stronger (weaker) influence of repression (activation) in the normal (tumor) case (see Figs. 11 and 12). $CD44$ showed variable influence while observing the normal and tumor cases. [23] could not derive proper inferences for $LEF1$ but the sensitivity indices indicate that the influence of $LEF1$ in tumor samples to be higher than in normal samples. This points to the $LEF1$'s major role in tumor cases. Finally, for the family of DKK, $DKK1$ and $DKK3-2$ show similar behaviour of expression (repression) in normal (tumor cases) (see [23]). For the former, the prominence of the influence is shown in the higher (lower) sensitivity for tumor (normal) case. For the latter higher (lower) sensitivity was recorded for normal (tumor) case. This implies that the latter has more influential role in normal while the former has more influential role in tumor case. $DKK3-1$ was found to be expressed (repressed) in normal (tumor) and its dominant role is prominent from the higher bar sensitivity bar for normal than the tumor. Similar behavior of $DKK2$ was inferred by [23] but the sensitivity indices point to varied results and thus a conclusion cannot be drawn. Note that greater the value of the sensitivity index, greater is an input factor's contribution to the model.

The first order indices generated by sobol functions implemented in sobol2002 (Fig. 13), sobol2007 (Fig. 14), soboljansen (Fig. 15), sobolmartinez (Fig. 16) and sobol

(Fig. 17) do not point to significant dependencies of the input factors. This can be attributed to the fact that there are less number of samples that help in the estimation of the sensitivity indicies. Finally, the total order indices need to be investigated in the context of the first order indices. It can be observed , sobol2002 (Fig. 18) and sobol2007 (Fig. 18) give much better estimates than soboljansen (Fig. 19) and sobolmartinez (Fig. 20). Most importantly, it is the former two that closely match with the sensitivity indices estimated using the HSIC distance measures. Interpretations from sobol2002 (Fig. 18) and sobol2007 (Fig. 21) are the same as those described above using the laplace and the rbf kernels from density based HSIC measure.

In summary, the sensitivity indices confirm the inferred results in [23] but do not help in inferring the causal relations using the static data. In combination with the results obtained from the Bayesian network models in [23] it is possible to study the effect of the input factors for the pathway in both normal and tumor cases. The results of sensitivity indices indicate how much these factors influence the pathway in normal and tumor cases. Again, not all indices reveal important information. So users must be cautious of results and see which measures reveal information that are close to already established or computationally estimated biological facts. Here the density based sensitivity indices captured information more precisely than the variance based indices (except for the total order indices from sobol2002/7 which gave similar
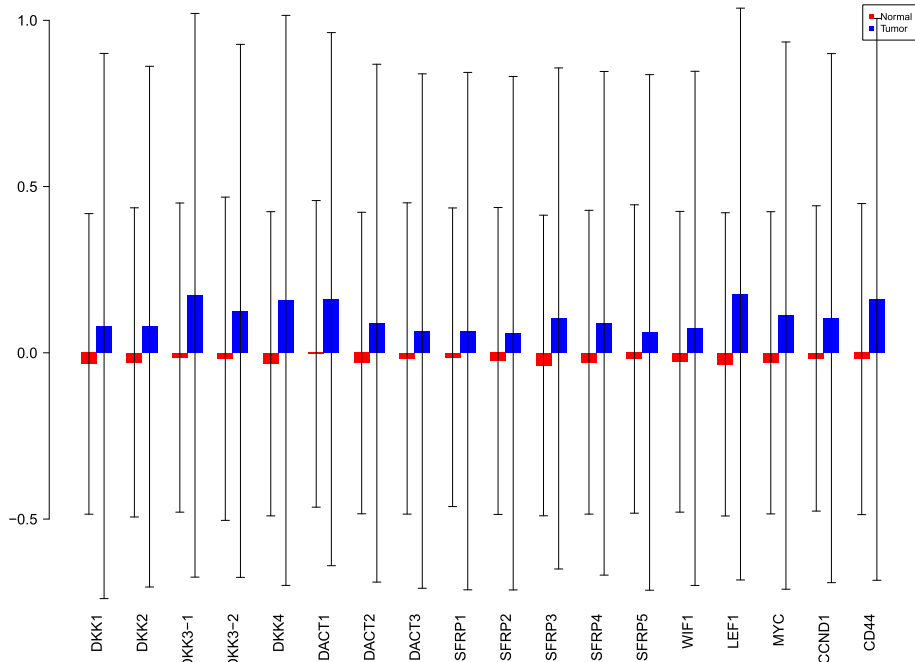

Fig. 13 Sobol 2002 first order indices. Red - indices for normal. Blue - indices for tumor

Fig. 14 Sobol 2007 first order indices. Red - indices for normal. Blue - indices for tumor

