Coupling kinetic models and advection–diffusion equations. 2. Sensitivity analysis of an advection–diffusion–reaction model

Lafras Uys, Jan-Hendrik S. Hofmeyr, and Johann M. Rohwer*

Laboratory for Molecular Systems Biology, Department of Biochemistry, Stellenbosch University, Private Bag X1, Matieland, Stellenbosch 7602, Western Cape, South Africa

*Corresponding author's e-mail address: jr@sun.ac.za

Citation: Uys L, Hofmeyr J-HS, Rohwer JM. 2021. Coupling kinetic models and advection–diffusion equations. 2. Sensitivity analysis of an advection–diffusion–reaction model. In Silico Plants 2021: diab014; doi: 10.1093/insilicoplants/diab014

ABSTRACT

The accompanying paper (Uys et al., in silico Plants, 2021: diab013) presented a core model of sucrose accumulation within the advection–diffusion–reaction framework, which is able to capture the spatio-temporal evolution of the system from a set of initial conditions. This paper presents a sensitivity analysis of this model. Because this is a non-steady-state model based on partial differential equations, we performed the sensitivity analysis using two approaches from engineering. The Morris method is based on a one-at-a-time design, perturbing parameters individually and calculating the influence on model output in terms of elementary effects. Fourier amplitude sensitivity test (FAST) is a global sensitivity analysis method, where all parameters are perturbed simultaneously, oscillating at different frequencies, enabling the calculation of the contribution of each parameter through Fourier analysis. Overall, both methods gave similar results. Perturbations in reactions tended to have a large influence on their own rate, as well as on directly connected metabolites. Sensitivities varied both with the time of the simulation and the position along the sugarcane stalk. Our results suggest that vacuolar sucrose concentrations are most sensitive to vacuolar invertase in the centre of the stalk, but that phloem unloading and vacuolar sucrose uptake also contribute, especially towards the stalk edges. Sucrose in the phloem was most sensitive to phloem loading at the nodes, but most sensitive to phloem unloading in the middle of the internodes. Sink concentrations of sucrose in the symplast were most sensitive to phloem unloading in the middle of the internodes, but at the nodes cytosolic invertase had the greatest effect.

KEYWORDS: Elementary effect; Fourier amplitude; global sensitivity analysis; one-at-a-time; sugarcane.

1. INTRODUCTION

The flow of mass between the source and sink tissues of plants is governed by a complex relationship that exists between transport and reaction steps. As a consequence, mathematical models of these processes are often complex. For example, systems of strongly coupled, non-linear differential equations, which simultaneously require a large number of parameters, are often employed to model such systems. In these cases, an algorithmic method is required to decide which parameters exert the greatest influence on the model output or model predictions.

The phloem and xylem are the main conduits through which mass flow is transacted. In the particular case of sugarcane, sucrose is synthesized in the leaves, translocated to the storage tissue via the phloem and stored in the culm (van Bel 1993; Komor 2000; Walsh et al. 2005). In the accompanying paper (Uys et al. 2021) we developed a framework to couple metabolic pathways to phloem translocation, and more specifically, to couple reaction kinetic models to advection–diffusion models. In this paper two methods of sensitivity analysis are introduced to quantify the effect of various parameters on output variables from the model when investigating its transient behaviour.

Some general background to sensitivity analysis is introduced first. Following this, an explanation of the two methods of sensitivity analysis used in this paper, i.e. the Morris method and Fourier amplitude sensitivity test (FAST), will be given. The advection–diffusion–reaction (ADR) model is then subjected to sensitivity analysis and the results are discussed.
2. METHODS

2.1 Basic principles of sensitivity analysis

One of the main reasons for subjecting a model to sensitivity analysis is that it is sometimes difficult to infer meaningful conclusions about the model’s behaviour from a functional analysis point of view. Sensitivity analysis can help in this case, because it allows one to interrogate such a model to find out which parameters have the greatest influence on determining a particular set of model outcomes. Importantly, sensitivity analysis in the sense employed here does not say anything about the existence of stable states, the convergence of the model to a stable state or the behaviour around a stable state. In our case we only require our system of equations to be numerically integrable.

A dynamic model requires two sets of input data, namely a parameter set and a set of initial conditions. One can gauge the sensitivity of a model’s outputs to either of these sets. Here we only consider the set of parameters. In the case where the parameters are continuously defined on R (or some discrete, possibly finite, yet large set) it becomes necessary to sample points from R^k (or the k-dimensional lattice generated by a discrete set), where k is the number of parameters of interest. More than one point is usually picked and the set of these points is called a design. The different processes by which successive points are chosen are collectively known as experimental design. The objective of picking a subset from R^k as a design is to find a way to choose the fewest number of points that maximizes the amount of information obtained from the model output. The motivation for choosing a small number of points is to minimize the number of model evaluations, since repeated model evaluations can quickly become expensive in terms of computational time and/or resources. Given a particular model, this is not always a well-defined problem to solve.

Typically, points are sampled from the unit cube and converted to appropriate parameter values for the model. This keeps the method general and independent of the actual model that is to be analysed. The set of sampled points is typically given as a design matrix. In this paper we will denote this matrix with X, where the columns correspond to a particular parameter and the rows to the sampled points.

A so-called transfer function converts the normalized points to parameter values. The functions we use are from the family of inverse cumulative distribution functions (abbreviated as a cdf −1). A cdf −1 is defined on the interval [0, 1], i.e. exactly the points provided by the design.

Let f(p) denote the model output and p be a parameter vector distributed around some expected value, E(p), with variance, Var(p). The following three points need to be noted. First, the algorithm does not demand a particular distribution, any distribution will do. The variance is chosen to reflect the uncertainty in a parameter—typically the standard error in an experimental measurement. Second, Var(p) is the diagonal of the covariance matrix and we treat all parameters as independent of each other (off-diagonal entries are zero). In other words, we assume that there are no confounding influences amongst parameters and that changing one parameter does not influence another parameter. Neither of the two sensitivity analysis methods, as implemented here, can deal with parameters that interact with each other.

