Foliar Uptake Models for Biocides: Testing Structural and Practical Identifiability

Enrico Sangoia, Federica Cattanib, Faheem Padiab, Federico Galvanina\*

aDepartment of Chemical Engineering, University College London (UCL), Torrington Place, WC1E 7JE London, United Kingdom

bProcess Studies Group, Global Sourcing and Production, Syngenta, Jealott's Hill International Research Centre, Berkshire, Bracknell, RG42 6EY, United Kingdom

f.galvanin@ucl.ac.uk

Abstract

As the global population grows and resources become scarcer, ensuring adequate food production is an urgent challenge. Developing safer biocides is crucial for optimizing crop yields and meeting rising food demands. Mathematical models play a pivotal role in understanding complex biological systems. This work focuses on developing a reliable model for characterizing the process of foliar biocide uptake in plants. A systematic modelling strategy involves the following steps: 1) formulating candidate models for foliar uptake; 2) conducting identifiability tests to verify that candidate model parameters can be determined from observations; 3) selecting the best model based on a-posteriori statistics from observations; 4) design experiments for achieving a precise estimation of model parameters from data. This paper presents a study on structural and practical identifiability of compartmental models for foliar biocide uptake (steps 1 and 2), considering experimental data limitations and initial condition variability. These findings will guide further model-based experimentation.

**Keywords**: model identification, foliar uptake, identifiability analysis, compartmental models.

* 1. Introduction

As the world's population continues to grow and the planet's resources remain limited, ensuring sufficient food production becomes a crucial challenge both in the present and for the coming decades. In tackling this issue, the development of improved and safer biocides will be essential to optimize crop yields and meet the increasing demand for food (Sharma et al., 2019). Mathematical models prove to be valuable tools to better understand the phenomena at play and formulate innovative solutions for the creation of new agricultural products. Several studies have been conducted on the foliar uptake of pesticides to develop mathematical models capable of describing the system behaviour. Among the models in the literature there are phenomena-specific empirical correlations (Forster et al., 2004), compartmental models (Bridges and Farrington, 1974), and diffusion-based models (Tredenick et al., 2019).

However previous studies did not assess systematically whether the model parameters can be accurately and precisely estimated, a key aspect to consider to ensure reliability of model predictions. This work aims at developing a mathematical model that can be effectively applied to describe the system and predict the dynamic uptake profile on different combinations of plant and product, while ensuring identifiability of the model parameters, i.e. that they can be uniquely determined from the system inputs and outputs (Miao et al., 2011), including in the model building strategy also the practical experimental limitations that will act as constraints when calibrating the model.

* 1. Methodology

The general model identification procedure that is used in this project to model the foliar uptake of pesticides, based on Franceschini and Macchietto (2008), is presented in Figure 1. The whole modelling procedure is divided in 5 key steps: i) the formulation of candidate models, ii) the conduction of model identifiability tests, iii) model-based design of experiments (MBDoE) for model discrimination, iv) MBDoE for precise parameter estimation, and v) the model validation. This paper focuses on the step (2) “model identifiability tests”, highlighted in Figure 1.



Figure 1 Model building strategy applied in the project, based on Franceschini and Macchietto (2008).

The general form considered for a dynamic model formulated as a set of differential and algebraic equations is:

|  |  |
| --- | --- |
|  | (1) |

In Eq. (1), is a vector of state variables, the vector of predicted model outputs, the vector of known system inputs, the model parameters vector. The foliar uptake models that will be presented in this work are expressed as systems of ODEs, therefore can be expressed as in Eq. (1).

The question that is under assessment when testing the identifiability of a mathematical model is if the model parameters can be uniquely determined from the given system input and the measurable system output (Miao et al., 2011). Identifiability can be assessed in different ways and identifiability tests can be classified in two main categories: *a-priori* structural identifiability (not requiring preliminary experimental data) and *a-posteriori* practical identifiability tests. An a-priori structural analysis test is conducted in ideal conditions, assuming total observability of the system, i.e. all the states can be observed at any time. On the other hand, a-posteriori practical identifiability test considers limitations on the quantity and quality of the data that can be practically retrieved from the system and that will be used for the estimation of the model parameters.

* + 1. A-priori structural identifiability

Several methods can be applied for testing structural identifiability, among which Taylor’s series, generating series, similarity transformation, direct test, differential algebra approaches. Miao et al. (2011) and Chis et al. (2011) give an extended overview of the different methods, discussing also advantages and disadvantages. Chis et al. (2011) showed that most of the methods scale badly with the number of states and parameters, and in this work the differential algebra approach is chosen because it was shown to have less limitations than the other methods. To test a priori structural identifiability with the differential algebra method the software DAISY (Bellu et al., 2007) is used.