Fig. 15 Sobol jansen first order indices. Red - indices for normal. Blue - indices for tumor
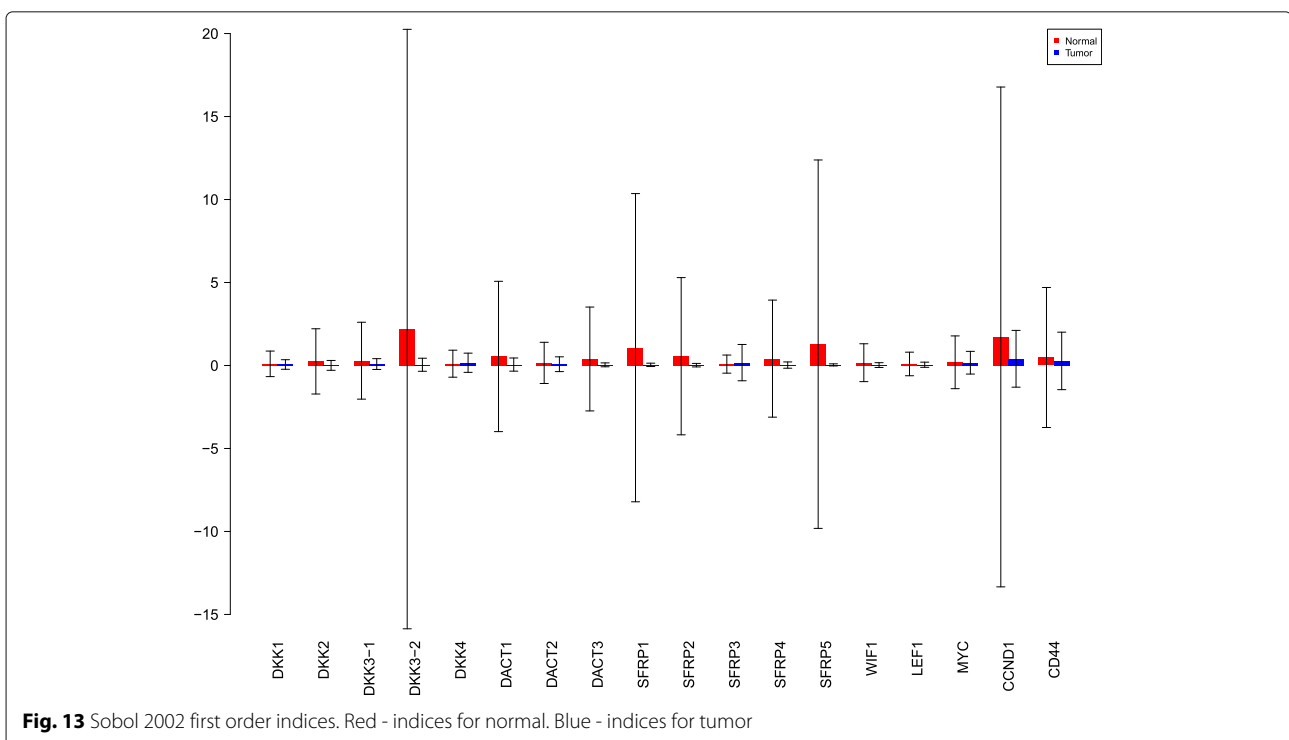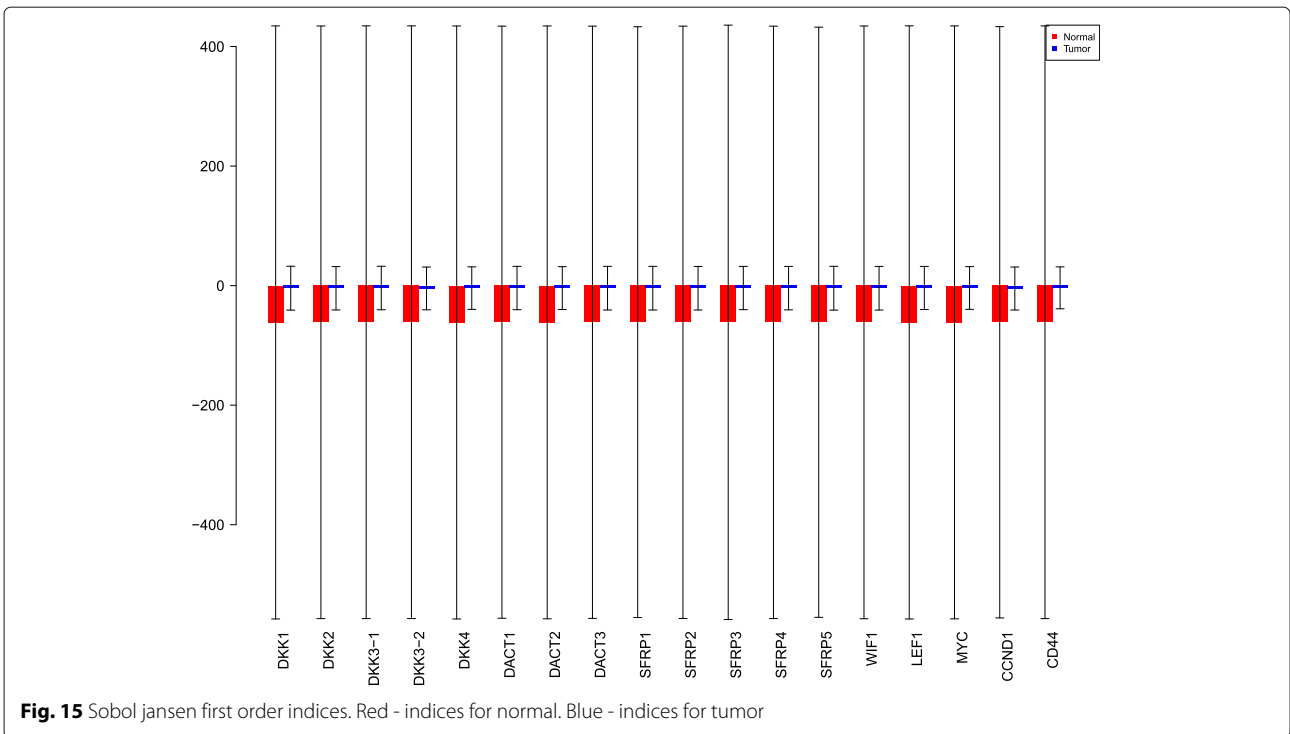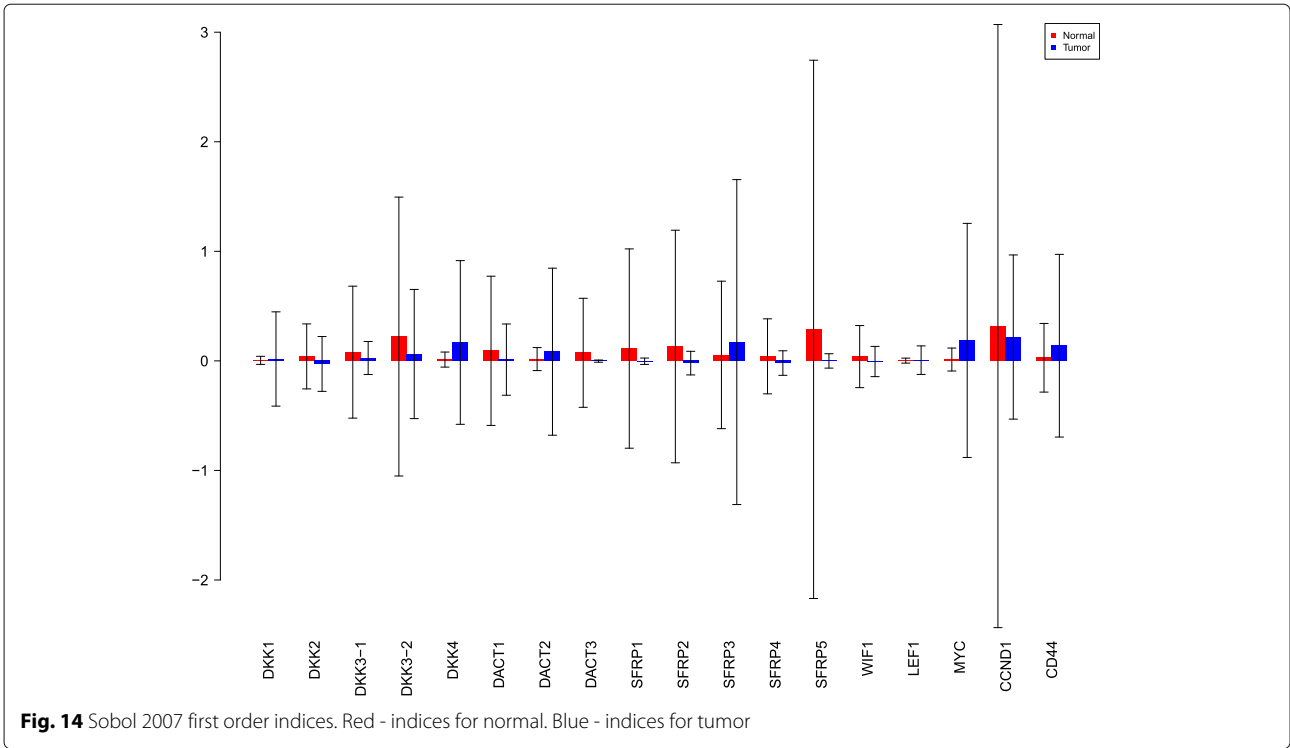
Fig. 16 Sobol martinez first order indices. Red - indices for normal. Blue - indices for tumor
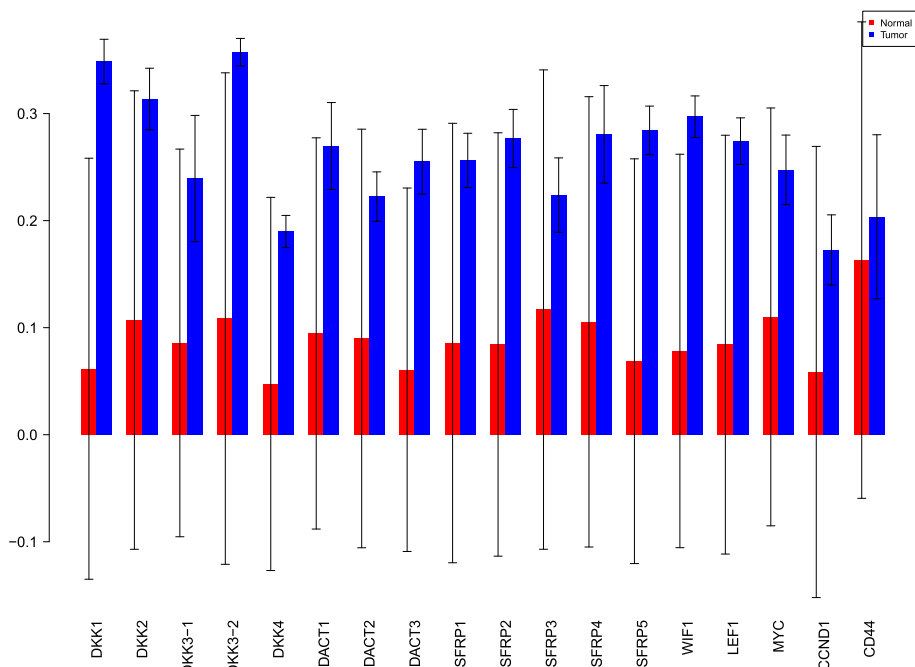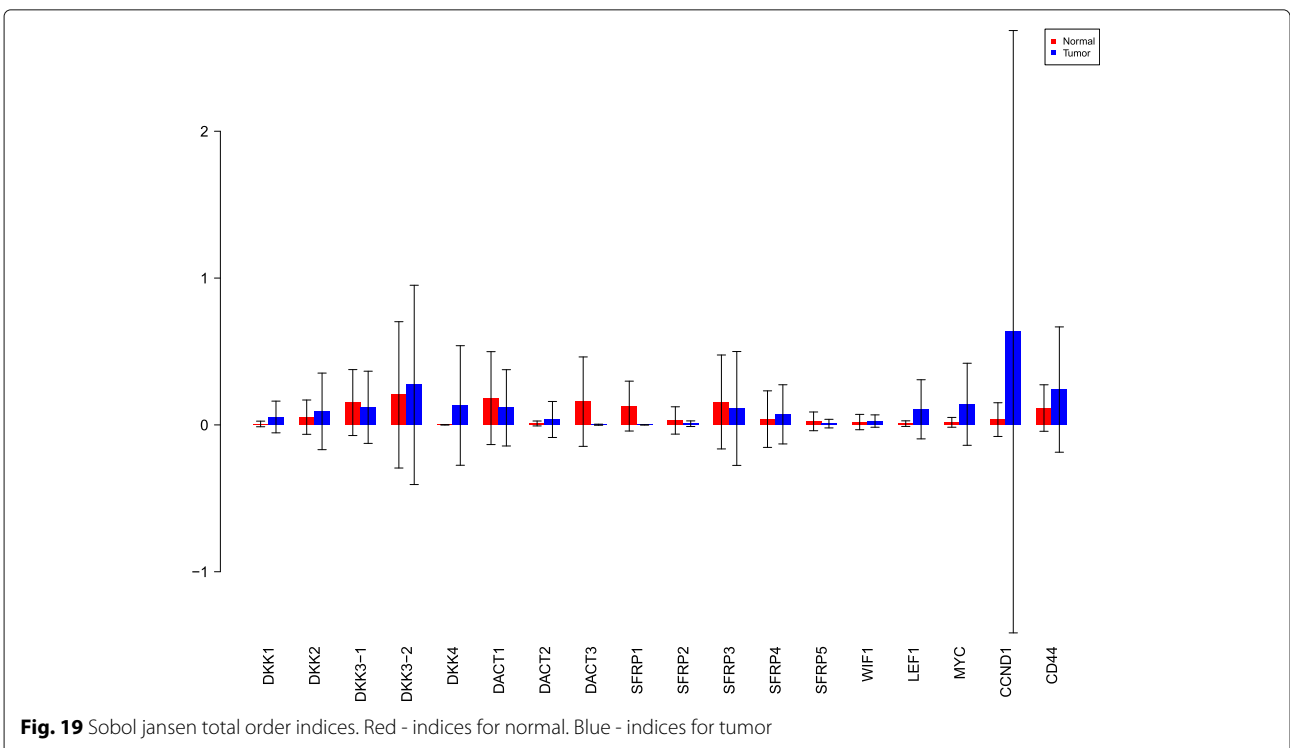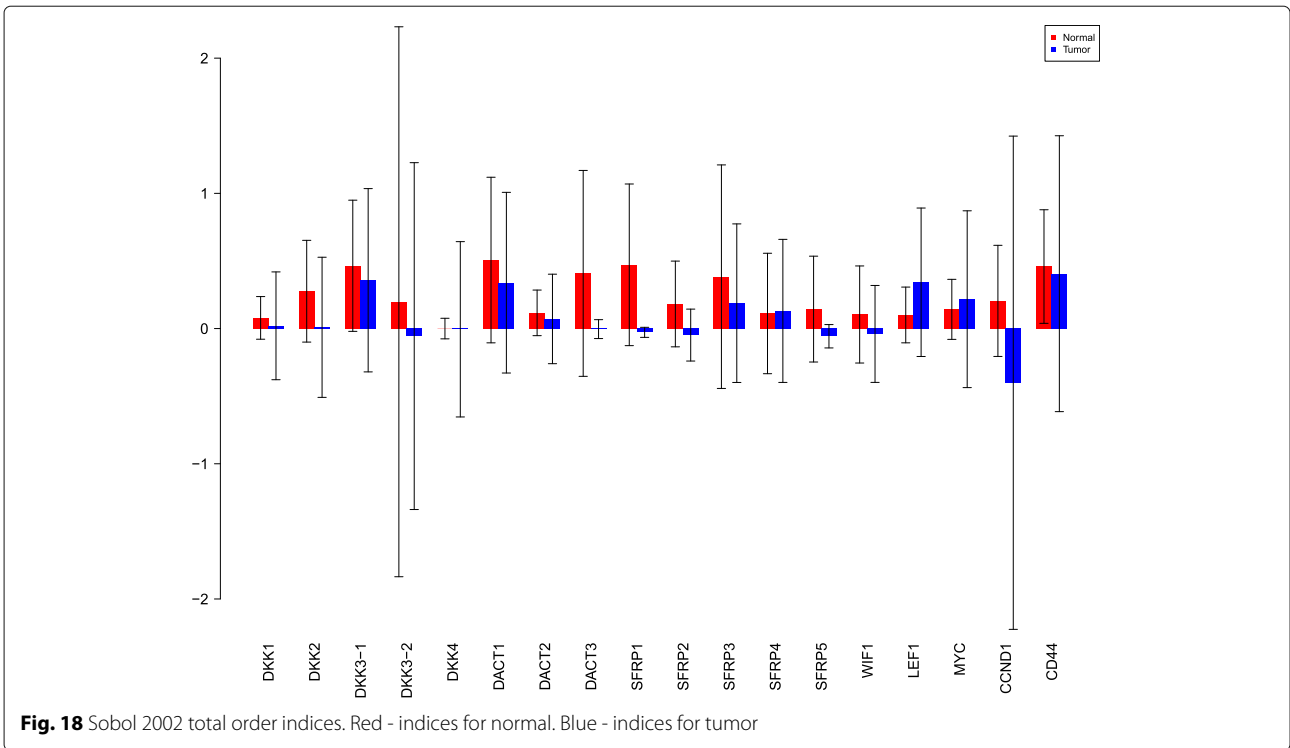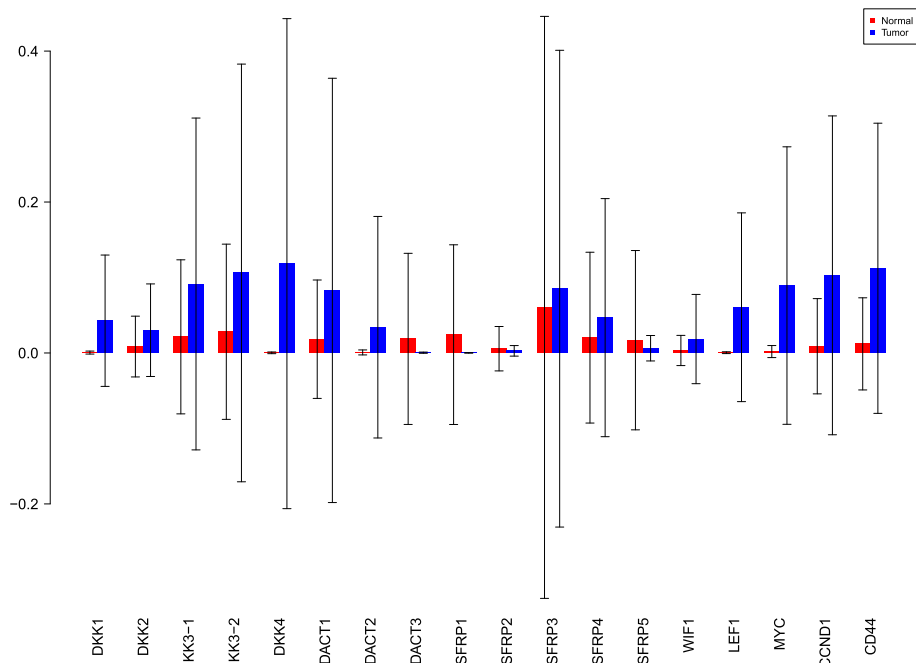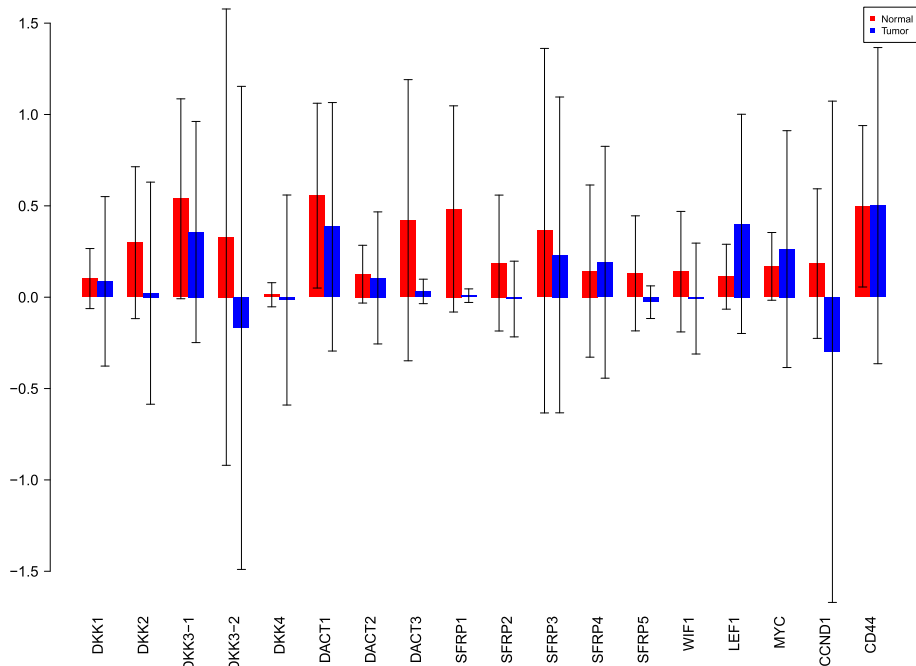

Fig. 17 Sobol first order indices. Red - indices for normal. Blue - indices for tumor

Fig. 18 Sobol 2002 total order indices. Red - indices for normal. Blue - indices for tumor

Fig. 19 Sobol jansen total order indices. Red - indices for normal. Blue - indices for tumor

Fig. 20 Sobol martinez total order indices. Red - indices for normal. Blue - indices for tumor

Fig. 21 Sobol 2007 total order indices. Red - indices for normal. Blue - indices for tumor

results as sensiHSIC). This is attributed to the analytical strength provided by the distance measures using the kernel trick via RKHS that capture nonlinear relations in higher dimensional space, more precisely. Finally, in a recent unpublished work by [87], it has been validated that the HSIC indices prove to be more sensitive to the global behaviour than the Sobol indices.

Time series data

Next, the analysis of the time series data is addressed using the sensitivity indices. THERE ARE TWO EXPERIMENTS THAT HAVE BEEN PERFORMED. FIRST IS RELATED TO THE ANALYSIS OF A PAIR OF THE FOLD CHANGES RECORDED AT TWO DIFFERENT CONSECUTIVE TIME POINTS I.E t_i & t_{i+1}. THE SECOND IS RELATED TO THE ANALYSIS OF A PAIR OF DEVIATIONS IN FOLD CHANGES RECORDED AT t_i & t_{i+1} AND t_{i+1} & t_{i+2}. The former compares the measurements in time while the latter takes into account the deviations that happens in time. For each measurement at a time point a normal distribution was generated with original recorded value as the mean, a standard deviation of 0.005 and an added noise in the form of jitter (see function jitter in R langauge). For the time measurements of each of the genes recorded in [27] an analysis of the sensitivity indices for both the scaled and the non-scaled data was done. Here the analysis for non-scaled data is presented. The reason for not presenting the scaled data is that the sample measurements did not vary drastically as found in the case of static data which caused troubles in the estimation of indices earlier. Another reason for not reporting the results on the scaled data is that the non-scaled ones present raw sensitive information which might be lost in scaling via normalization. Note that [27] uses self organizing maps (SOM) to cluster data and use correlational analysis to derive their conclusions. In this work, the idea of clustering is abandoned and sensitivity indices are estimated for recorded factors participating in the pathway. Also the simple correlational analysis is dropped in favour of highly analytical kernel based distance measures which easily capture the nonlinearities inherent in the data. Figure 22 represents the experimental setup in a pictorial format.
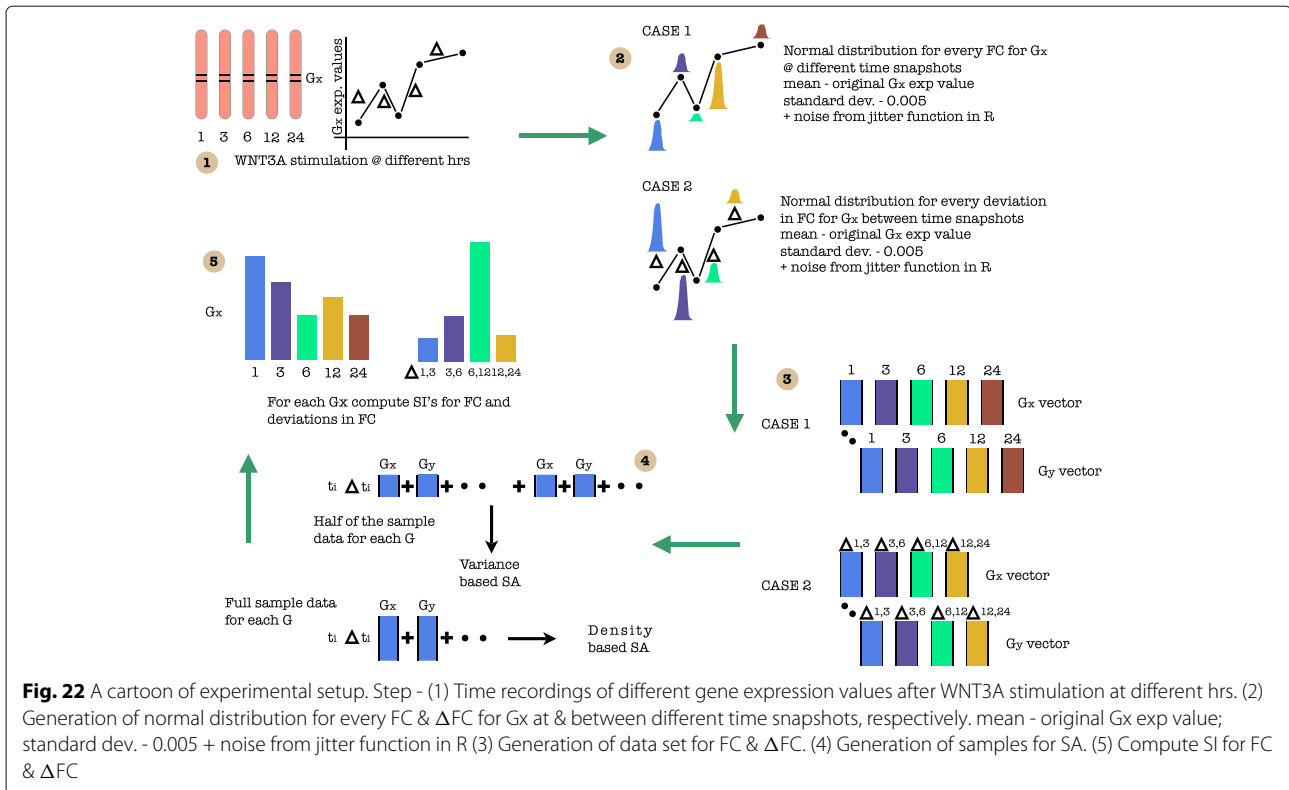
Also, in a recent development, [25] point to two findings namely, • the robust fold changes of *β-catenin* and • the transcriptional machinery of the Wnt pathway depends on the fold changes in *β-catenin* instead of absolute levels of the same and some gene transcription networks must respond to fold changes in signals according to the Weber's law in sensory physiology. The second study also carries a weight in the fact that due to the study of the deviations in the fold changes it is possible to check if the recently observed and reported natural psychophysical laws in the signaling pathway hold true or not. Finally, using the sensitivity indicies an effort is made to confirm

the existing biological causal relations that have been shown in [23].