Third, in the analysis of the ADR model, the output of interest will be the concentrations of the respective metabolites and the rates of the biochemical and transport reactions. Nothing prevents the output from being something else, for example the output could be the ratio of two concentrations, the sum of a subset of the fluxes or the concentrations and rates when the model is at steady state.

In the model introduced in Uys et al. (2021), the output is actually a function of the parameter vector p, position z along the stalk, and time t, i.e. f(p) = f(p, z, t). Notationally, we write f(p) = f(p)|t,z to refer to the model evaluated at a specific point, z, and time, t.

The direct output from evaluating the model with the design points is usually irrelevant. However, sensitivity analysis methods generally specify a means of analysing the output that is generated by evaluating the model for each point in the design. The result of this allows one to rank the relative importance of the parameters in determining a particular model output.

2.2 Global versus local sensitivity analysis

A sensitivity analysis uses the perturbation of one or more parameters to find a quantifiable measure of the subsequent change in model output. The response of a variable to perturbations in a set of parameters can be compared to find the parameter that has the greatest influence on its value. Qian and Mahdi (2020) provide an overview of the most important sensitivity analysis methods used in the biomedical sciences. They distinguish between global and local sensitivity analysis methods and their brief introductions explaining each of the methods (including Morris and FAST) provide excellent starting points for delving deeper into a particular analysis method. Various global sensitivity analysis methods have been reviewed in the specific context of systems biology (Marino et al. 2008).

A global sensitivity analysis aims to explore the whole parameter space (i.e. the space of all inputs), as opposed to a local analysis, which only explores a small neighbourhood around a reference point in the parameter space. Methods such as metabolic control analysis (MCA, Kacer and Burns 1973; Heinrich and Rapoport 1974) exist for performing a local sensitivity analysis around a steady state for ordinary differential equation (ODE)-based reaction kinetic models. The model described in this and the accompanying paper is neither at steady state nor ODE-based, and therefore a variational sensitivity analysis method is needed.

The simplest perturbation is simply to change one parameter at a time, i.e. a so-called one-at-a-time (OAT) design. This approach neglects the possibility that some effects are only noticeable when two or more parameters are perturbed at the same time. In order to do a global sensitivity analysis it is therefore important to choose points so that perturbations occur simultaneously.

There is usually a trade-off between the computational cost and accuracy of running a sensitivity analysis. For this reason it is often advisable to run a so-called screening test for models with a large number of parameters. This establishes the parameters with the largest elementary effects and allows one to focus on these in further tests.

One such screening method has been proposed by Morris, based on an OAT design. The number of times a model has to be evaluated is linear in the number of parameters. A global sensitivity analysis method such as FAST (which we consider later), on the other hand, requires model evaluations exponential in parameter number.
We here propose the use of Morris, followed by FAST, to analyse the sugarcane ADR model. Because the model is of a size that both methods are computationally feasible, they are both applied to the complete model, which allows for a detailed comparison. The sensitivity of the model output to its parameters is calculated, with specific focus on how sensitivities change along the length of the domain and with time. The relative computational requirements of both methods, as well as the potential of Morris as an OAT screening method for FAST analysis of more extensive models, are addressed in the Discussion. Below we first provide an outline of both Morris and FAST.

2.3 Method of Morris
Morris (1991) proposed a method of sensitivity analysis to extract elementary effects of parameters on model outputs. Given a model with \( k \) parameters of interest, an elementary effect is defined as

\[
d_{i}(p) = (f(p_{i+1}) - f(p_{i})) \cdot \text{diag}(\Delta)^{-1}, \quad \text{with } i = 1 \ldots k,
\]

where \( \Delta \) is a vector of perturbations and \( p \) is a \( k \)-dimensional parameter vector, \( f(\cdot) \) is a vector of model outputs and \( d_{i}(p) \) is the vector of elementary effects of the \( i \)-th parameter on the model output. The subscript \( i \) refers to a perturbation in the \( i \)-th parameter of the parameter vector, i.e. parameter vectors \( p_{i} \) and \( p \) differ only in position \( i \). The matrix \( \text{diag}(\Delta) \) is a diagonal matrix constructed from \( \Delta \); the inverse of \( \Delta \) matrix \( \text{diag}(\Delta) \) is the number of levels \( l \) in the design (an even number) and \( \Delta \) is the perturbation size, which is usually chosen as \( l/(2(l-1)) \) (Morris 1991).

The algorithm for generating \( p \) and \( \Delta \) is described in Morris (1991). Briefly, start with a \((k+1) \times k\) sampling matrix,

\[
B = \begin{pmatrix}
0 & 0 & \ldots & 0 \\
1 & 0 & \ldots & 0 \\
1 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
1 & 1 & \ldots & 1
\end{pmatrix},
\]

the successive rows of which can be seen to differ only in the \( i \)-th position. Generate a new matrix,

\[
B^{\prime} = \{1x^{*} + \Delta/2((2B - 1)D^{*} + 1)\}P^{*},
\]

where \( 1 \) is the all-ones matrix, \( D^{*} \) is a diagonal matrix with independently and identically distributed entries \( \pm 1 \) and \( P^{*} \) is a random permutation matrix. In practice the last step, the permutation step, is sometimes omitted. To eliminate bias as a result of the random selection of sampling points, the analysis can be run more than once. For a total of \( r \) runs, the matrix \( B^{\prime} \) (Equation (3)) is generated \( r \) times to obtain a design matrix

\[
X = \begin{pmatrix}
B_{1}^{*} \\
\vdots \\
B_{r}^{*}
\end{pmatrix}.
\]

A graphical example of such a design matrix is given in Fig. 1. Each row of \( X \) contains a permutation vector \( x \) which is in turn substituted into Equation (2) for model evaluation. After model evaluation, the elementary effects are calculated according to Equation (1) and the mean is calculated over the \( r \) runs.