* + 1. A-posteriori practical identifiability

Practical identifiability has been tested with the correlation matrix method. The correlation matrix approach looks at the correlation between pairs of parameters in the model by means of the correlation coefficient between parameters and . To build the correlation matrix , the -th entry is obtained from the variance-covariance matrix of the model parameters as in Eq. (2).

|  |  |
| --- | --- |
|  | (2) |

This identifiability technique is applied a-posteriori since it requires experimental data. In fact, the matrix is obtained starting from the dynamic sensitivity matrix , which brings the information about the model structure and the distribution of the sampling points, and the variance-covariance matrix of the measurements , characterizing the variability in the experimental data. The sensitivity of the *i*-th response to the *j*-th parameter at the *k*-th sampling time is calculated as in Equation (3).

|  |  |
| --- | --- |
|  | (3) |

The matrices and are combined to evaluate the Fisher Information Matrix (**FIM**), Eq.(4). The matrix is then obtained as the inverse of the FIM (Eq. 5).

|  |  |  |
| --- | --- | --- |
|  | (4) |  |
|  | (5) |

The correlation coefficient describes how closely linked the parameters and are. From the definition in Equation (2), it can be seen that high values of are associated to parameter pairs having a covariance comparable with their variances. Values of the correlation coefficient higher than 0.99 lead to singular FIM, therefore are a symptom of practical non-identifiability (Rodriguez-Fernandez et al. 2006).

* 1. Model description

A compartmental model for foliar uptake of pesticides is formulated starting from a description of the system structure, as shown in Figure 2. The system is divided into the following compartments: droplet, store, cuticle, and leaf tissue. Three model structures (A, B, C) will be considered in this study. In case C, an additional compartment “environment” is included in the model, representing the AI lost form the droplet due to volatility, while other loss terms preventing the AI penetration such as photo- or chemical-instability of the molecule are represented with the transfer rate . The store compartment refers to AI crystallized on the leaf surface. The model equations are a system of ODEs describing the dynamic change in mass of AI in the different compartments, as expressed in Equation (6).

|  |  |
| --- | --- |
|  | (6) |

Figure 2 Graphical representation of the three compartmental model structures A, B and C. The pins departing from the compartments indicate the observed states in the system.

In (7) [µg] indicates the mass of AI in compartment, and [µg/min] the transfer rate of AI from compartment to compartment. With respect to the generic formulation presented in Eq. (1), the vector of state variables is, and the model parameters are . No inputs are present in the system. The vector of the observables output changes among the three models: in the case A and C three states are observed, i.e. (being ), and . In case B only the total amount of AI inside the leaf is measured, not decoupling the cuticle and the cellular tissue, so the observed states are and the sum .

* 1. Results

The results of the identifiability tests are here presented along with a discussion of the findings and a comparison between the structural and practical identifiability results.

* + 1. A-priori structural identifiability

For the structural identifiability analysis let indicate with the true value of the parameters . In the context of structural identifiability a model is defined globally identifiable if the parameters are uniquely distinguished, locally identifiable if a finite number of solutions larger than one is obtained for , and nonidentifiable if the system input-outputs lead to an infinite number of parameters values (Bellu et al., 2007). The results of the structural identifiability test conducted with the software DAISY are summarised in Table 1. Model parameters result to be non-identifiable in all the three scenarios considered if no information about initial conditions is available. From this result it can be concluded that knowing the initial conditions of the system is crucial to ensure model identifiability, otherwise parameters would not be identifiable even in the ideal case of noise-free measurements and total observability of the system.

When initial conditions of the system are known, the model results to be globally identifiable in scenario A, locally identifiable in scenario B and non-identifiable in the scenario C. In scenario C the model is non-identifiable even with known initial conditions

Table 1 Summary of the a-priori structural identifiability results obtained with DAISY for the three scenarios A, B and C considering known and unknown initial conditions in the system.

|  |  |  |  |
| --- | --- | --- | --- |
| **Structural identifiability** | **Scenario A** | **Scenario B** | **Scenario C** |
| Unknown initial conditions | Model non-identifiable | Model non-identifiable | Model non-identifiable |
| Known initial conditions | Model globally identifiable | Model locally identifiable | Model non-identifiable |

 

 (a) (b)

Figure 3 (a) The t-values of the parameter estimates confidence intervals, compared to the reference t-test at 0.95 significance. (b) Heatmap showing the parameters correlation matrix for scenario A.

because the parameters and cannot be distinguished and infinite solutions in the form of Eq. (7) are obtained

|  |  |
| --- | --- |
|  | (7) |

The structural analysis results described above indicate that these two parameters cannot be uniquely determined from the system outputs considered.