Analysis of fold changes at different time points

Lets begin with the gene *WNT3A* as changes in its concentration lead to recording of the measurements of the different genes by [27]. Of the list of genes recorded, the indices of the those which are influenced by the concentration of *WNT3A* are analysed. Next based on these confirmations and patterns of indices over time, conclusions for other enlisted genes are drawn. For the former list, the following genes *FZD1, FZD2, LEF1, TCF7, TCF7L1, LRP6, DVL1, SFRP1, SFRP4, CTBP1, CTBP2, PORCN, GSK3β, MYC, APC* and *CTNNB1* are considered. Figures 23 and 24 represent the indices computed over time. Columns represent the different kinds of indices computed while the rows show the respective genes. Each graph contains the sensitivity index computed at a particular time point (represented by a coloured bar). It should be observed from the aforementioned figures that the variants of the Sobol first order (FO) and the total order (TO) indices computed under different formulations were not very informative. This can be seen in graphs were some indices are negative and at some places the behaviour across time and genes remain the same. In contrast to this, the indices generated via the original Sobol function (under the column Sobol-SBL) as well as the sensiHSIC were found to be more reliable. Again, the rbf and laplace kernels under the HSIC formulations showed similar behaviour in comparison to the use of the linear kernel.

Gujral and MacBeath [27] simulate the serum starved HEK293 cells with 200 ng/mL of *WNT3A* at different lengths of time. After the first hour (t_1), (under HSIC-rbf/laplace) it was observed that the sensitivity of *WNT3A* was low (red bar). The maximum contribution of *WNT3A* can be recoreded after the 12th stimulation. But due to increased stimulation by *WNT3A* later on, there is an increased sensitivity of *FZD-1/2* as well as *LRP6*. The *FZD* or the frizzled family of 7-transmember protein [88] works in tandem with *LRP-5/6* as binding parameters for the Wnt ligands to initiate the Wnt signaling. Consistent with the findings of [89] and [90], *FZD1* was found to be expressed. But there is a fair decrease in the contribution of the same in the next two time frames i.e after 3rd and the 6th hour. The maximum contribution of *FZD1* is found after the *WNT3A* simulation at 12th hour. This probably points to repetitive involvement of *FZD1* after a certain period of time to initiate the working of the signaling pathway. *FZD2* showed increasing significance in contribution after the first two time frames. The contribution drops significantly after the 3rd simulation and gradually increases in the next two time frames. The repetitive behaviour is similar to *FZD1*, yet it's role is not

Fig. 22 A cartoon of experimental setup. Step - (1) Time recordings of different gene expression values after WNT3A stimulation at different hrs. (2) Generation of normal distribution for every FC & ΔFC for Gx at & between different time snapshots, respectively. mean - original Gx exp value; standard dev. - 0.005 + noise from jitter function in R (3) Generation of data set for FC & ΔFC. (4) Generation of samples for SA. (5) Compute SI for FC & ΔFC

well studied as it appears to bind to both *WNT3A* which promotes *Wnt/beta-catenin* signaling and *WNT5A* which inhibits it as shown by [91], respectively.

Klapholz-Brown et al. [92] and [93] show that there is increased *β-catenin* due to *WNT3A* stimulation which is depicted by the increased sensitivity of *CTNNB*1 expression in one of the above mentioned figures. *MYC* (i.e

$c - MYC$) is known to be over expressed in colorectal cancer cases mainly due to the activation of $TCF - 4$ transcription factor via intra nuclear binding of *β-catenin* [94], either by *APC* mutations [95] or *β-catenin* mutations [96]. The sensitivity of *MYC* increased monotonically but after the 6th h it dropped significantly. Probably *MYC* does not play important role at later stages. As found in [97] and
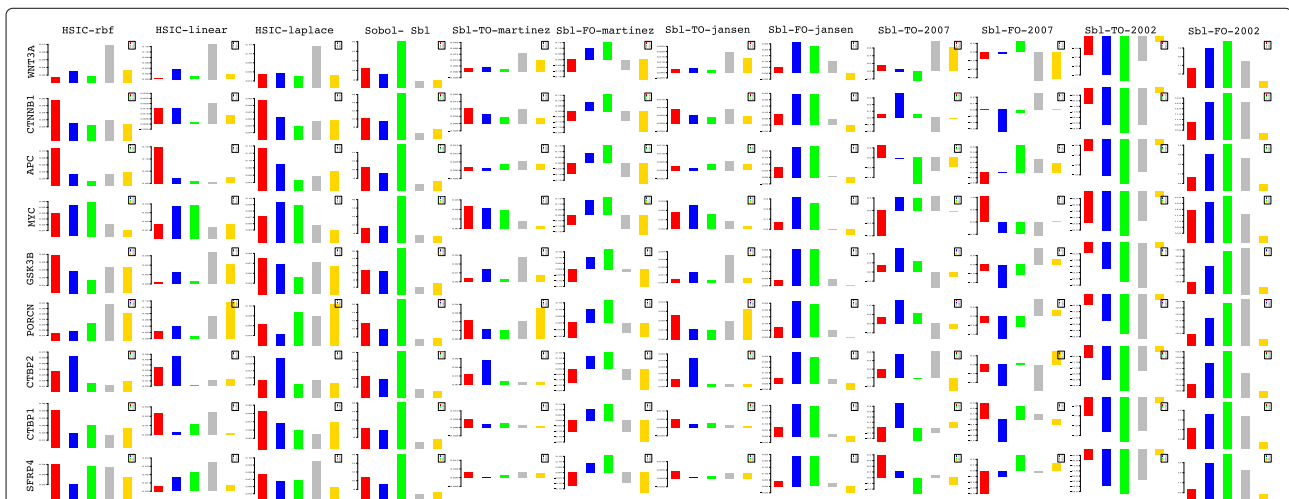
Fig. 23 Column wise - methods to estimate sensitivity indices. Row wise - sensitivity indicies for each gene. For each graph, the bars represent sensitivity indices computed at t1 (red), t2 (blue), t3 (green), t4 (gray) and t5 (yellow). Indices were computed using non scaled time series data. TO - total order; FO - first order; SBL - Sobol

Fig. 24 Column wise - methods to estimate sensitivity indices. Row wise - sensitivity indicies for each gene. For each graph, the bars represent sensitivity indices computed at t1 (red), t2 (blue), t3 (green), t4 (gray) and t5 (yellow). Indices were computed using non scaled time series data. TO - total order; FO - first order; SBL - Sobol
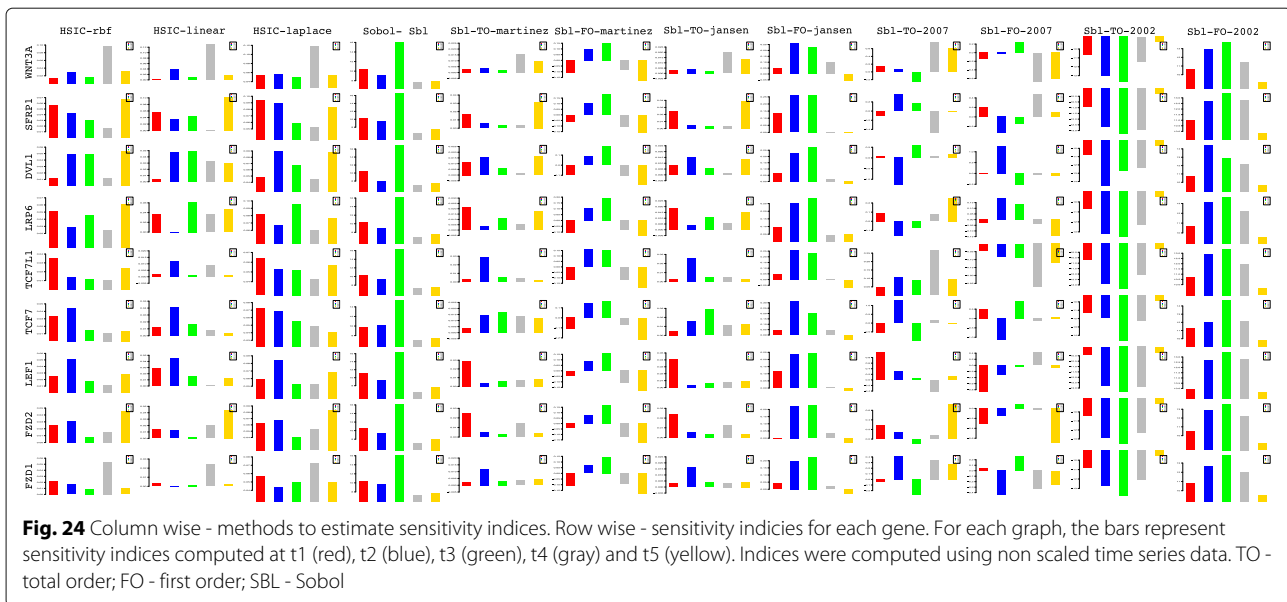
[98], *DVL* family interacts with the frizzled *FZD* members leading to disassembly of the *β-catenin* destruction complex and subsequent translocation of *β-catenin* to the nucleus. Development on *DVL* family have been extensively recorded in [99] and [100], and significance of *DVL*1 in Taiwanese colorectal cancer in [101]. *DVL*1 shows a marked increase in sensitivity as the concentration of the *WNT*3A increases in time. This is supported by the fact that ligand binding at the membrane leads to formation of complex including *DVL*1, *FZD* and *AXIN*.

Negative regulators like *SFRP*4 were found to have lower sensitivity as *WNT*3A concentration increases, but remained constant for most period. Meanwhile the significance of Wnt antagonist *SFRP*1 ([102], [103] and [104]) decreases over the period as the concentration of *WNT*3A increases. [105] reviews the co-repressor ability of the *CTBP* family, while [106] shows *CTBP* as a binding factor that interacts with *APC* thus lowering the availability of free nuclear *β-catenin*. This interaction is further confirmed in the recent research work by [107]. As shown by [93] *CTBP*1 showed increased sensitivity with increased stimulation of *WNT*3A in the first hour. The latter stages show a decreased contribution of *CTBP*1 as the concentration of *WNT*3A was increased. This is in line with what [27] show in their manuscript and indicate the lowering of the co-repressor effect of *CTBP* at later stages. On the other hand, *CTBP*2 showed reverse behaviour of sensitivity in comparison to *CTBP*1 across different time points. Increased significance of *CTBP*2 was observed in the first two time frames, i.e after 1st and 3rd hour of stimulation, followed by lower contribution to the pathway at the latter stages. In both cases, the diminishing co-repressive nature of *CTBP* in time is observed. Contrary to these finding, recent results in [108] suggest

that both *CTBP*1 and *CTBP*2 are up-regulated in colon cancer stem cells.

PORCN showed less sensitivity in the initial stages than in final stages indicating its importance in the contribution to Wnt secretion which is necessary for signaling [109]. The sensitivity of *GSK*3*β* and *APC* decreased in time indicating the lowering of its significance in later stages due to no formation of the degradation complex. Activity of *TCF* gains greater prominence in the first and the second time frames after the initial *WNT*3A stimulation. This is in conjugation with the pattern showed by *CTBP*2. Regarding *TCF*7*L*2, the activity is observed to be maximum during the first time frame with decrease in the contribution in the later time frames.

Indicies for remaining 57 genes as well as analysis of the same will be presented in the following B part of this manuscript. Graphs for these 57 genes have been presented in Figures 27 and 28 in Appendix.

Analysis of deviations in fold changes

In comparison to the contributions estimated via the sensitivity indices using fold changes at different time points separately, this section analyzes the contributions due to deviations in the fold changes recorded between two time points i.e t_i & t_{i+1}. These analyses are also a way to test the efficacy of deviations in fold changes versus the absolute levels that have been stressed upon in [25]. I PRESENT HERE A DETAILED ANALYSIS OF HOW SOME OF THE COMPONENTS OF THE PATHWAY INFLUENCE THE PATHWAY AT WHICH TIME POINT AND TIME PERIOD. OF THE EXPRESSION PROFILES OF 71 GENES RECORDED DURING THE STIMULATION, I PRESENT ONLY A FEW OF THEM AS AN EXAMPLE OF SYSTEM WIDE ANALYSIS AT THE COMPUTATIONAL

LEVEL. NOTE THAT IT IS THESE TIME PERIODS WHICH HAVE BEEN IDENTIFIED BY OBSERVING THE NUMERICAL VALUES OF THE SENSITIVITY INDICES IN WHICH THE INFLUENCE OF A PARTICULAR FACTOR/COMPONENT IS REPORTED. IDENTIFICATION OF EARLY TIME POINTS AND TIME PERIODS INDICATE HOW THE PATHWAY IS AFFECTED AT AN EARLY STAGE AND VICE VERSA. THUS, THE SYSTEM WIDE ANALYSIS CONDUCTED AT TIME COURSE LEVEL, GIVES A DEEPER PICTURE OF THE INFLUENCES OF THE COMPONENTS IN THE WORKING OF THE PATHWAY. Some of these findings are in line with [27]. But some provide more deeper analysis where [27] fails to do so.