2.4 Fourier amplitude sensitivity test
The FAST is a global sensitivity analysis method pioneered by Cukier, Shuler and co-workers (Cukier et al. 1973, 1975, 1978; Schaibly and Shuler 1973), with improvements proposed, amongst others, in Saltelli and Bolado (1998); Fang et al. (2003); Xu and Gertner (2007, 2008). Fourier amplitude sensitivity test has long been used in analysing chemical systems (Cukier et al. 1973, 1975; Schaibly and Shuler 1973; Koda et al. 1979a, b; Campolongo et al. 2000; Saltelli et al. 2005; Zádor et al. 2006; Marino et al. 2008). The method is described in detail below, but the main idea behind FAST is this: parameters are allowed to oscillate around a point at an assigned frequency; the frequency can be isolated from the model output and allows the contribution of a parameter to the total model variation to be calculated.

As with the Morris method, we define a search function to step through the unit hypercube. Let \( \Omega \) be a vector of frequencies with elements pairwise co-prime. The \( i \)-th frequency, \( \omega_{i} \), corresponds to the \( i \)-th parameter of the model. Each parameter \( p_{i} \) is perturbed once per repetition. Generate a new search function to step through the unit hypercube. Let \( \Omega \) be a vector of frequencies with elements pairwise co-prime. The \( i \)-th frequency, \( \omega_{i} \), corresponds to the \( i \)-th parameter of the model. Each parameter \( p_{i} \) is perturbed once per repetition. Generate a new search function to step through the unit hypercube. Let \( \Omega \) be a vector of frequencies with elements pairwise co-prime. The \( i \)-th frequency, \( \omega_{i} \), corresponds to the \( i \)-th parameter of the model. Each parameter \( p_{i} \) is perturbed once per repetition.
parameter, \( p \). The frequency \( \omega \) ensures that each parameter oscillates at a unique frequency, so that the contribution of that parameter to the model variation can be isolated. The different \( \omega \) are chosen according to an algorithm by Cukier et al. (1973) so as to minimize the amount of interference between frequencies, thereby isolating the effect of a parameter perturbation in the model output. An example illustrating the method is given in Fig. 2.

The following search function gives a set of points sampled from the unit hypercube,

\[
x(s) = \frac{1}{2} + \frac{\arcsin(\sin(\Omega \cdot s))}{\pi},
\]

where \( s \) is defined on the interval \([-\pi, +\pi]\); \( s \) is also continuous on this interval, so in order to build a discrete design matrix, \( X \), \( N \) discrete, evenly spaced points need to sampled from \([-\pi, +\pi]\). The design matrix is then

\[
X = \begin{pmatrix}
x_1^T \\
x_2^T \\
\vdots \\
x_N^T
\end{pmatrix}.
\]

One way of determining the size of \( N \) is to define it to be free of interference up to order \( M \) if \( N \) satisfies

\[
N = 2M\omega_{\text{max}} + 1,
\]

where \( M \) is typically taken to be 4 or 6 and \( \omega_{\text{max}} = \max \Omega \) is the largest frequency that has been assigned (Cukier et al. 1973; Saltelli et al. 1999). This means that there will be little or no interference up to the \( M \)th harmonics for the chosen frequencies. For example, if \( M = 4 \) and \( \omega_{\text{max}} = 23 \) then \( N = 185 \). Such a choice of \( N \) would allow the search function to take on values of 0 or 1 for certain frequencies. In itself this is not a problem, but it does mean that if a normal distribution is used then \( \text{cdf}^{-1}(0) \to -\infty \) and \( \text{cdf}^{-1}(1) \to \infty \), which is a problem. The number of runs can then be incremented by some small integer. As noted (Cukier et al. 1973), \( N \) can be chosen to be even smaller: if there are \( k \) oscillating parameters then the number of model runs required is given by \( N \approx 2.6k^{2.5} \).

As with the Morris method, using an appropriate cdf and Equation (2), and then iterating over all the rows in the design matrix, values for \( p \) can be obtained. An example of periodic model output is given in Fig. 3 for two variables of the ADR model considered here.

![Figure 2. The workings of the search function used in the FAST algorithm illustrated with two example parameters, say \( V_1 \) and \( V_2 \), two maximal velocities of enzymatic rate equations. The diagram shows the relationship between the interval on which \( s \) is defined (top left-hand corner) and the parameter search space (lower right-hand corner). As an example, following the red line from a value of \( s \) to the right, a particular value is obtained for \( V_2^* \) which is oscillating at a frequency of \( \omega_2 = 5 \) between 0 and 1. Moving down the red line to the cumulative density function for \( V_2^* \), a value is obtained for \( V_2 \) in the bottom left-hand corner as a function of \( s \). If these values are combined with those of \( V_1 \), then the parameter space is obtained in the lower right-hand corner.](https://academic.oup.com/insilicoplants/article/3/1/diab014/6206764)
Once the output is obtained, a discrete Fourier transform (DFT) is performed, giving the complex coefficients, \( \xi_j \), for \( j = 0, 1, 2, \ldots, (N - 1)/2 \), where \( j \) denotes the positive or negative frequency and \( N \) is the number of parameter sets generated by the search function and is therefore also the number of times the model needs to be evaluated to produce the output.

Because the DFT is performed on real input data, the output is symmetric, i.e. \( \Re(\xi_j) = \Re(\xi_{-j}) \) and \( \Im(\xi_j) = -\Im(\xi_{-j}) \), so that the positive and negative frequencies are complex conjugates of each other.