* + 1. A-posteriori practical identifiability

The correlation matrix method presented in the methodology section is a local test performed around a nominal value of the parameters . The values of used in this study are obtained from a preliminary parameter estimation on foliar uptake data provided by Syngenta and a *t*-test with 95% significance is conducted to assess the precision of parameter estimates. The *t*-test results are represented in Figure 3a, where low *t*-values compared to a reference are associated to parameters with very large variance. In all the scenarios the three parameters , and are difficult to estimate with a good precision from data, while the other parameters have acceptable *t*-values i.e. larger than *t*-reference. In scenario B, differently form A and C, also parameters and fail the *t*-test, and this result is in good agreement with the a-priori structural analysis presented in section 4.1. In fact, the result obtained for model B suggests that the transport-barrier role of the cuticle is hard to describe with this compartmental model if experimental information about the AI distribution within the leaf structure is not available.

The parameter correlation matrix is reported with the heatmap in Figure 3b only for scenario A for sake of conciseness. The model passed the a-priori structural identifiability test under these conditions, however when looking at the practical identifiability results, strong correlations between parameters emerge. The correlation coefficient among all the pairs of parameters in the set {, , } result to be 0.99 or higher, indicating practical non-identifiability of these parameters (Rodriguez-Fernandez et al. 2006). From the *t*-test (Fig. 3a) high variance for these parameters was observed, while the analysis of correlation coefficient (Eq. 2) in the correlation matrix (Fig. 3b) lead to conclude that also their covariance is extremely large. From these results it emerges that higher model complexity must be sustained by more experimental information to ensure the identifiability of the full set of model parameters. Other strategies such as model reparameterization can be applied to overcome the identifiability issues (Quaglio et al., 2020).

* 1. Conclusions

In this paper, a systematic framework to develop a mathematical model for the description of foliar uptake of biocides is presented, focusing on the first steps of the methodology shown in Figure 1. The model building strategy adopted includes model parameters identifiability testing, which is done both a-priori and a-posteriori. Verifying model identifiability is an essential prerequisite to apply the model in the following steps of the procedure, i.e. MBDoE and parameter estimation. Model identifiability has been tested in three candidate models with different number of parameters and observed states in the system. Model B and C were not globally identifiable from the structural analysis, while model A passed the structural test but not the practical identifiability test. These results indicate that the model complexity achievable is strongly limited by the observability on the system. Future work will include the formulation of alternative model structures, also considering diffusion-based models, to ensure parameter identifiability as a requirement for the application of model-based design of experiments techniques in foliar uptake.

Acknowledgements

The authors gratefully acknowledge the support of the Department of Chemical Engineering, University College London, UK and Syngenta.

References

Bellu, G., Saccomani, M.P., Audoly, S., D’Angiò, L., 2007. DAISY: A new software tool to test global identifiability of biological and physiological systems. Computer Methods and Programs in Biomedicine 88, 52–61.

Bridges, R.C., Farrington, J.A., 1974. A compartmental computer model of foliar uptake of pesticides. Pestic. Sci. 5, 365–381

Chis, O.-T., Banga, J.R., Balsa-Canto, E., 2011. Structural identifiability of systems biology models: A critical comparison of methods. PLoS ONE 6.

Forster, W.A., Zabkiewicz, J.A., Riederer, M., 2004. Mechanisms of cuticular uptake of xenobiotics into living plants: 1. Influence of xenobiotic dose on the uptake of three model compounds applied in the absence and presence of surfactants into Chenopodium album, Hedera helix and Stephanotis floribunda leaves. Pest Manag. Sci. 60, 1105–1113

Franceschini, G., Macchietto, S., 2008. Model-based design of experiments for parameter precision: State of the art. Chemical Engineering Science, Model-Based Experimental Analysis 63, 4846–4872.

Miao, H., Xia, X., Perelson, A.S., Wu, H., 2011. On Identifiability of Nonlinear ODE Models and Applications in Viral Dynamics. SIAM Review 53, 3–39

Quaglio, M., Fraga, E.S., Galvanin, F., 2020. A diagnostic procedure for improving the structure of approximated kinetic models. Comput. Chem. Eng. 133, 106659.

Rodriguez-Fernandez, M., Egea, J.A., Banga, J.R., 2006. Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. BMC Bioinformatics 7, 483

Sharma, A., Kumar, V., Shahzad, B., Tanveer, M., Sidhu, G.P.S., Handa, N., Kohli, S.K., Yadav, P., Bali, A.S., Parihar, R.D., Dar, O.I., Singh, K., Jasrotia, S., Bakshi, P., Ramakrishnan, M., Kumar, S., Bhardwaj, R., Thukral, A.K., 2019. Worldwide pesticide usage and its impacts on ecosystem. SN Appl. Sci. 1, 1446

Tredenick, E.C., Farrell, T.W., Forster, W.A., 2019. Mathematical Modelling of Hydrophilic Ionic Fertiliser Diffusion in Plant Cuticles: Lipophilic Surfactant Effects. Plants 8, 202.