In such an already extensive computational study at systems level, it is not possible to provide wet lab tests to validate the above findings and neither does the author currently have the resources to test the same in a wet lab setting. INSTEAD, SUCH INSILICO FINDINGS FACILITATE THE BIOLOGISTS TO VERIFY AND STUDY THE PATHWAY DEEPLY IN WET LAB. IN THE AUTHOR'S LIMITED AWARENESS, SUCH A STUDY HAS NOT BEEN UNDERTAKEN BEFORE. FINALLY, THE FOLLOWING DESCRIPTIONS JUST DO NOT EXPLAIN THE GRAPH. THEY RATHER POINT TO THE TAKE HOME MESSAGES REGARDING THE INFLUENCE OF FACTORS/COMPONENTS AT DIFFERENT TIME POINTS AND TIME PERIODS AT INSILICO LEVEL.

As with the analysis of the fold changes at different time points, the estimates obtained using the rbf, linear and laplacian kernels in the HSIC based sensitivity analysis have been used here. Of these, the rbf and laplacian kernels give almost similar results. Plots of the time series expression profiles from [27] have been relegated to the Appendix and shown in Figures 31 and 32.

■ WNT3A - Figure 31 in Appendix shows the profile of mRNA expression levels of *WNT3A* after external stimulation. There is a series of $(+ - ++)$ deviations in the fold change recordings at different time points. A repetitive behaviour is observed in the contribution of the deviations in fold changes for *WNT3A* estimated via the sensitivity indices. For intervals in t_1, t_3 and t_6 there is an increase in the significance of the contribution of *WNT3A* in Fig. 25 (see the first two bars for $< t_1, t_3 >$ & $< t_3, t_6 >$), even though in the first three time frames levels of *WNT3A* are shown to be down-regulated (see Figure 31 in Appendix). This behaviour is again repeated in Fig. 25 for intervals t_6, t_{12} and t_{24} (see the next two bars for $< t_6, t_{12} >$ & $< t_{12}, t_{24} >$). In both cases one finds an increase in the contribution of the deviation in the fold change. Comparing the contribution of levels of fold changes in Fig. 23 were it was found that there is a dip in contribution of *WNT3A* after t_3 and then a further increase in the contribution at a latter time frame, one finds that the deviations in fold changes involving $< t_1, t_3 >$ & $< t_3, t_6 >$ have higher

significance than the deviations in fold changes involving $< t_6, t_{12} >$ & $< t_{12}, t_{24} >$. It can be noted that even in the deregulated state from $< t_1$, t_3 and $t_6 >$ the deviations are minimal and the contributions are significantly high. In case of the regulated states from $< t_6$, t_{12} and $t_{24} >$ the deviation is extremely high between the first two time frames and low in the next two. This results in greater significant contribution in the latter deviation than the former deviation. Thus when deviations are low and the fold changes over time do not vary much, the contributions of the involved factor to the signaling pathway is expected to be high and vice versa. This points to the fact that low variations in fold changes over time have a stablizing influence of *WNT3A* rather than abrupt high variations in fold changes that might not have the same influence. Thus measurments of deviations in fold changes provide greater support for studying the affect of *WNT3A* over time.

■ CTNNB1 - Figure 31 in Appendix shows the profile of mRNA expression levels of *CTNNB*1 after external stimulation. There is a series of $(+ + -+)$ deviations in the fold change recordings at different time points. An initial increase in the influence of *CTNNB*1 is observed from $< t_1, t_3 >$ to $< t_3, t_6 >$ (first two bars in Fig. 25) followed by a gradual decrease of influence from $< t_3, t_6 >$ to $< t_6, t_{12} >$ to $< t_{12}, t_{24} >$ (last three bars in Fig. 25). This is observed even though there is an up-regulation in levels of *CTNNB*1 with a slight dip at t_{12} (see Figure 31 in Appendix). In comparison to the contribution of levels of fold changes in Fig. 23 were it was found that there is a gradual decrease in the influence of *CTNNB*1 till t_6 and then a further increase in the contribution at latter time frames, one finds that the deviations in fold changes involving $< t_3, t_6 >$ have the highest significance with an almost exponential decrease in the deviations in fold changes involving $< t_6, t_{12} >$ & $< t_{12}, t_{24} >$. Even though in a regulated state the influence of deviations in the fold changes indicate a different scenario altogether in comparison to influences of fold changes at distinct time frames. This might point to the fact that the affect of *CTNNB*1 is maximum during $< t_3, t_6 >$ in comparison to other stages even after constantly increasing external stimulation with *WNT3A* at different time points. Exponential decrease in the influence in the deviations in latter time frames points to the ineffectiveness of *CTNNB*1 in the pathway at later stages. Finally, in contrast to the behaviour of influence of *WNT3A* in the foregoing paragraph, *CTNNB*1 showed higher (lower) influence for greater (lesser) deviations in fold changes.

■ APC - Figure 31 in Appendix shows the profile of mRNA expression levels of *APC* after external stimulation. The profile of the deviations of *APC* in a deregulated state show the following $(- + +-)$ pattern. While the *CTNNB*1 expression profile shows non-monotonic

Fig. 25 Column wise - methods to estimate sensitivity indices. Row wise - sensitivity indicies for each gene on deviations in fold change. For each graph, the bars represent sensitivity indices computed at $< t1, t2 >$ (red), $< t2, t3 >$ (blue), $< t3, t4 >$ (green) and $< t4, t5 >$ (gray). Indices were computed using non scaled time series data. TO - total order; FO - first order; SBL - Sobol
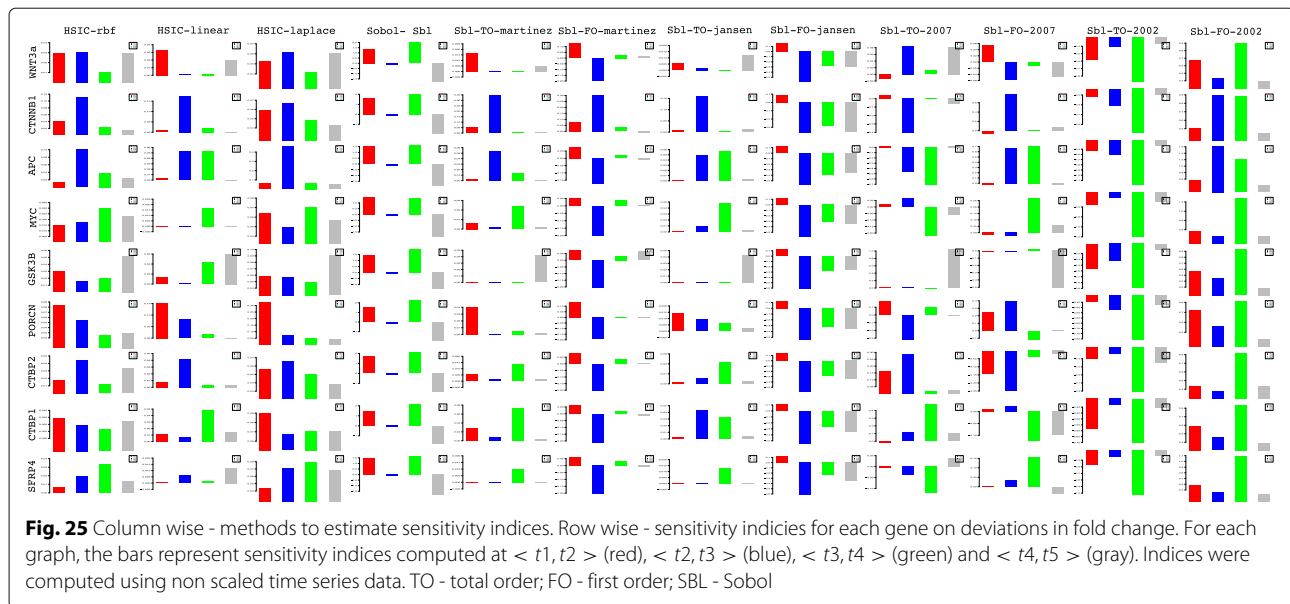
increase in levels of fold changes in upregulated state, *APC* expression profile shows a nonlinear behaviour in levels of fold changes in down regulated state. The significance of deviation in fold changes for *APC* is maximum during $< t_3, t_6 >$ when the downregulation is weakened. Further weaking of the downregulation during $< t_6, t_{12} >$ does not have much significance. This attenuation in significance of deviations in fold change might support the fact that *APC*'s weaking in downregulation amplifies shutting down of the Wnt pathway after the intial strong downregulation (where Wnt activity is high). This is corroborated by the finding of [27] which observes the initial (later) positive (negative) feedback that strenghtens (weakens) the Wnt pathway activity. An initial increase in the influence of *APC* is observed from $< t_1, t_3 >$ to $< t_3, t_6 >$ (first two bars in Fig. 25) followed by a gradual decrease of influence from $< t_3, t_6 >$ to $< t_6, t_{12} >$ to $< t_{12}, t_{24} >$ (last three bars in Fig. 25). This is observed even though there is an down-regulation in levels of *APC* with slight weaking at t_6, t_{12} and t_{24} in comparison to recordings at other time frames (see Figure 31 in Appendix). In comparison to the contribution of levels of fold changes in Fig. 23 where it was found that there is a gradual decrease in the influence of *APC* till t_6 and then a further increase in the contribution at latter time frames, one finds that the deviations in fold changes involving $< t_3, t_6 >$ have the highest significance with an almost exponential decrease in the deviations in fold changes involving $< t_6, t_{12} >$ &$< t_{12}, t_{24} >$.

■ **MYC** - Figure 31 in Appendix shows the profile of mRNA expression levels of *MYC* after external stimulation. The profile of the deviations in fold changes of *MYC* in an up-regulated state show the following $(- + ++)$ pattern. After an initial dip in the up-regulation at

t_6 there is an exponential increase in the fold changes of *MYC* as time progresses. While Fig. 23 shows an increasing sensitivity of *MYC* for the first three time frames, later up-regulated state of *MYC* due to increasing *WNT*3*A* stimulations do not hold much significance. In contrast, it is not possible to observe a pattern in the sensitivity of deviations in fold changes for *MYC* except for the fact that the maximum contribution of deviation in fold change is observed for the period of $< t_6, t_{12} >$. This is a period when *MYC*'s significance in the pathway is maximum.

■ **GSK3B** - Figure 31 in Appendix shows the profile of mRNA expression levels of $GSK3\beta$ after external stimulation. The profile of the deviations in fold changes of $GSK3\beta$ in an varied regulated state show the following $(- + ++)$ pattern. After an initial up-regulation at t_3 there is down regulation at t_6 before which up-regulation follows for latter stages. It is widely known that *WNT* stimulation leads to inhibition of $GSK3\beta$. In contrast to this regard $GSK3\beta$ shows a up-regulated levels at t_3, t_{12} and t_{24}. The author is currently unaware of why this contasting behaviour is exhibited. Later upregulation might point to the fact that the effectiveness of Wnt stimulation has decreased and $GSK3\beta$ plays the role of stabilizing and controlling the behaviour of the pathway by working against the Wnt stimulation and preventing further degradation. While work by [27] does not shed light on this aspect, contrasting models of inhibitions for $GSK3$ has been recently proposed in [110] which might support this behaviour. Figure 23 shows an decreasing sensitivity of $GSK3\beta$ for the first two time frames, after which there is an increasing sensitivity for the next three time frames. Comparing this with plots in Fig. 25 it is found that there is greater significance of deviations in fold changes of $GSK3\beta$ during later stages of $< t_6, t_{12} >$ and $< t_{12}, t_{24} >$.

■ **PORCN** - Figure 31 in Appendix shows the profile of mRNA expression levels of *PORCN* after external stimulation. The profile of the deviations in fold changes of *PORCN* in an up regulated state show the following $(+ - --)$ pattern. After an initial hike in up-regulation at t_3 there is continuous decrease in the up regulation. *PORCN* is known to help in the secretion of the Wnt ligands that later on help in the instigation of the signaling activity [111]. Sustanined stimulation by *WNT*3A over a period of time might lead to decrease in the up regulation of *PORCN* which helps in Wnt secretion. Graph for *PORCN* in Fig. 23 shows increasing significance of the influence of *PORCN* as time passes, even though there is lower regulation of the same at later stages (Figure 31 in Appendix). The highly significant influence of lower regulation at later stages indicates the lessened effectiveness of *PORCN* due to sustained *WNT*3A stimulation that might have suppressed the functionality of secretion carried out via *PORCN*. Contrary to this, the influences of the deviations in the fold changes over time show the reverse behaviour. The maximum influence is during the first two time frames of $< t_1, t_3 >$ and this influence of deviations decreases at later stages. This points to the fact that the deviations in the fold changes at intial stage has greater significance in the pathway than the deviations at later stages. It follows that in initial stages of Wnt stimulation the expression of *PORCN* has significant influence.

■ **CTBP2** - Figure 31 in Appendix shows the profile of mRNA expression levels of *CTBP*2 after external stimulation. The profile of the deviations in fold changes of *CTBP*2 in an up regulated state show the following $(- + + -)$ pattern. It is known that *CTBP*2 shows co-repressive nature [112] and the pattern of sensitivity indicates this heigthened effect at $< t_3, t_6 >$ and $< t_{12}, t_{24} >$. In contrast to this in Fig. 23 one finds that the significance of upregulation at t_{12} and t_{24} is minimal and yet the sensitivity for the deviations in fold changes for this period is second to $< t_3, t_6 >$. The probable explanation for this might be that for higher upregulation (in terms of numeric representations), even small deviations might play a sigificant role while the sensitivity at individual time frames remains low.
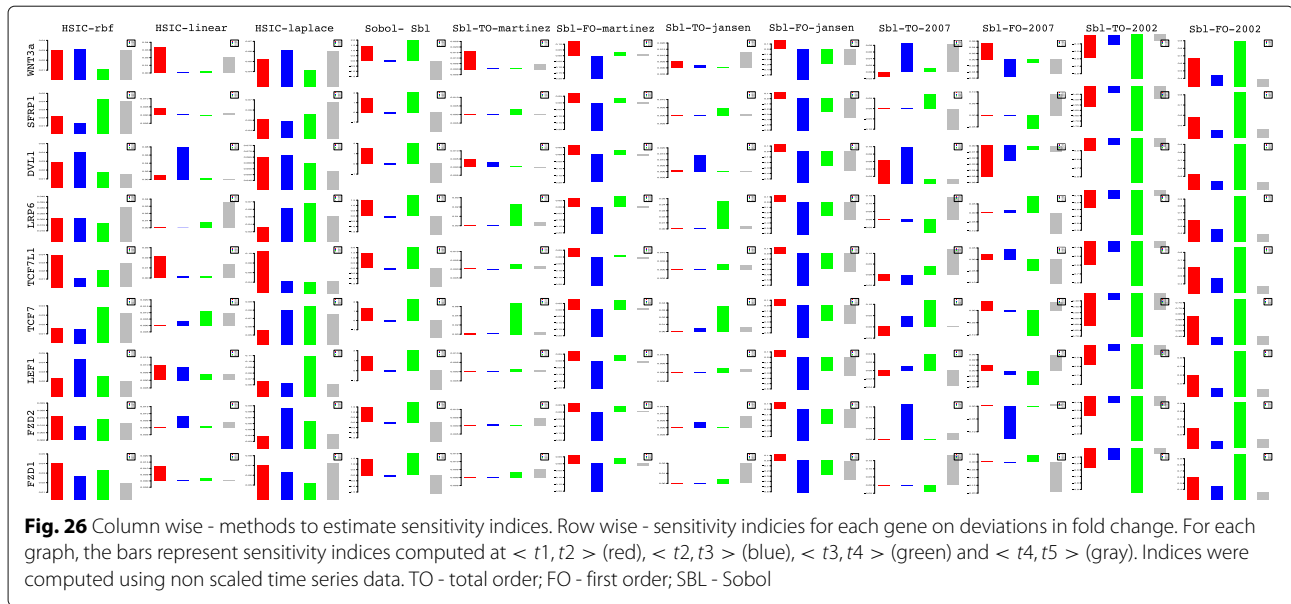
■ **CTBP1** - Figure 31 in Appendix shows the profile of mRNA expression levels of *CTBP*1 after external stimulation. The profile of the deviations in fold changes of *CTBP*1 in an up regulated state show the following $(+ + + -)$ pattern. As with the heightened sensitivity at time frame t_1 the sensitivity of deviations in the fold changes exhibits heightened effect in the pathway at $< t_1, t_3 >$. Further analysis might not be possible as one finds lowered sensitivity even at heightened up regulation for individual recordings as well as deviations.