The Fourier spectrum (\( \Lambda \)) is then defined as the set of all \( \Lambda_j \), where

\[
\Lambda_j = 2\left(\xi_j \cdot \xi_{-j}\right)
\]

(8)

\[
= 2(|\xi_j|^2)
\]

(9)

For every output in \( f(p) \), at a given position \( z \) and time \( t \) a DFT is performed,

\[
\Xi_{z,t} = \begin{pmatrix}
\xi_{1,0} & \xi_{1,1} & \cdots & \xi_{1,(N-1)/2} \\
\xi_{2,0} & \xi_{2,1} & \cdots & \xi_{2,(N-1)/2} \\
& \cdots & \cdots & \cdots \\
\xi_{m,0} & \xi_{m,1} & \cdots & \xi_{m,(N-1)/2}
\end{pmatrix}
\]

(11)

The expected values of the model output are

\[
E(f(p)) = \frac{1}{N} \sum_{i=1}^{N} f(p_i) = \begin{pmatrix}
|\xi_{0,0}| \\
|\xi_{1,0}| \\
& \cdots \\
|\xi_{m,0}|
\end{pmatrix}
\]

(12)

in other words, the mean value of the outputs is equal to the zero frequency coefficients. Recall that the zero frequency is real, i.e. it has zero imaginary part.

The total model covariance is

\[
\text{Cov}(f(p)) = \frac{1}{N} \sum_{i=1}^{N} ((f(p_i) - E(f(p_i))) \cdot (f(p_i) - E(f(p_i))))^T,
\]

(13)

and the variance is

\[
\text{Var}(f(p)) = \text{diag}(\text{Cov}(f(p))) = \sum_j \Lambda_{jj}
\]

(14)

The fraction of the total variance of the model that is due to the variance in a particular parameter \( p_i \) can then be defined as

\[
D_{p_i} = \frac{\text{Var}(f(p_i))}{\text{Var}(f(p))} = \frac{\sum_j \Lambda_{ji} \omega_i}{\sum_j \Lambda_{jj}},
\]

(15)

For \( l = 1, 2, \ldots \) and \( l \omega_i \leq (N - 1)/2 \). The value of \( \omega_i \) is the frequency at which the parameter \( p_i \) oscillates in the design; thus, the \( l \omega_i \) are integers that account for all the multiples of \( \omega_i \)—the harmonics of \( \omega_i \).

Because the elements of \( \Omega \) are pairwise co-prime, the contribution of the parameter \( p_i \) can be isolated in this way. From the above it is clear that \( D_{p_i} \) takes on a value between 0 and 1.

Fourier amplitude sensitivity test sensitivities have a purely statistical interpretation; in other words, they lack the clear operational interpretation of response coefficients of MCA. If a parameter is allowed
to vary in the defined range, then its contribution to the variation in the model output is given by the FAST sensitivity. Evaluation of the sensitivities therefore allows the parameters to be ranked according to their contribution to the model variance.

Biologically, this means that of all the possible factors in a set (e.g. a group of kinetic parameters) that could possibly affect model behaviour (e.g. homeostasis), those with the highest sensitivity indices will have the greatest effect. Another way of interpreting sensitivity indices, modified from Saltelli et al. (2005), could be as follows: suppose it is desirable that a certain model output is minimized (or maximized), then the first parameters that should to be changed, for the greatest effect, are the ones with the highest sensitivity indices.

2.5 Software and computation
All calculations were coded in the Python programming language, making use of the NumPy and SciPy libraries of numerical routines for scientific computing. The Morris and FAST sensitivity analysis methods were coded as Python classes. The discrete fast Fourier transform (FFT) routines from NumPy, numpy.fft, and probability distributions from the SciPy statistics module, scipy.stats, were used in the FAST analysis. The model was also implemented in a Python class and simulations were done with FiPy (Guyer et al. 2009, 2017) as described in the accompanying paper (Uys et al. 2021). Codes are available under an open-source BSD licence and are included as Supporting Information.

Computations were performed using Stellenbosch University’s high performance computing cluster. Parallel simulations were carried out using the ipyparallel framework from IPython (Pérez and Granger 2007). Due to the size of the resulting data sets (~5 GB), hierarchical data format (HDF) files were used for data storage, management and access using the Python h5py module (https://www.h5py.org). In all cases, model output was evaluated for a total of 512 (2^9) steps with a time delta of 0.2 per step to give an end time of 102.

2.6 Experimental design for Morris method
The design matrix (Equation (4)) was constructed with 6 levels and 30 runs (l = 6, r = 30). For each model evaluation (i.e. each row of the design matrix), the parameter vector was calculated with Equation (2). The parameters were varied by ±20%, i.e. a uniform distribution over the interval [0.8, 1.2] was used in Equation (2). Since the model has 12 parameters (k = 12), a total of 390 = r (k + 1) model evaluations were required to complete the analysis. The elementary effects were then calculated according to Equation (1), and the means calculated over the 30 runs. To further remove sampling bias, this analysis was repeated three times, and the results presented in Figs 5 and 6 reflect the means of the elementary effects from these three repetitions.

2.7 Experimental design for FAST
The vector of frequencies used in Equation (5) was Ω = (31, 87, 113, 151, 191, 223, 245, 265, 275, 281, 289, 293) (cf. Schaibly and Shuler 1973). Note that the length of Ω corresponds to the number of parameters. The number of sampling points N was taken as 2349 (2 × 4 × 293 + S, cf. Equation (7)) and the design matrix constructed according to Equation (6). As for Morris, a uniform distribution over [0.8, 1.2] was used in Equation (2) to calculate the parameter vectors. After model evaluation for 2349 runs, the model output was subjected to a DFT using NumPy and the FAST coefficients were calculated from Equation (15). To further reduce sampling bias, we employed the extended FAST analysis (Saltelli et al. 1999) by making two additional changes: first, introducing a random phase shift for each of the frequencies ω, when setting up the design matrix (cf. Equation (5)), and second, repeating the analysis three times and taking the means of the partial variance and total variance to calculate the FAST coefficients (cf. Equation (15)).

3. RESULTS
3.1 Sensitivity analysis with the Morris method
The core model describing sucrose accumulation in sugarcane within the ADR framework, which was introduced in the preceding paper (Uys et al. 2021), was subject to sensitivity analysis using Morris and FAST. To facilitate easy cross-reference, the pathway scheme of the model is repeated here (Fig. 4). The parameters that were manipulated were the maximal activities (limiting rates) of each of the biochemical reactions and intercompartmental transport steps. The justification for this was that these are the easiest to manipulate in an experimental setting (e.g. via genetic means through overexpression or knock-down of the genes for metabolic enzymes and transport proteins). Other parameters such as $K_m$ values for enzymes are much harder to change (if at all possible), and physical constants are set by nature. If the rationale of a model sensitivity analysis is to identify parameters that can possibly be manipulated (e.g. to increase sucrose accumulation in sugarcane), then the maximal activities are a natural choice.

The maximal activities enter the respective rate equations as multipliers; hence, the elementary effects quantify the changes in the model outputs (variables) as a result of perturbations in the activities of each of these steps. According to Equation (1), the elementary effects capture the magnitude as well as the direction of the change in a particular model output. Hence, the results of analysing the ADR model with the Morris method are visualized in two ways: Fig. 5 shows the signed values of the elementary effects to emphasize which variables are increasing and which ones are decreasing as a consequence of the perturbations, and Fig. 6 shows the absolute values to facilitate a direct comparison of the magnitude of the changes. For example, in Fig. 5, $S_{S_{vc}}$ responds positively to perturbations in $R_{S_{vc}}$ and $R_{S_{ps}}$ but negatively to $R_{S_{np}}$. As discussed below, it is impossible to obtain this information from the FAST method.

A number of features in Figs 5 and 6 deserve further mention. First, each of the little squares describing the effect of a reaction on a model variable actually summarizes a total of 25 100 elementary effects (50 points along the distance axis, z, and 502 points along the time axis, t). These effects change with position and time, and the figure thus provides a spatio-temporal overview of the magnitude of the sensitivities.

Second, reactions frequently have large effects on metabolites that are directly connected (e.g. the effect of $R_{s_{nu}}$ on $S_{nu}$, the effect of $R_{on}$ on $S_{on}$ and $P_{on}$, or the effect of $R_{on}$ on $M_{on}$ and $N_{on}$). Furthermore, the profiles for substrates and products of a particular reaction frequently mirror each other, as an increase in the activity of a reaction would lead
to a decrease in its substrate and increase in its product (cf. the effect of $R_3$ on $S_{ph}$ and $S_{sk}$).

Third, in the lower part of Figs 5 and 6, a diagonal is clearly visible, i.e. reactions tend to have a positive effect on their own flux. There are, however, some off-diagonal effects, which will be discussed further below. For example, reactions feeding carbon into the sink ($R_1$, $R_2$, $R_3$, and $R_5$) have a negative effect on hexose remobilization from the vacuole ($R_9$), while $R_{10}$ has very similar positive effects on $R_9$, $R_{10}$ and $R_{11}$.

Finally, some elementary effects are strong at the beginning of the simulation and weaken as the metabolism progresses, while other effects start out weak and become stronger. In some cases, effects also become more prominent in some parts of the stalk while decreasing in other parts as time progresses. Therefore, this analysis allows conclusions beyond, for example 'It decreases vacuolar sucrose accumulation' to quantifying the extent of this effect in both space and time.

3.2 Sensitivity analysis with FAST

The partial variates (FAST coefficients) calculated with using the FAST method are summarized in Fig. 7. For easy comparison, this figure has the same layout as Fig. 6 depicting the absolute values of the Morris elementary effects. There is, however, a crucial difference: FAST sensitivities take values on the interval 0 to 1; in some cases, they also sum...
Figure 5. Morris's signed elementary effects showing the sensitivity of concentrations and reaction rates to variation in the 12 maximal activities. The parameters that were varied are arrayed along the columns and the state variables (model output) are...
3.3 Comparing Morris and FAST

To compare Morris and FAST, it is necessary to look at Figs 6 and 7; as explained above, only relative comparisons are possible. In both figures a characteristic, diagonal band can be seen for the sensitivity of a reaction rate to a perturbation of its own maximal activity. This reflects the intuitive fact that all rate equations are naturally sensitive to their own parameters and has already been discussed under the Morris results.

Both methods also showed that parameter sensitivities change with geometry and time. Specifically, sensitivities show the characteristic ‘sawtooth’ profile that coincides with sugarcane node and internode positions and was also observed for many of the model variables (e.g. $S_{u}$, $S_{a}$, and $P_{v}$) in the time simulation of the model (Uys et al. 2021). Each moiety concentration pair, $M_{u}/N_{a}$ or $H_{u}/H_{a}$, had an identical sensitivity profile. The exception is for the signed Morris results (Fig. 5), where the profiles were identical but opposite in sign. This can be explained by the fact that the moiety couples are stoichiometrically linked; when one member of this pair increases, the other member will decrease by the same amount.

The sensitivities for $S_{u}$ and $P_{v}$ with respect to $R_{s}$ show a difference between Morris and FAST. Morris predicts that the sensitivities should, in general, increase with time, while FAST predicts the exact opposite. Another difference is that $R_{s}$ is virtually unresponsive to any perturbations according to Morris, but is sensitive to $R_{u}$, $R_{v}$, and itself according to FAST. These differences are discussed further below.