■ **SFRP4** - Figure 31 in Appendix shows the profile of mRNA expression levels of *SFRP*4 after external stimulation. The profile of the deviations in fold changes of *SFRP*4

in an down regulated (except in the last time frame) state show the following $(- + ++)$ pattern. Known to be a negative regulator of the Wnt pathway, it was found that its sensitivity in extremely high as it is down regulated during the stimulation. This is depicted in the figures plotted in Fig. 23. This heightened sensitivity during most of the down regulation points towards the significant role of hypermethylation that leads to silencing of this gene. Contrary to this, there is a monotonically increasing sensitivity for deviations in the fold changes during down regulation from t_1 to t_{12}. A dip in the sensitivity of the deviation for the final time frame $< t_{12}, t_{24} >$ happens when the up regulation is recorded in the last time frame. Based on the maximum sensitivity for deviation in fold change during $< t_6, t_{12} >$, up regulation of *SFRP*4 in this period is expected to have greatest reverse effect on the activation of the Wnt pathway. It appears that the hypermethylation that causes the silencing of the *SFRP*4 is maximal during this stage and thus a potential period for the pathway to be inhibited via reversal of silencing [103].

■ **SFRP1** - Figure 32 in Appendix shows the profile of mRNA expression levels of *SFRP*1 after external stimulation. The profile of the deviations in fold changes of *SFRP*4 in an up regulated (except in during the second time frame) state show the following $(- + --)$ pattern. It is widely known that *SFRP*1 is a Wnt antagonist and is known for inactivation in the canconical Wnt pathway due to hypermethylation thus leading to upregulation of the pathway [104]. [103] further indicates that *SFRP*1 is thought to silence ligand-dependent Wnt signaling by binding of the cysteine rich-domain (CRD) to Wnt proteins, thus preventing interaction with *FRZ* receptors. Recent in silico results by [113] confirm hypermethylation of *SFRP*1 in colorectal cancers. Given the above profile, it is possible to see that there is a down regulation at t_3 but the significance of its influence on the pathway is not much as revealed in Fig 24. Figure 24 shows a decreasing significance in the influence of the *SFRP*1 with the maximum influence in the last stage of the *WNT*3A stimulation, where there is an up regulation. In similar way, in Fig. 26 there is significance of the sensitivity in the deviations in fold changes during the up regulation of *SFRP*1 during the $< t_6, t_{12} >$ and $< t_{12}, t_{24} >$. The activation at later stages show that *SFRP*1 has a greater antagonistic effect on the Wnt pathway. In comparison to up regulation, one finds the down regulation at t_3 does not play a significant role. These are in line with [27]'s claim regarding reversal of behaviour at different time stages.

■ **DVL1** - Figure 32 in Appendix shows the profile of mRNA expression levels of *DVL*1 after external stimulation. The profile of the deviations in fold changes of *DVL*1 in an up regulated state show the following $(- + + -)$ pattern. *DVL*1 is an adaptor protein that helps in signal transmission that leads to stabilization of cytosolic

Fig. 26 Column wise - methods to estimate sensitivity indices. Row wise - sensitivity indicies for each gene on deviations in fold change. For each graph, the bars represent sensitivity indices computed at $< t1, t2 >$ (red), $< t2, t3 >$ (blue), $< t3, t4 >$ (green) and $< t4, t5 >$ (gray). Indices were computed using non scaled time series data. TO - total order; FO - first order; SBL - Sobol

β*-catenin* for further processing. [101] report high expression of *DVL*1 in Taiwanese colorectal cancer patients with liver metastasis and it has also been observed as a potential biomarker in CRC [114]. In Fig 24 the time frame t_{12} at which *DVL*1 shows maximum up regulation is the most insignificant one due to the lowest sensitivity while moderate upregulation during t_3 and t_6 show high sensitivity. The same is true for upregulation at t_{24}. Comparing this with the deviations in the fold changes in 26, one finds that there is maximum sensitivity during the period of $< t_3, t_6 >$ preceded by a lower sensitivity index for the periood of $< t_1, t_3 >$. At other intervals there was a decreased sensitivity even though the deviations in the fold changes were very high. This indicates that the high deviations might not influence the signaling activity significantly. Also the best period of intervention is at $< t_3, t_6 >$.

■ **LRP6** - Figure 32 in Appendix shows the profile of mRNA expression levels of *Lrp*6 after external stimulation. The profile of the deviations in fold changes of *Lrp*6 in an up regulated state (except at t_3) show the following $(- + + -)$ pattern. In an extensive work on the molecular differences of *LRP*s [115] investigate and show that *LRP*5/6 along with the frizzled family members form a Wnt inducible co-receptor complex that helps in signal transmission after *LRP* phosphorylation. Earlier wet lab work by [116] and in silico findings by [117] have shown highly expressed participation of *LRP*5/6 in Wnt signaling pathway. Latest work by [118] shows that KRAS signaling promotes canonical Wnt activity via *LRP*6. In Fig. 24, *LRP*6 shows significant influences during the t_1, t_6 and t_{12}. The only period in which it is down regulated during t_3 has little significance in comparison to the significance of the up regulated states. It is not known why *LRP*6 shows

down regulation at this stage. Finally, for unkown reasons, the influence of *LRP*6 during t_{12} was found to be the lowest. It was not possible to read into the sensitivity of *LRP*6 for deviations in fold changes the HSIC based indices. The laplace kernel shows a pattern of increasing sensitivity with the heightest during $< t_6, t_{12} >$. But this is not so in the other two formulations. Thus wet lab experiments might aid in confirming these results and shedding more light on the duration during which a drug could be administered.

■ **TCF7L16** - Figure 32 in Appendix shows the profile of mRNA expression levels of *TCF7L*1 (also known as *TCF*3) after external stimulation. The profile of the deviations in fold changes of *TCF7L*1 in a down regulated state (except at t_3) show the following $(+ - + -)$ pattern. It is known that Wnt stimulation promotes the phosphorylation of repressor-acting *TCF7L*1 by homeodomain-interacting protein kinase (HIPK2), which results in its dissociation from the WRE [119]. Gujral and MacBeath [27]'s results also indicate the same repression of *TCF7L*1 during *WNT*3*A* simulation as shown in Figure 32 in Appendix. But in contradiction to this recent findings of [120], *TCF7L*1 is found to be expressed in the colon crypt and in colon cancer. Their results indicate that *TCF7L*1 may have an as yet unidentified role in transmission of tumor-related β*-catenin* signals. Evidence of this up regulation is found at t_3 time period as shown in Figure 32 in Appendix. From Fig. 24 it can be seen that the sensitiviy of *TCF7L*1 is maximum during the first time period. Later on the sensitivity subsides as time passes untill the sensitivity shots up in the last time frame t_{24}. In comparison to this with respect to the deviations in fold changes over time in Fig. 26, $< t_1, t_3 >$ showed the maximum period of influence. Later on there is a drop in the sensitivity which is

followed by an approximately monotonic increase. It is the first transition from down regulation to up regulation $< t_1, t_3 >$ that might be the time for intervention. Also the last stage might be of some value but during down regulation only.

■ TCF7 - Figure 32 in Appendix shows the profile of mRNA expression levels of TCF7 after external stimulation. The profile of the deviations in fold changes of TCF7 in an up regulated state show the following $(- - + +)$ pattern. TCF7 is found to be regulated upon Wnt stimulation as it binds with LEFs to activate the transcription procedure after interacting with β-*catenin* [121]. In Fig. 24, the sensitivity of the activation of TCF7 decreases monotonically as time progresses. But this behaviour is not the same for deviations in the fold changes. The maximum influence is found for the duration of $< t_6, t_{12} >$. The next best consistent influence is in the duration $< t_{12}, t_{24} >$. These are the two time periods when the influence of the deviations in the fold changes is maximum and thus susceptible to therapeutic interference.

■ LEF1 - Figure 32 in Appendix shows the profile of mRNA expression levels of LEF1 after external stimulation. The profile of the deviations in fold changes of LEF1 in an up regulated state (except at t_3) show the following $(- - + +)$ pattern. Generally, LEF1 is found to be up regulated upon Wnt stimulation when it works in tandem with TCF7 [121]. Yet, in Fig. 24, the sensitivity of the activation of LEF1 is not similar to that of TCF7. In contradiction to what is observed, one finds a down regulation during the time period t_3. More importantly this is the period in which the most significant influence of down regulated LEF1 is observed. The Initial down regulation at this subinterval indicates that LEF1 is not facilitating the Wnt pathway positiviely. Conclusive results cannot be stated regarding the deviations in fold changes from Fig. 26.

■ FZD2 - Figure 32 in Appendix shows the profile of mRNA expression levels of FZD2 after external stimulation. The profile of the deviations in fold changes of FZD2 in an up regulated state (except at t_3 and t_6) show the following $(- - + +)$ pattern. The FZD or the frizzled family of 7-transmember protein [88] works in tandem with LRP-5/6 as binding parameters for the Wnt ligands to initiate the Wnt signaling. In comparison to the repetitive behavior shown in Fig. 24 it is not possible to draw conclusions on the deviations in fold changes.

■ FZD1 - Figure 32 in Appendix shows the profile of mRNA expression levels of FZD1 after external stimulation. The profile of the deviations in fold changes of FZD1 in an down regulated state (except at t_3) show the following $(+ - - +)$ pattern. Consistent with the findings of [89] and [90], FZD1 was found to be expressed at t_3. In the rest of the time periods, it was down regulated. But the significance of the influence shows a different pattern in Fig. 24 with the down regulation at t_{12} being the most

influencial. In contrast to this, while observing the deviations in the fold changes, it was found that the first two durations $< t_1, t_3 >$ and $< t_3, t_6 >$ showed consistent decreasing behaviour in terms of influence. It is during the first period that the deviations in fold changes are significant and thus it is possible to intervene therapeutically during the activation stage. Indicies for remaining 57 genes as well as analysis of the same will be presented in the following B part of this manuscript. Graphs for these 57 genes have been presented in Figures 29 and 30 in the Appendix.

Conclusion

COMPUTATIONAL SIGNIFICANCE

Local and global sensitivity analysis on static and time series measurements in Wnt signaling pathway for colorectal cancer is done. Density based Hilbert-Schmidt Information Criterion indices outperformed the variance based Sobol indices. This is attributed to the employment of distance measures & the kernel trick via Reproducing kernel Hilbert space (RKHS) that captures nonlinear relations among various intra/extracellular factors of the pathway in a higher dimensional space. The gained advantage is confirmed on the inferred results obtained via a Bayesian network model based on prior biological knowledge and static gene expression data. In time series data, using these indices it is now possible to observe when and in which period of time and to what degree a factor gets influenced & contributes to the pathway, as changes in concentration of the another factor is made. This facilitates in time based administration of target therapeutic drugs & reveal hidden biological information within colorectal cancer samples.

DEVIATIONS IN FORMULATION OF PSYCHOPHYSICAL LAW

In context of [25]'s work regarding the recent development of observation of Weber's law working downstream of the pathway, it has been found that the law is governed by the ratio of the deviation in the input and the absolute input value. More importantly, it is these deviations in input that are of significance in studing any phemomena. The current manuscript explores the sensitivity of deviation in the fold changes between measurements of fold changes at consecutive time points to explore in what duration of time i.e $< t_i, t_{i+1} >$, a particular factor is affecting the pathway in a major way. This has deeper implications in the fact that one is now able to observe when in time an intervention can be made or a gene be perturbed to study the behaviour of the pathway in tumorous cases. Thus sensitivity analysis of deviations in mathematical formulations of the psychophysical law can lead to insights into the time period based influence of the involved factors

in the pathway. This will also shed light on the duration in which the psychophysical laws might be most prevalent.

SPECIFIC EXAMPLES OF BIOLOGICAL INTERPRETATIONS

GSK3β - It is widely known that *WNT* stimulation leads to inhibition of *GSK3β*. In contrast to this regard *GSK3β* shows a up-regulated levels at t_3, t_{12} and t_{24}. The author is currently unaware of why this contasting behaviour is exhibited. Later upregulation might point to the fact that the effectiveness of Wnt stimulation has decreased and *GSK3β* plays the role of stabilizing and controlling the behaviour of the pathway by working against the Wnt stimulation and preventing further degradation. While work by [27] does not shed light on this aspect, contrasting models of inhibitions for *GSK3* has been recently proposed in [110] which might support this behaviour. Considering analysis of fold changes at different time points, decreasing sensitivity of *GSK3β* was observed for the first two time frames, after which there is an increasing sensitivity for the next three time frames. Comparing this with plots of analysis of deviations in fold changes, it is observed that there is greater significance of deviations in fold changes of *GSK3β* during later stages of $< t_6, t_{12} >$ and $< t_{12}, t_{24} >$. It is in these periods that one might be able to pertube and study significant affects on the pathway.

PORCN - *PORCN* is known to help in the secretion of the Wnt ligands that later on help in the instigation of the signaling activity. Sustanined stimulation by *WNT3A* over a period of time might lead to decrease in the up regulation of *PORCN* which helps in Wnt secretion. Graph for *PORCN* in analysis of fold changes shows increasing significance of the influence of *PORCN* as time passes, even though there is lower regulation of the same at later stages. The highly significant influence of lower regulation at later stages indicates the lessened effectiveness of *PORCN* due to sustained *WNT3A* stimulation that might have suppressed the functionality of secretion carried out via *PORCN*. Contrary to this, the influences of the deviations in the fold changes over time show the reverse behaviour. The maximum influence is during the first two time frames of $< t_1, t_3 >$ and this influence of deviations decreases at later stages. This points to the fact that the deviations in the fold changes at intial stage has greater significance in the pathway than the deviations at later stages. It follows that in initial stages of Wnt stimulation the expression of *PORCN* has significant influence.