4. DISCUSSION

In this paper, we have applied two methods of global sensitivity analysis, i.e. the method of Morris and FAST, to the ADR model of sucrose accumulation in sugarcane that was developed in the accompanying paper (Uys et al. 2021). These methods allowed us to identify which parameters had the greatest influence on model outputs in spatio-temporal detail.

Any uncertainty in parameters translates into uncertainty or variation in model output. While parameters may vary legitimately within a range and this range is known, so that there is nothing uncertain about them, it is nevertheless desirable to quantify the contribution from a particular parameter to variation in the overall model output, which is something that both sensitivity analysis methods provide. In this case the uniform distribution [0.8, 1.2], which was used in the transfer function (Equation (2)), can be replaced with a statistical distribution of experimentally observed errors in order to better approximate the effect of parameter variations on model output.

Overall, there was very good agreement between the results from Morris and FAST, which lends additional credence to the results. However, as summarized in Section 3.3, there were some differences, which we now discuss further. In the case of $R_{s}$, one explanation would be that an OAT method is not strong enough to pick up any effects of perturbations on this step, probably because changes in $R_{s}$ would depend on changing more than one reaction simultaneously. In the case of $S_{u}$ and $P_{v}$, it is probably also best to trust the FAST predictions since FAST is a global sensitivity analysis method. Because of the non-linear nature of the model, the distribution of model outputs will be more accurate where combinations of parameters are perturbed as compared to an OAT design. Even so, the FAST implementation used here cannot deal with interacting parameters. If a significant proportion of the model variance is unaccounted for, this may be due to higher-order interactions between parameters and modifications of FAST accommodating this are available (Xu and Gertner 2007, 2008). This was, however, not required for the present analysis.

The greater reliability of FAST comes at an increased computational cost. In our setup, the experimental design for Morris required 390 model evaluations, while FAST required 6-fold more (2349 evaluations). As a consequence, it is recommended to use Morris in an exploratory analysis to identify the most important parameters, and then focus on those with a detailed analysis using FAST. In addition the number of runs in our Morris implementation ($r = 30$) was deliberately chosen quite high to minimize sampling bias. Reducing this number to 10 introduced slightly more variability into successive repeats of the analysis, but the overall conclusions were unchanged [see Supporting Information—Fig. S1]. This would reduce the computational requirements by a factor of three if resources are limited.

Global sensitivity analysis has a significant recent history in plant science, as exemplified by the following studies. Cerasuolo et al. (2016) developed a model of source–sink interactions in willow and analysed the genotype × environment effects using the Morris method. A number of studies employed the Agricultural Production Systems...
Figure 6. Morris’s absolute elementary effects showing the sensitivity of concentrations and reaction rates to variation in the 12 maximal activities. The layout is the same as for Fig. 5; the figure merely re-plots the absolute values of the data. The colour scale is normalized from zero (dark blue) to the maximal value indicated on the right-hand side (yellow). Raw data are provided as Supporting Information.
Figure 7. Fourier amplitude sensitivity test coefficients showing the sensitivity of concentrations and reaction rates to variation in the 12 maximal activities. The layout is the same as for Fig. 5. The colour scale is normalized from zero (dark blue) to the one (yellow). Simulations were performed as described in the main text; model code and the raw data for this figure are provided as Supporting Information.
SIMulator (APSIM) platform (Keating et al. 2003; Holzworth et al. 2014) to develop crop improvement strategies: Elli et al. (2020) used the Morris method to identify suitable Eucalyptus traits for adaptation to climate change; and two studies (Sexton et al. 2017; Gladish et al. 2018) used so-called emulators, which are statistical approximations of more complex models that require fewer model evaluations (and are thus less computationally expensive to run), yet yield variance-based global sensitivities of similar accuracy. Sexton et al. (2017) applied this approach to determine the sensitivity of a sugarcane growth model (APSIM-Sugar) to cultivar trait parameters for cultivation in two production environments that differed in terms of their climatic and soil conditions. Importantly, all these cited studies employ phenomenological models. To our knowledge, ours is the first report of a model that combines mechanistic simulation of enzyme-catalysed reactions and intercompartmental transport with biophysical models of phloem flow within the ADR framework, and we have shown that global sensitivity analysis methods such as Morris and FAST are equally suited to such ADR models.

4.1 Reactions affecting sucrose accumulation

The reactions affecting sucrose accumulation merit a more detailed discussion. The variables $S_{\text{s}}$ and $P_{\text{su}}$ illustrate that sensitivity to parameters can vary quite dramatically across the stalk. At the end of the simulation ($t = 102$), the sensitivity of both of these concentrations to perturbations in $R_6$ changed from near zero at the top and bottom of the stalk to near maximal in the centre of the stalk (Figs 5–7). This may suggest that vacuolar sucrose is least sensitive to vacuolar breakdown when concentrations are very low or near saturation (recall that $S_{\text{s}}$ increased from the top to the bottom of the stalk, cf. Figure 5 in Uys et al. 2021).

A central aim of our model was to describe how sucrose ($S$ in the model) moves from where it is produced to where it is stored in the stalk to near maximal in the centre of the stalk (Figs 5–7). This is most sensitive to reaction 2 (at the nodes and least in the middle of an internode, and inversely with respect to reaction 3. Stated in another way, the further sucrose is from the source along an internode, the more palpable parenchymal uptake becomes. This is important, because it impacts the creation of sucrose gradients in the phloem.

At the nodes, $S_{\text{s}}$ is most sensitive to phloem unloading, while in the middle of an internode, $S_{\text{s}}$ is the most sensitive to the breakdown into hexoses (reaction 3). This is potentially important, because if sucrose accumulation is to increase, the effect of sucrose inversion should be as small as possible. This effect becomes bigger as the internodes mature; the sensitivity to breakdown increases along the stalk. $S_{\text{s}}$ is consistently sensitive to its own breakdown by vacuolar acid soluble invertase ($R_6$), and to a lesser extent to $R_7$ (phloem unloading), $R_8$ (vacuolar uptake) and $R_9$ (regeneration of the vacuolar proton gradient). Sucrose breakdown persists in the vacuole, with the Morris elementary effects for $S_{\text{s}}$ and $P_{\text{su}}$ numerically similar but opposite in sign with respect to $R_6$, reflecting the fact that $S_{\text{s}}$ is the substrate of $R_6$, while $P_{\text{su}}$ is the product.

In almost all cases, reaction rates are the most sensitive to their own parameters. One exception to this is reaction 8 (the step for transporting $S$ into the vacuole), which, according to both FAST and Morris, is sensitive to phloem unloading ($R_7$) and the creation of proton gradient ($R_9$). The sensitivity shifts from $R_7$ to $R_9$ as the simulation time progresses. Both of these are transport steps, with reaction 7 requiring an input of metabolic energy to regenerate the proton gradient across the tonoplast. This sensitivity is to be expected, since reaction 8 cannot accumulate $S_{\text{s}}$ against a gradient without an input of free energy, and without the coupling in reaction 7, this proton gradient would rapidly deplete.

Overall, the simulations suggest that acid soluble invertase ($R_6$) has the largest effect on vacuolar sucrose concentrations ($S_{\text{s}}$), with increases in the enzyme activity leading to a decrease in $S_{\text{s}}$ levels (the Morris elementary effects are negative, cf. Fig. S), and vice versa. A way to test this would be to lower the $V_j$ of $R_6$ by 20 or 30 %, and to re-compute and re-visualize the effects on the model behaviour in space and time to verify that $S_{\text{s}}$ is indeed increased. The same could be done for the other reactions that do have an impact but show smaller sensitivity coefficients ($R_7, R_9, R_6$ see above). An alternative is to consider derived variables that relate to the net accumulation of sucrose in various compartments. For example, we could consider:

- $y_1 = R_6 - R_3$, quantifying net carbohydrate accumulation in the vacuole;
- $y_2 = R_9 - R_6$, quantifying net synthesis of sucrose in parenchymal cells (symplast plus vacuole) minus net degradation, which could also be considered as the extent of futile cycling of sucrose.

The sensitivity analysis could then be repeated for these variables, which each relate to a different aspect of sucrose accumulation. This is left for future work. In general, though, absolute values of sensitivity coefficients reported here should be treated with caution as this is a core model, and we restrict ourselves to the description of trends.

5. CONCLUSIONS

The ADR framework and model of sucrose accumulation in sugarcane developed in the previous paper (Uys et al. 2021) has been successfully subjected to global sensitivity analysis using two generic methods, i.e. Morris and FAST. This allowed us to elucidate the sensitivities of model outputs to input parameters. While this is still a core model that does not include the detail of every biochemical reaction, the model was constructed to capture the most important biochemical features including futile cycling of sucrose, the energetics of sucrose accumulation and interconversion between mono- and disaccharides, as well as compartmentation. Specifically, the greater detail in terms of spatio-temporal separation, compartmentation and the mechanics of phloem flow bring it a step closer to implementing a full-scale realistic model of sucrose accumulation in sugarcane.
SUPPORTING INFORMATION

Supplementary text document

- fast_supplementary.pdf Summary document with additional figures and instructions for setting up the computational environment and performing the simulations.

Code

- adr.py Model definition.
- morris.py Class implementation of the Morris method.
- fast.py Class implementation of the FAST method.
- run_morris.py Script for model execution for the Morris method.
- run_fast.py Script for model execution for FAST.
- calc_morris_effects.py Script for calculating the Morris elementary effects.
- calc_fast_variates.py Script for calculating the FAST coefficients.
- LICENSE.txt Software licence.

Data

- morris_ee_mean.npy Data for mean values of Morris's elementary effects.
- fast_variates_mean.npy Data for mean values of FAST coefficients.

SOURCES OF FUNDING

This work is supported in part by funds from the South African National Research Foundation (NRF: # 81129 and # 114748) and the South African National Bioinformatics Network. The funding bodies played no role in the design of the study, collection, analysis and interpretation of data, decision to publish and in writing the manuscript.

CONFLICT OF INTEREST

None declared.

CONTRIBUTIONS BY THE AUTHORS

L.U. and J.M.R. conceived the analysis, performed the simulations, analysed and visualized the results. J.M.R. procured funding. L.U. wrote the first draft of the manuscript. J.M.R. and J.-H.S.H. supervised the study. All authors reviewed and edited the manuscript.

DATA AVAILABILITY

Model code and simulation data are provided as Supporting Information.

LITERATURE CITED

Campolongo F, Tarantola S, Saltelli A. 2000. Sensitivity analysis as an ingredient of modeling. Stat. Sci. 150:377–395.

Cerasuolo M, Richter GM, Richard B, Cunniff J, Girbau S, Shield I, Purdy S, Karp A. 2016. Development of a sink-source interaction model for the growth of short-rotation coppice willow and in silico exploration of genotype × environment effects. Journal of Experimental Botany 670(3):961–977.

Cukier RI, Fortuin CM, Shuler KE, Petschek AG, Schaibly JH. 1973. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I. Theory. Journal of Chemical Physics 590:3873–3878.

Cukier RI, Levine HB, Shuler KE. 1978. Nonlinear sensitivity analysis of multivariate model systems. Journal of Computational Physics 260:1–42.

Cukier RI, Schaibly JH, Shuler KE. 1975. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. III. Analysis of the approximations. Journal of Chemical Physics 630:1140–1149.