Appendix
Choice of sensitivity indices

The **SENSITIVITY PACKAGE** ([122] and [31]) in R langauge provides a range of functions to compute the indices and the following indices will be taken into account for addressing the posed questions in this manuscript.

1. sensiFdiv - conducts a density-based sensitivity analysis where the impact of an input variable is defined in terms of dissimilarity between the original output density function and the output density function when the input variable is fixed. The dissimilarity between density functions is measured with Csiszar f-divergences. Estimation is performed through kernel density estimation and the function kde of the package ks [63] and [66].

2. sensiHSIC - conducts a sensitivity analysis where the impact of an input variable is defined in terms of the distance between the input/output joint probability distribution and the product of their marginals when they are embedded in a Reproducing Kernel Hilbert Space (RKHS). This distance corresponds to HSIC proposed by [68] and serves as a dependence measure between random variables.

3. soboljansen - implements the Monte Carlo estimation of the Sobol indices for both first-order and total indices at the same time (all together 2p indices), at a total cost of (p+2) × n model evaluations. These are called the Jansen estimators [58] and [50].

4. sobol2002 - implements the Monte Carlo estimation of the Sobol indices for both first-order and total indices at the same time (all together 2p indices), at a total cost of (p+2) ×n model evaluations. These are called the Saltelli estimators. This estimator suffers from a conditioning problem when estimating the variances behind the indices computations. This can seriously affect the Sobol indices estimates in case of largely non-centered output. To avoid this effect, you have to center the model output before applying "sobol2002". Functions "soboljansen" and "sobolmartinez" do not suffer from this problem [44].

5. sobol2007 - implements the Monte Carlo estimation of the Sobol indices for both first-order and total indices at the same time (all together 2p indices), at a total cost of (p+2) × n model evaluations. These are called the Mauntz estimators [57].

6. sobolmartinez - implements the Monte Carlo estimation of the Sobol indices for both first-order and total indices using correlation coefficients-based formulas, at a total cost of (p + 2) × n model evaluations. These are called the Martinez estimators.

7. sobol - implements the Monte Carlo estimation of the Sobol sensitivity indices. Allows the estimation of the indices of the variance decomposition up to a given order, at a total cost of (N + 1) × n where N is the number of indices to estimate [30].
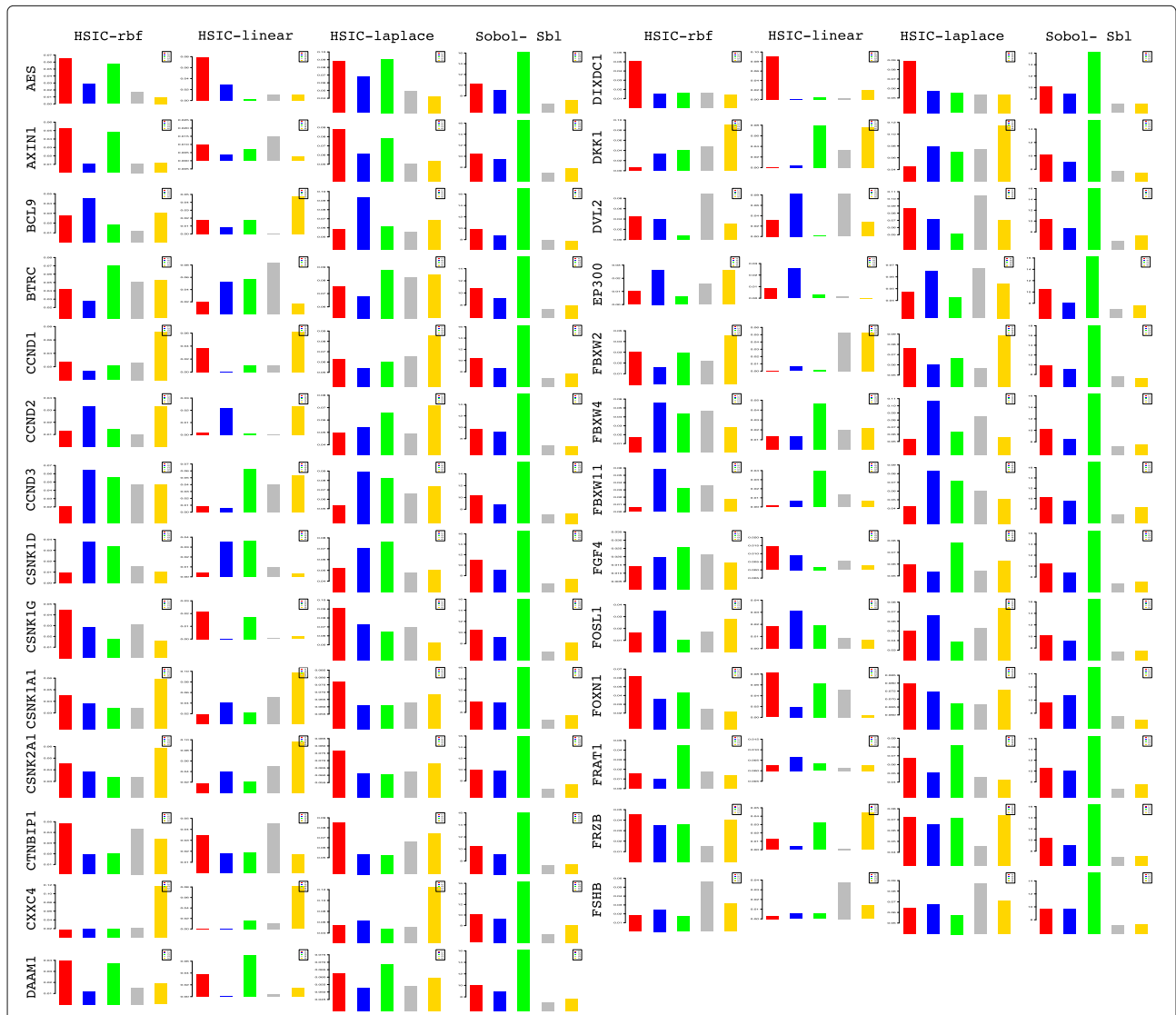
Fig. 27 Column wise - methods to estimate sensitivity indices. Row wise - sensitivity indicies for each gene. For each graph, the bars represent sensitivity indices computed at t1 (red), t2 (blue), t3 (green), t4 (gray) and t5 (yellow). Indices were computed using non scaled time series data. TO - total order; FO - first order; SBL - Sobol
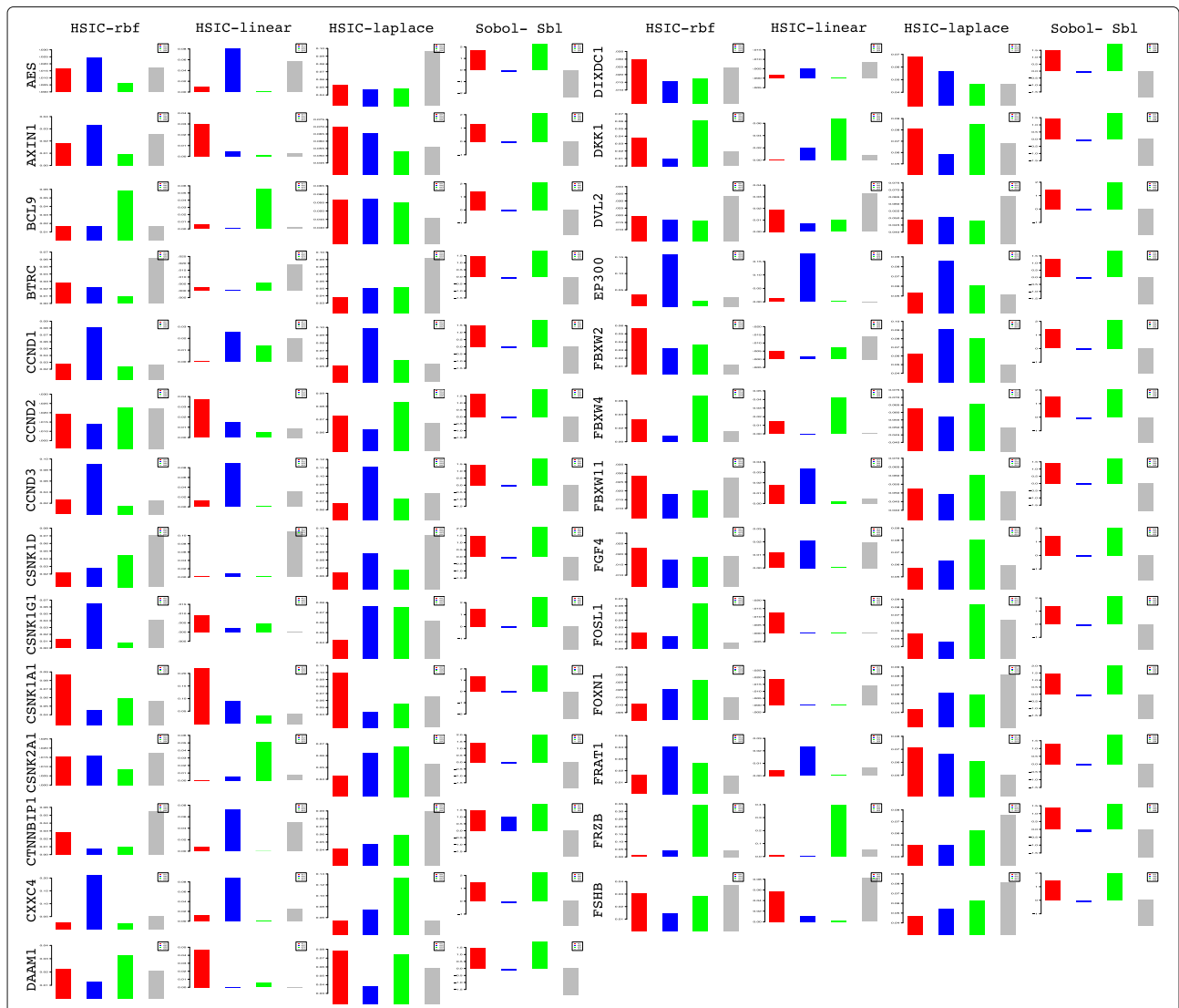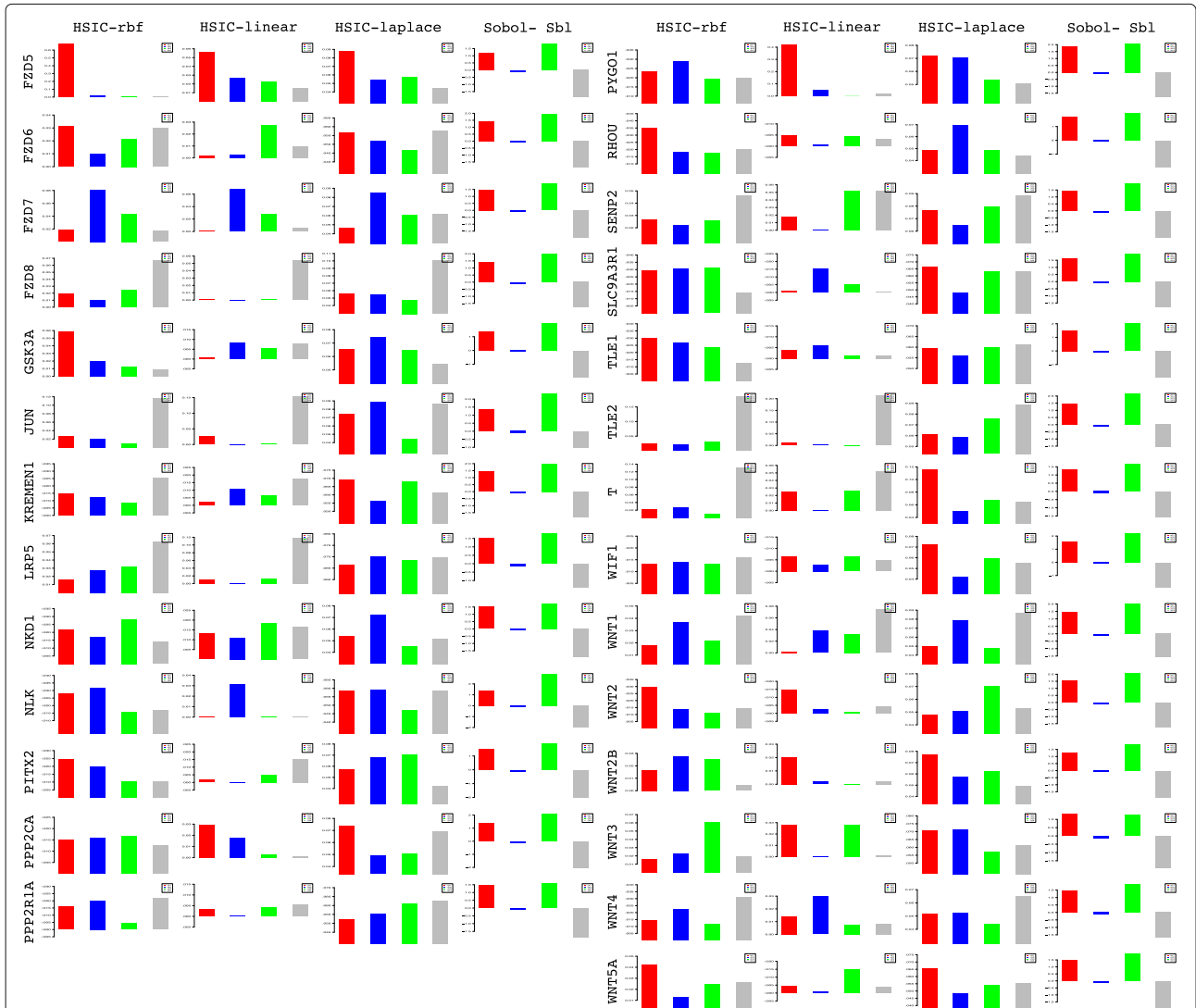
**Fig. 28** Column wise - methods to estimate sensitivity indices. Row wise - sensitivity indicies for each gene. For each graph, the bars represent sensitivity indices computed at t1 (red), t2 (blue), t3 (green), t4 (gray) and t5 (yellow). Indices were computed using non scaled time series data. TO - total order; FO - first order; SBL - Sobol

Fig. 29 Column wise - methods to estimate sensitivity indices. Row wise - sensitivity indicies for each gene on deviations in fold change. For each graph, the bars represent sensitivity indices computed at $< t1, t2 >$ (red), $< t2, t3 >$ (blue), $< t3, t4 >$ (green) and $< t4, t5 >$ (gray). Indices were computed using non scaled time series data. TO - total order; FO - first order; SBL - Sobol

Fig. 30 Column wise - methods to estimate sensitivity indices. Row wise - sensitivity indicies for each gene on deviations in fold change. For each graph, the bars represent sensitivity indices computed at $< t1, t2 >$ (red), $< t2, t3 >$ (blue), $< t3, t4 >$ (green) and $< t4, t5 >$ (gray). Indices were computed using non scaled time series data. TO - total order; FO - first order; SBL - Sobol
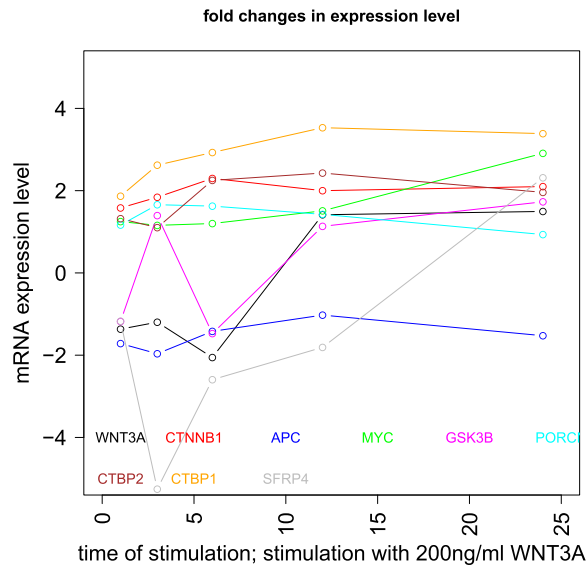
Fig. 31 mRNA expression levels of genes at 1st, 3rd, 6th, 12th and 24th hour from [27]
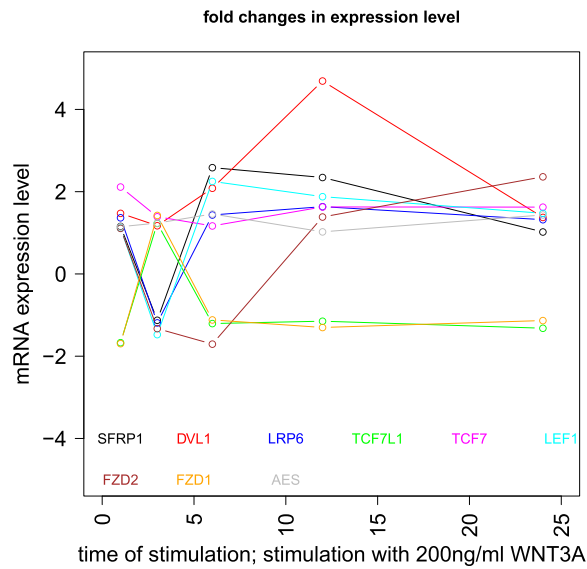
Fig. 32 mRNA expression levels of genes at 1st, 3rd, 6th, 12th and 24th hour from [27]

Abbreviations

AES: amino-terminal enhancer of split; ANOVA: Analysis of variance; APC: Adenomatous polyposis coli or Wnt pathway regulator; AXIN: AXIN or Protein phosphatase 1 regulatory subunits; β-TRCP: β-Transducin repeat containing E3 Ubiquitin protein ligase; BCL9: B-cell CLL/lymphoma 9; BRTC: beta-transducin repeat containing E3 ubiquitin protein ligase; CCND1: Cyclin D1; CCND2/3: Cyclin D2/3; CD44: CD44 antigen (homing function and Indian blood group system); CK1: Casein kinase 1 or serine/threonine-selective enzymes; CSNK1D: casein kinase 1 delta; CSNK1G: casein kinase 1 gamma; CSNK1A1: casein kinase 1 alpha 1; CSNK2A1: casein kinase 2 alpha 1; CTBP: C-terminal binding proteins; CTNNB1: Catenin beta 1; CTNNBIP1: Catenin beta interacting protein 1; CXXC4: CXXC finger protein 4; DAAM1: dishevelled associated activator of morphogenesis 1; DIXDC1: DIX domain containing 1; DVL: Dishevelled segment polarity proteins; DKK: Dickkopf WNT signaling pathway inhibitor; DACT: Dishevelled binding antagonist of beta catenin; EP300: E1A binding protein p300; FBXW: F-box and WD repeat domain containing; FGF: fibroblast growth factor; FOSL1: FOS like 1, AP-1 transcription factor subunit; FOXN1: forkhead box N1; FRAT1: FRAT1, WNT signaling pathway regulator; FRZB: frizzled-related protein; FSHB: follicle stimulating hormone beta subunit; FZD: Frizzled receptors or G protein-coupled receptor proteins; FBOX/WD: F: box and WD repeat domain; GSK3: Glycogen synthase kinase 3 or serine/threonine protein kinase; GROUCHO: Groucho, a transcription-inhibiting factor; HEK 293: Human embryonic kidney cells 293; HSIC: Hilbert-Schmidt Information Criterion; JUN: Jun proto-oncogene, AP-1 transcription factor subunit; KL-divergence: Kullback–Leibler divergence; KREMEN1: kringle containing transmembrane protein 1; LRP: Low density lipoprotein receptor-related proteins; LEF1: Lymphoid enhancer binding factor 1; MYC: MYC proto-oncogene, bHLH transcription factor; NKD1: naked cuticle homolog 1; NLK: nemo like kinase; OAT: One-at-a-time; PITX2: paired like homeodomain 2; PORCN: Porcupine homolog (Drosophila); PPP2CA: protein phosphatase 2 catalytic subunit alpha; PPP2R1A: protein phosphatase 2 scaffold subunit Aalpha; PYGO1: pygopus family PHD finger 1; qPCR: Quantitative polymerase chain reaction; RHOU: ras homolog family member U; RKHS: Reproducing kernel Hilbert space; SENP2: SUMO1/sentrin/SMT3 specific peptidase 2; SLC9A3R1: SLC9A3 regulator 1; SFRP: Secreted frizzled-related protein; SOM: Self organizing maps; T: T brachyury transcription factor; TCF: Transcription factor; TLE: transducin like enhancer of split; TCF7L2: Transcription factor 7-like 2 (T-cell specific, HMG-box); WNT: Wingless-type Mouse Mammary Tumor Virus; WIF1: WNT inhibitory factor 1; WNT-3A: Wnt family member 3A

Availability of data and materials

Code has been made available on Google drive at https://drive.google.com/folderview?id=0B7Kkv8wlhPU-Q2NBZGt1ZERrSVE&usp=sharing Audio file along with the poster presented at SBHD 2016 has been made available on Google drive at https://drive.google.com/drive/folders/0B7Kkv8wlhPU-aVR0eFJqTkNUOFE The datasets generated and/or analysed during the current study are available in the GEO public database (accession number GSE10972) repository, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE10972 for [28]. The datasets generated and/or analysed during the current study are available in Table S3, http://dx.doi.org/10.1371/journal.pone.0010024 for [27].

Authors' information

shriprakash sinha holds a MSc from Utrecht University (The Netherlands) in Applied Computing Science and MS in Computer Science from Oregon State University (USA). He has worked as an intern in Siemens Information Systems Limited, Computer Aided Diagnosis Research Center (Bangalore, India), as a project associate in the prestigious Indian Institute of Science (IISc, Bangalore, India) and as a scientific researcher/fMRI data analyst at the Neuroimaging Center (NiC) in UMCG hospital (Groningen, The Netherlands). He also had the opportunity to work at the TuDelft and Philips (The Netherlands) before returning temporarily to India and working as a researcher independently. As an independent researcher he has been interested in working on systems biology of the Wnt pathway using sensitivity analysis and machine learning methods as well as development of search engine to prioritize extra/intracellular nth order interactions that affect the pathway. It is hoped that these prioritization in the form of rankings will help reduce wet lab experiments to test crucial interactions in the pathway and thus save significant costs in wet lab experiments. As a further development, observing the changing rankings in time will lead to time based intervention in the pathway and development of target based drugs for cancer. He currently works as a Senior Lecturer in the faculty of Maths and IT at Royal Thimphu College Bhutan and teaches data mining and problem solving skills. Apart from scientific pursuits, sinha is also engaged in conducting 15 days intensive meditations programs based on the Heart Sutra.

Authors' contributions

SS conceived and designed the experiments; performed the experiments; analyzed the data; wrote the paper.

Ethics approval and consent to participate

The datasets from the articles [27] and [28] were used in the computational experiments. [27] states the following -

```
N/A. An ethics statement is not required for this
work. An informed consent from participants
involved is also not applicable for this work.
```

This has been made available on the PLOS journal at http://dx.doi.org/10.1371/journal.pone.0010024. [28] states the following -

```
Human tissue samples were obtained from Singapore
Tissue Network using protocols approved by
institutional Review Board of National University
of Singapore; informed consent was obtained from
each individual who provided the tissues. The
colorectal cance cell lines and non-transformed
cell lines used in this study were purchased from
the American Type Culture Collection (Manassas,
VA).
```

Consent for publication

[28] states the following -

```
Human tissue samples were obtained from Singapore
Tissue Network using protocols approved by
institutional Review Board of National University
of Singapore; informed consent was obtained from
each individual who provided the tissues. The
colorectal cance cell lines and non-transformed
cell lines used in this study were purchased from
the American Type Culture Collection (Manassas,
VA).
```

Under the License number 4085451080321, the author of this article is using Table S1 of [28] containing static expression profiles. For [27] the following statement holds -

```
Copyright: ©2010 Gujral, MacBeath. This is an
open-access article distributed under the terms of
the Creative Commons Attribution License, which
permits unrestricted use, distribution, and
reproduction in any medium, provided the original
author and source are credited.
```

The dynamic data from [27] that is used in this manuscript is available and can be downloaded from Table S3 of PLOS journal website.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Sharma R. Wingless a new mutant in drosophila melanogaster. Drosophila Inf Serv. 1973;50:134–4.
2. Thorstensen L, Lind GE, Løvig T, Diep CB, Meling GI, Rognum TO, Lothe RA. Genetic and epigenetic changes of components affecting the wnt pathway in colorectal carcinomas stratified by microsatellite instability. Neoplasia. 2005;7(2):99–108.
3. Baron R, Kneissel M. Wnt signaling in bone homeostasis and disease: from human mutations to treatments. Nat Med. 2013;19(2):179–92.
4. Clevers H. Wnt/[β]-catenin signaling in development and disease. Cell. 2006;127(3):469–80.
5. Sokol S. Wnt Signaling in Embryonic Development, vol 17: Elsevier; 2011.
6. Pinto D, Gregorieff A, Begthel H, Clevers H. Canonical wnt signals are essential for homeostasis of the intestinal epithelium. Gene Dev. 2003;17(14):1709–13.
7. Zhong Z, Ethen NJ, Williams BO. Wnt signaling in bone development and homeostasis. Wiley Interdiscip Rev Dev Biol. 2014;3(6):489–500.
8. Pećina-Šlaus N. Wnt signal transduction pathway and apoptosis: a review. Cancer Cell Int. 2010;10(1):1–5.
9. Kahn M. Can we safely target the wnt pathway? Nat Rev Drug Discov. 2014;13(7):513–32.
10. Garber K. Drugging the wnt pathway: problems and progress. J Natl Cancer Inst. 2009;101(8):548–50.
11. Voronkov A, Krauss S. Wnt/beta-catenin signaling and small molecule inhibitors. Curr Pharm Des. 2012;19(4):634.
12. Blagodatski A, Poteryaev D, Katanaev V. Targeting the wnt pathways for therapies. Mol Cell Ther. 2014;2:28.
13. Curtin JC, Lorenzi MV. Drug discovery approaches to target wnt signaling in cancer stem cells. Oncotarget. 2010;1(7):552.
14. Rao TP, Kühl M. An updated overview on wnt signaling pathways a prelude for more. Circ Res. 2010;106(12):1798–1806.
15. Yu J, Virshup DM. Updating the wnt pathways. Biosci Rep. 2014;34(5):593–607.
16. Antebi YE, Nandagopal N, Elowitz MB. An operational view of intercellular signaling pathways. Curr Opin Syst Biol. 2017;1:16–24.
17. Goentoro L. Cross-hierarchy systems principles. Curr Opin Syst Biol. 2016;1:80–83.
18. Lee E, Salic A, Kruger R, Heinrich R, Kirschner MW. The roles of apc and axin derived from experimental and theoretical analysis of the wnt pathway. PLoS Biol. 2004;2(3):405–6.
19. Kogan Y, Halevi-Tobias KE, Hochman G, Baczmanska AK, Leyns L, Agur Z. A new validated mathematical model of the wnt signalling pathway predicts effective combinational therapy by sfrp and dkk. Biochem J. 2012;444(1):115–25.
20. Lee M, Chen GT, Puttock E, Wang K, Edwards RA, Waterman ML, Lowengrub J. Mathematical modeling links wnt signaling to emergent patterns of metabolism in colon cancer. Mol Syst Biol. 2017;13(2):912.
21. MacLean AL, Rosen Z, Byrne HM, Harrington HA. Parameter-free methods distinguish wnt pathway models and guide design of experiments. Proc Natl Acad Sci. 2015;112(9):2652–7.
22. Koutroumpas K, Ballarini P, Votsi I, Cournède PH. Bayesian parameter estimation for the wnt pathway: an infinite mixture models approach. Bioinformatics. 2016;32(17):781–9.
23. Sinha S. Integration of prior biological knowledge and epigenetic information enhances the prediction accuracy of the bayesian wnt pathway. Integr Biol. 2014;6:1034–48. doi:10.1039/c4ib00124a.
24. Sinha S. A pedagogical walkthrough of computational modeling and simulation of wnt signaling pathway using static causal models in matlab. EURASIP J Bioinforma Syst Biol. 2016;2017(1):1.
25. Goentoro L, Kirschner MW. Evidence that fold-change, and not absolute level, of β-catenin dictates wnt signaling. Mol Cell. 2009;36:872–84.
26. Azam M, Bhatti A, Arshad A, Babar M. Sensitivity analysis of wnt signaling pathway. In: Applied Sciences and Technology (IBCAST), 2013 10th International Bhurban Conference On. IEEE. 2013. p. 122–7.
27. Gujral TS, MacBeath G. A system-wide investigation of the dynamics of wnt signaling reveals novel phases of transcriptional regulation. PloS ONE. 2010;5(4):10024.
28. Jiang X, Tan J, Li J, Kivimäe S, Yang X, Zhuang L, Lee PL, Chan MT, Stanton LW, Liu ET, et al. Dact3 is an epigenetic regulator of wnt/β-catenin signaling in colorectal cancer and is a therapeutic target of histone modifications. Cancer Cell. 2008;13(6):529–41.