Elli EF, Huth N, Sentelhas PC, Carneiro RL, Alves CA. 2020. Global sensitivity-based modelling approach to identify suitable Eucalyptus traits for adaptation to climate variability and change. In Silico Plants 2:diaa003; doi:10.1093/insilicoplants/diaa003.

Fang S, Gertner GZ, Shinkareva S, Wang G, Anderson A. 2003. Improved generalized Fourier amplitude sensitivity test (FAST) for model assessment. Statistics and Computing 130:221–226.

Gladish DW, Darnell R, Thorburn PJ, Haldankar B. 2018. Emulated multivariate global sensitivity analysis for complex computer models applied to agricultural simulators. Journal of Agricultural, Biological and Environmental Statistics 240:130–153.

Guyer JE, Wheeler D, Warren JA. 2009. FiPy: partial differential equations with Python. Computing in Science & Engineering 11:6–14.

Guyer JE, Wheeler D, Warren JA. 2017. FiPy manual, release 3.1.3. Gaithersburg: Materials Science and Engineering Division and the Center for Theoretical and Computational Materials Science, Material Measurement Laboratory, NIST. http://www.ctcms.nist.gov/fipy (1 March 2021).

Heinrich R, Rapoport TA. 1974. A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. European Journal of Biochemistry 42:89–95.

Holzworth DP, Huth NI, deVoil PG, Zurcher EJ, Herrmann NI, McLean G, Chenu K, van Oosterom EJ, Snow V, Murphy C, Moore AD, Brown H, Whish JPM, Verrall S, Fainges J, Bell LW, Peake AS, Poulton PL, Hochman Z, Thorburn PJ, Gaydon DS, Dalgliesh NP, Rodriguez D, Cox H, Chapman S, Doherty A, Teixeira E, Sharp J, Cichota R, Vogeler I, Li FY, Wang E, Hammer GL, Robertson MJ, Dimes JP, Whitbread AM, Hunt J, van Rees H, McClelland T, Carberry PS, Hargreaves JNG, MacLeod N, McDonald C, Hawksford J, Wedgwood S, Keating BA. 2014. APSIM – evolution towards a new generation of agricultural systems simulation. Environmental Modelling & Software 62:327–350.

Kacser H, Burns JA. 1973. The control of flux. Symposia of the Society for Experimental Biology 27:65–104.

Keating BA, Carberry PS, Hammer GL, Probert ME, Robertson MJ, Holzworth D, Huth NI, Hargreaves JNG, Meinke H, Hochman Z, McLean G, Verburg K, Snow V, Dimes JP, Silburn M, Wang E, Brown S, Bristow KL, Asseng S, Chapman S, McCown RL, Freebairn DM, Smith CJ. 2003. An overview of APSIM, a model designed for farming systems simulation. European Journal of Agronomy 180:267–288.

Koda M, Dogru AH, Seinfeld JH. 1979a. Sensitivity analysis of partial differential equations with application to reaction and diffusion processes. Journal of Computational Physics 300:259–282.
Koda M, Mcrae GJ, Seinfeld JH. 1979b. Automatic sensitivity analysis of kinetic mechanisms. *International Journal of Chemical Kinetics* **110**:427–444.

Komor E. 2000. Source physiology and assimilate transport: the interaction of sucrose metabolism, starch storage and phloem export in source leaves and the effects on sugar status in phloem. *Australian Journal of Plant Physiology* **27**:497–505.

Marino S, Hogue IB, Ray CJ, Kirschner DE. 2008. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology* **254**:178–196.

Morris MD. 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics* **33**:161–174.

Pérez F, Granger BE. 2007. IPython: a system for interactive scientific computing. *Computing in Science & Engineering* **90**:21–29.

Qian G, Mahdi A. 2020. Sensitivity analysis methods in the biomedical sciences. *Mathematical Biosciences* **323**:108306.

Saltelli A, Bolado R. 1998. An alternative way to compute Fourier amplitude sensitivity test (FAST). *Computational Statistics and Data Analysis* **260**:445–460.

Saltelli A, Ratto M, Tarantola S, Campolongo F. 2005. Sensitivity analysis for chemical models. *Chemical Reviews* **105**:2811–2828.

Saltelli A, Tarantola S, Chan KPS. 1999. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics* **410**:39–56.

Schaibly JH, Shuler KE. 1973. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. II applications. *Journal of Chemical Physics* **590**:3879–3888.

Sexton J, Everingham Y, Inman-Bamber G. 2017. A global sensitivity analysis of cultivar trait parameters in a sugarcane growth model for contrasting production environments in Queensland, Australia. *European Journal of Agronomy* **88**:96–105.

Uys L, Hofmeyr JHS, Rohwer JM. 2021. Coupling kinetic models and advection-diffusion equations. 1. Framework development and application to sucrose translocation and metabolism in sugarcane. *In Silico Plants* 3:diab013. doi:10.1093/insilicoplants/diab013.

van Bel A. 1993. Strategies of phloem loading. *Annual Review of Plant Physiology* **440**:253–281.

Walsh KB, Sky RC, Brown SM. 2005. The anatomy of the pathway of sucrose unloading within the sugarcane stalk. *Functional Plant Biology* **32**:367–374.

Xu C, Gertner GZ. 2007. Extending a global sensitivity analysis technique to models with correlated parameters. *Computational Statistics and Data Analysis* **51**:5579–5590.

Xu C, Gertner GZ. 2008. A general first-order global sensitivity analysis method. *Reliability Engineering and System Safety* **93**:1060–1071.

Zádor J, Zsély IG, Turányi T. 2006. Local and global uncertainty analysis of complex chemical kinetic systems. *Reliability Engineering and System Safety* **910**:1232–1240.