29. Gregorieff A, Clevers H. Wnt signaling in the intestinal epithelium: from endoderm to cancer. Gene Dev. 2005;19(8):877–90.
30. Sobol' IM. On sensitivity estimation for nonlinear mathematical models. Matematicheskoe Modelirovanie. 1990;2(1):112–8.
31. Iooss B, Lemaître P. A review on global sensitivity analysis methods. 2014. arXiv preprint arXiv:1404.2405.
32. Morris MD. Factorial sampling plans for preliminary computational experiments. Technometrics. 1991;33(2):161–74.
33. Moon H, Dean AM, Santner TJ. Two-stage sensitivity-based group screening in computer experiments. Technometrics. 2012;54(4):376–87.
34. Dean A, Lewis S. Screening: Methods for Experimentation in Industry, Drug Discovery, and Genetics: Springer; 2006.
35. Andres TH, Hajas WC. Using iterated fractional factorial design to screen parameters in sensitivity analysis of a probabilistic risk assessment model. 1993.
36. Bettonvil B, Kleijnen JP. Searching for important factors in simulation models with many factors: Sequential bifurcation. Eur J Oper Res. 1997;96(1):180–94.
37. Cotter SC. A screening design for factorial experiments with interactions. Biometrika. 1979;66(2):317–20.
38. Christensen R. Linear Models for Multivariate, Time Series, and Spatial Data: Springer; 1991.
39. Saltelli A, Chan K, Scott E. Sensitivity analysis wiley series in probability and statistics. 2000.
40. Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab Eng Syst Saf. 2003;81(1):23–69.
41. McKay MD, Beckman RJ, Conover WJ. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics. 1979;21(2):239–45.
42. Homma T, Saltelli A. Importance measures in global sensitivity analysis of nonlinear models. Reliab Eng Syst Saf. 1996;52(1):1–17.
43. Sobol IM. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. Math Comput Simul. 2001;55(1):271–80.
44. Saltelli A. Making best use of model evaluations to compute sensitivity indices. Comput Phys Commun. 2002;145(2):280–97.
45. Saltelli A, Ratto M, Tarantola S, Campolongo F. Sensitivity analysis for chemical models. Chem Rev. 2005;105(7):2811–28.
46. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S. Global Sensitivity Analysis: the Primer: Wiley; 2008.
47. Cukier R, Fortuin C, Shuler KE, Petschek A, Schaibly J. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. i theory. J Chem Phys. 1973;59(8):3873–8.
48. Saltelli A, Tarantola S, Chan KS. A quantitative model-independent method for global sensitivity analysis of model output. Technometrics. 1999;41(1):39–56.
49. Tarantola S, Gatelli D, Mara TA. Random balance designs for the estimation of first order global sensitivity indices. Reliab Eng Syst Saf. 2006;91(6):717–27.
50. Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S. Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. Comput Phys Commun. 2010;181(2):259–70.
51. Janon A, Klein T, Lagnoux A, Nodet M, Prieur C. Asymptotic normality and efficiency of two sobol index estimators. ESAIM Probab Stat. 2014;18:342–64.
52. Owen AB. Better estimation of small sobol'sensitivity indices. ACM Trans Model Comput Simul (TOMACS). 2013;23(2):11.
53. Tissot JY, Prieur C. Bias correction for the estimation of sensitivity indices based on random balance designs. Reliab Eng Syst Saf. 2012;107:205–13.
54. Da Veiga S, Gamboa F. Efficient estimation of sensitivity indices. J Nonparametric Stat. 2013;25(3):573–95.
55. Archer G, Saltelli A, Sobol I. Sensitivity measures, anova-like techniques and the use of bootstrap. J Stat Comput Simul. 1997;58(2):99–120.
56. Tarantola S, Gatelli D, Kucherenko S, Mauntz W, et al. Estimating the approximation error when fixing unessential factors in global sensitivity analysis. Reliab Eng Syst Saf. 2007;92(7):957–60.
57. Saltelli A, Annoni P. How to avoid a perfunctory sensitivity analysis. Environ Model Softw. 2010;25(12):1508–17.
58. Jansen MJ. Analysis of variance designs for model output. Comput Phys Commun. 1999;117(1):35–43.

59. Storlie CB, Helton JC. Multiple predictor smoothing methods for sensitivity analysis: Description of techniques. Reliab Eng Syst Saf. 2008;93(1):28–54.

60. Da Veiga S, Wahl F, Gamboa F. Local polynomial estimation for sensitivity analysis on models with correlated inputs. Technometrics. 2009;51(4):452–63.

61. Li G, Rosenthal C, Rabitz H. High dimensional model representations. J Phys Chem A. 2001;105(33):7765–77.

62. Hajikolaei KH, Wang GG. High dimensional model representation with principal component analysis. J Mech Des. 2014;136(1):011003.

63. Borgonovo E. A new uncertainty importance measure. Reliab Eng Syst Saf. 2007;92(6):771–84.

64. Sobol IM, Kucherenko S. Derivative based global sensitivity measures and their link with global sensitivity indices. Math Comput Simul. 2009;79(10):3009–17.

65. Fort JC, Klein T, Rachdi N. New sensitivity analysis subordinated to a contrast. 2013. arXiv preprint arXiv:1305.2329.

66. Da Veiga S. Global sensitivity analysis with dependence measures. J Stat Comput Simul. 2015;85(7):1283–305.

67. Székely GJ, Rizzo ML, Bakirov NK, et al. Measuring and testing dependence by correlation of distances. Ann Stat. 2007;35(6):2769–794.

68. Gretton A, Bousquet O, Smola A, Schölkopf B. Measuring statistical dependence with hilbert-schmidt norms. In: Algorithmic Learning Theory. Springer. 2005. p. 63–77.

69. Csiszar I, et al. Information-type measures of difference of probability distributions and indirect observations. Studia Sci Math Hungar. 1967;2: 299–318.

70. Aizerman M, Braverman E, Rozonoer L. Theoretical foundations of the potential function method in pattern recognition learning. Autom Remote Control. 1964;25:821–37.

71. Sumner T, Shephard E, Bogle I. A methodology for global-sensitivity analysis of time-dependent outputs in systems biology modelling. J R Soc Interface. 2012;9(74):2156–66.

72. Zheng Y, Rundell A. Comparative study of parameter sensitivity analyses of the tcr-activated erk-mapk signalling pathway. IEE Proc-Syst Biol. 2006;153(4):201–11.

73. Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J Theor Biol. 2008;254(1):178–96.

74. Sinha S. Sensitivity analysis of wnt β-catenin based transcription complex might bolster power-logarithmic psychophysical law and reveal preserved gene gene interactions. 2015. bioRxiv, 015834. doi:10.1101/015834.

75. Adler M, Mayo A, Alon U. Logarithmic and power law input-output relations in sensory systems with fold-change detection. PLoS Comput Biol. 2014;10(8):1003781.

76. Masin SC, Zudini V, Antonelli M. Early alternative derivations of fechner's law. J Hist Behav Sci. 2009;45:56–65. doi:10.1002/jhbs.20349.

77. Fechner GT. Elemente der Psychophysik (2 Vols): Breitkopf and Hartel; 1860.

78. Weber EH. De Pulsu Resorptione, Auditu et Tactu: Annotationes anatomicae et physiologicae; 1834.

79. Bernoulli D. Specimen theoriae novae de mensura sortis. Commentarii Acad Sci Imperialis Petropolitanae. 1738;5:175–92.

80. Sobol S, andKucherenko IM. Global sensitivity indices for nonlinear mathematical models. review. Wilmott Magazine, 2–7.

81. Baucells M, Borgonovo E. Invariant probabilistic sensitivity analysis. Manag Sci. 2013;59(11):2536–49.

82. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. Phys Rev E. 2004;69(6):066138.

83. Sejdinovic D, Sriperumbudur B, Gretton A, Fukumizu K, et al. Equivalence of distance-based and rkhs-based statistics in hypothesis testing. Ann Stat. 2013;41(5):2263–91.

84. Daumé III H. From zero to reproducing kernel hilbert spaces in twelve pages or less. 2004.

85. Riesz F. Sur une espèce de géométrie analytique des systèmes de fonctions sommables. CR Acad Sci Paris. 1907;144:1409–11.

86. Taylor JS, Cristianini N. Properties of Kernels: Cambridge University Press; 2004. Chap. 3.

87. De Lozzo M, Marrel A. New improvements in the use of dependence measures for sensitivity analysis and screening. 2014. arXiv preprint arXiv:1412.1414.

88. Ueno K, Hirata H, Hinoda Y, Dahiya R. Frizzled homolog proteins, micrornas and wnt signaling in cancer. Int J Cancer. 2013;132(8):1731–40.

89. Holcombe R, Marsh J, Waterman M, Lin F, Milovanovic T, Truong T. Expression of wnt ligands and frizzled receptors in colonic mucosa and in colon carcinoma. Mol Pathol. 2002;55(4):220.

90. Planutis K, Planutiene M, Nguyen AV, Moyer MP, Holcombe RF. Invasive colon cancer, but not non-invasive adenomas induce a gradient effect of wnt pathway receptor frizzled 1 (fz1) expression in the tumor microenvironment. J Transl Med. 2013;11(50):10–1186.

91. Sato A, Yamamoto H, Sakane H, Koyama H, Kikuchi A. Wnt5a regulates distinct signalling pathways by binding to frizzled2. EMBO J. 2010;29(1): 41–54.

92. Klapholz-Brown Z, Walmsley GG, Nusse YM, Nusse R, Brown PO. Transcriptional program induced by wnt protein in human fibroblasts suggests mechanisms for cell cooperativity in defining tissue microenvironments. PLoS ONE. 2007;2(9):945.

93. Yokoyama N, Yin D, Malbon CC. Abundance, complexation, and trafficking of wnt/β-catenin signaling elements in response to wnt3a. J Mol Signal. 2007;2(1):11.

94. He TC, Sparks AB, Rago C, Hermeking H, Zawel L, da Costa LT, Morin PJ, Vogelstein B, Kinzler KW. Identification of c-myc as a target of the apc pathway. Science. 1998;281(5382):1509–12.

95. Korinek V, Barker N, Morin PJ, van Wichen D, de Weger R, Kinzler KW, Vogelstein B, Clevers H. Constitutive transcriptional activation by a β-catenin-tcf complex in apc-/- colon carcinoma. Science. 1997;275(5307):1784–7.

96. Morin PJ, Sparks AB, Korinek V, Barker N, Clevers H, Vogelstein B, Kinzler KW. Activation of β-catenin-tcf signaling in colon cancer by mutations in β-catenin or apc. Science. 1997;275(5307):1787–90.

97. Hino SI, Michiue T, Asashima M, Kikuchi A. Casein kinase iε enhances the binding of dvl-1 to frat-1 and is essential for wnt-3a-induced accumulation of β-catenin. J Biol Macromol. 2003;278(16):14066–73.

98. You XJ, Bryant PJ, Jurnak F, Holcombe RF. Expression of wnt pathway components frizzled and disheveled in colon cancer arising in patients with inflammatory bowel disease. Oncol Rep. 2007;18(3):691–4.

99. González-Sancho JM, Brennan KR, Castelo-Soccio LA, Brown AM. Wnt proteins induce dishevelled phosphorylation via an lrp5/6-independent mechanism, irrespective of their ability to stabilize β-catenin. Mol Cell Biol. 2004;24(11):4757–68.

100. Gao C, Chen YG. Dishevelled: The hub of wnt signaling. Cell Signal. 2010;22(5):717–27.

101. Huang MY, Yen LC, Liu HC, Liu PP, Chung FY, Wang TN, Wang JY, Lin SR. Significant overexpression of dvl1 in taiwanese colorectal cancer patients with liver metastasis. Int J Mol Sci. 2013;14(10):20492–507.

102. Galli LM, Barnes T, Cheng T, Acosta L, Anglade A, Willert K, Nusse R, Burrus LW. Differential inhibition of wnt-3a by sfrp-1, sfrp-2, and sfrp-3. Dev Dyn. 2006;235(3):681–90.

103. Suzuki H, Watkins DN, Jair KW, Schuebel KE, Markowitz SD, Chen WD, Pretlow TP, Yang B, Akiyama Y, van Engeland M, et al. Epigenetic inactivation of sfrp genes allows constitutive wnt signaling in colorectal cancer. Nat Genet. 2004;36(4):417–22.

104. Caldwell GM, Jones C, Gensberg K, Jan S, Hardy RG, Byrd P, Chughtai S, Wallis Y, Matthews GM, Morton DG. The wnt antagonist sfrp1 in colorectal tumorigenesis. Cancer Res. 2004;64(3):883–8.

105. Chinnadurai G. Ctbp, an unconventional transcriptional corepressor in development and oncogenesis. Mol Cell. 2002;9(2):213–24.

106. Hamada F, Bienz M. The apc tumor suppressor binds to c-terminal binding protein to divert nuclear β-catenin from tcf. Dev Cell. 2004;7(5): 677–85.

107. Schneikert J, Brauburger K, Behrens J. Apc mutations in colorectal tumours from fap patients are selected for ctbp-mediated oligomerization of truncated apc. Hum Mol Genet. 2011;20(18):3554–64.

108. Patel J, Baranwal S, Love IM, Patel NJ, Grossman SR, Patel BB. Inhibition of c-terminal binding protein attenuates transcription factor 4 signaling to selectively target colon cancer stem cells. Cell Cycle. 2014;13(22): 3506–18.

109. Willert K, Nusse R. Wnt proteins. Cold Spring Harb Perspect Biol. 2012;4(9):007864.

110. Metcalfe C, Bienz M. Inhibition of gsk3 by wnt signalling–two contrasting models. J Cell Sci. 2011;124(21):3537–44.

111. Lum L, Clevers H. The unusual case of porcupine. Science. 2012;337(6097):922–3.

112. Chinnadurai G. Ctbp family proteins: more than transcriptional corepressors. Bioessays. 2003;25(1):9–12.

113. Kim J, Kim S. In silico identification of sfrp1 as a hypermethylated gene in colorectal cancers. Genomics Inf. 2014;12(4):171–80.

114. Wu CH, Chung FY, Chang JY, Wang JY. Rapid detection of gene expression by a colorectal cancer enzymatic gene chip detection kit. Biomark Genomic Med. 2013;5(3):87–91.

115. MacDonald BT, Semenov MV, Huang H, He X. Dissecting molecular differences between wnt coreceptors lrp5 and lrp6. PLoS ONE. 2011;6(8): 23537.

116. Liu G, Bafico A, Harris VK, Aaronson SA. A novel mechanism for wnt activation of canonical signaling through the lrp6 receptor. Mol Cell Biol. 2003;23(16):5825–35.

117. Watanabe T, Kobunai T, Toda E, Kanazawa T, Kazama Y, Tanaka J, Tanaka T, Yamamoto Y, Hata K, Kojima T, et al. Gene expression signature and the prediction of ulcerative colitis–associated colorectal cancer by dna microarray. Clin Cancer Res. 2007;13(2):415–20.

118. Lemieux E, Cagnol S, Beaudry K, Carrier J, Rivard N. Oncogenic kras signalling promotes the wnt/β-catenin pathway through lrp6 in colorectal cancer. Oncogene. 2014;34:4914–27.

119. Hikasa H, Sokol SY. Phosphorylation of tcf proteins by homeodomain-interacting protein kinase 2. J Biol Chem. 2011;286(14): 12093–100.

120. Leushacke M, Spörle R, Bernemann C, Brouwer-Lehmitz A, Fritzmann J, Theis M, Buchholz F, Herrmann BG, Morkel M. An rna interference phenotypic screen identifies a role for fgf signals in colon camangancer progression. PLoS ONE. 2011;6(8):23381.

121. Cadigan KM, Waterman ML. Tcf/lefs and wnt signaling in the nucleus. Cold Spring Harb Perspect Biol. 2012;4(11):007906.

122. Faivre R, Iooss B, Mahévas S, Makowski D, Monod H. Analyse de Sensibilité et Exploration de Modèles: Application aux Sciences de la Nature et de L'environnement: Editions Quae; 2013.