



Modelling in agronomy

Christophe Lecarpentier

► To cite this version:

Christophe Lecarpentier. Modelling in agronomy: Model calibration : how to estimate model parameters. Master. France. 2022. hal-03789540

HAL Id: hal-03789540

<https://hal.science/hal-03789540>

Submitted on 27 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



➤ Modeling in agronomy

Model calibration : how to estimate model parameters

Christophe Lecarpentier, chercheur INRAE

Unité Génétique Diversité et Ecophysiologie des Céréales (GDEC)

christophe.lecarpentier@inrae.fr

> Plan du cours

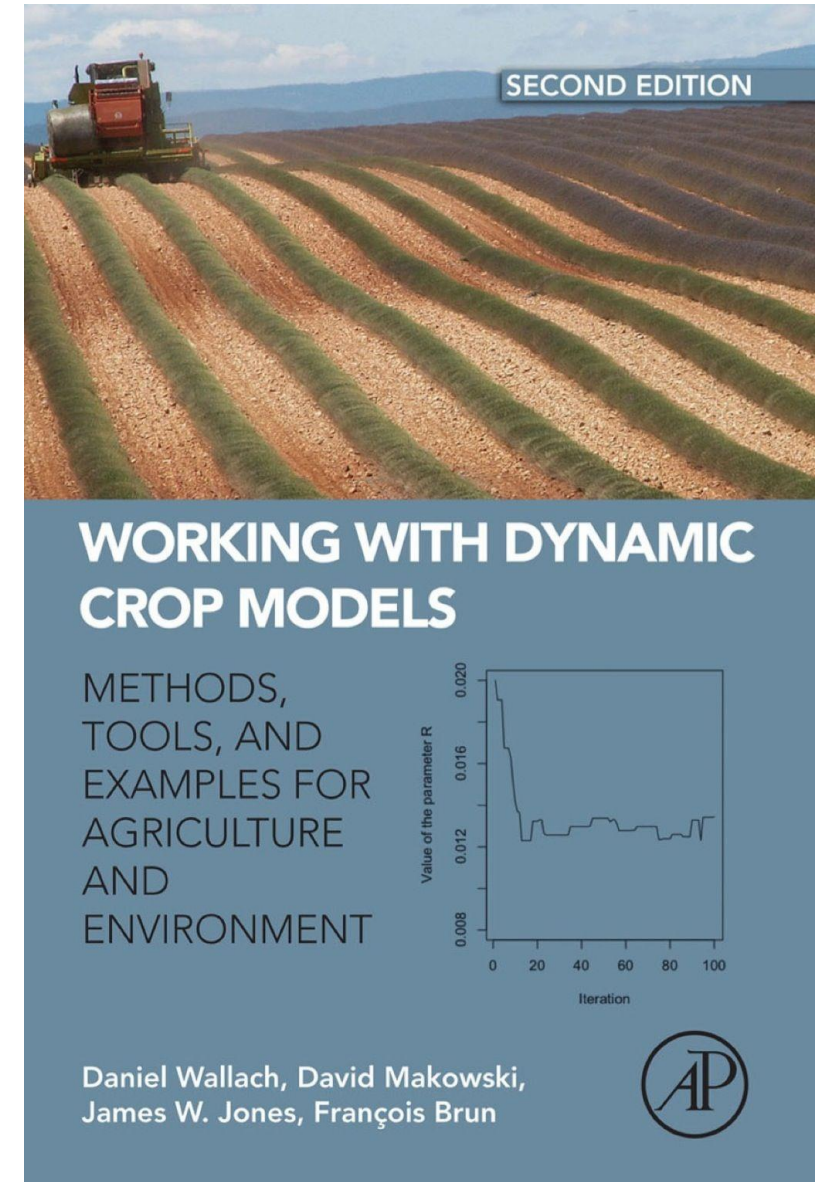
Modeling in agronomy : Parameter estimation

1. Introduction

1. Reminder of modeling bases
2. What is a parameter ?
3. What is model calibration ?
4. Why do we need calibration ?

2. Estimate the parameters of a function

3. Estimate the parameters of a complex model





Introduction

Definitions

Model ? Parameter ?

Models for agronomy

Calibration ?

Why do we need to calibrate ?

What is a model ?



> Introduction

Reminders: what is a model ?

A model is a **simplified formalisation** of an existant process.

The considered process is represented only by this simplified relation and then we **compare the prediction with the reality** to evaluate the relevance of our formalisation.

« All models are wrong, but some are useful »

Georges E. P. Box

A model is not designed to be "true" or to represent reality, but to answer to a specific use.
One of the major interest of modeling is to isolate formalisms and to test them successively

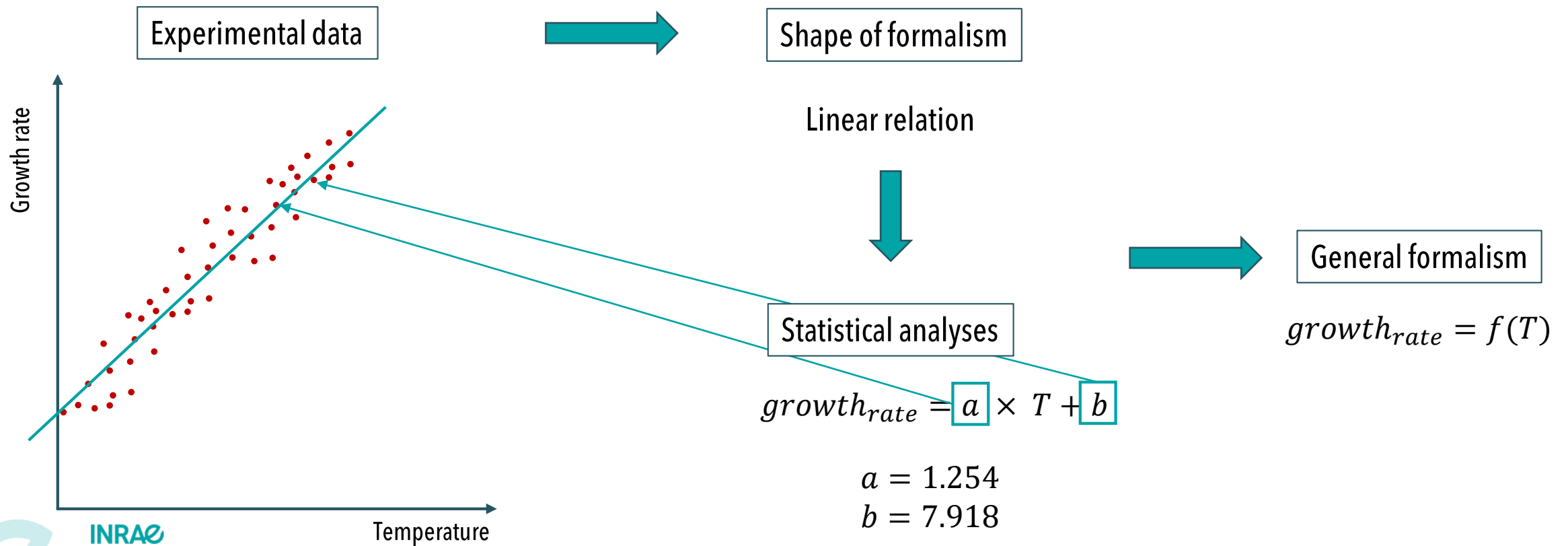
If the representation of the process is too far from experimental data:

- We test **another formalism**
- We add a second relation to complete the first one

> Introduction

Reminders: what is a model ?

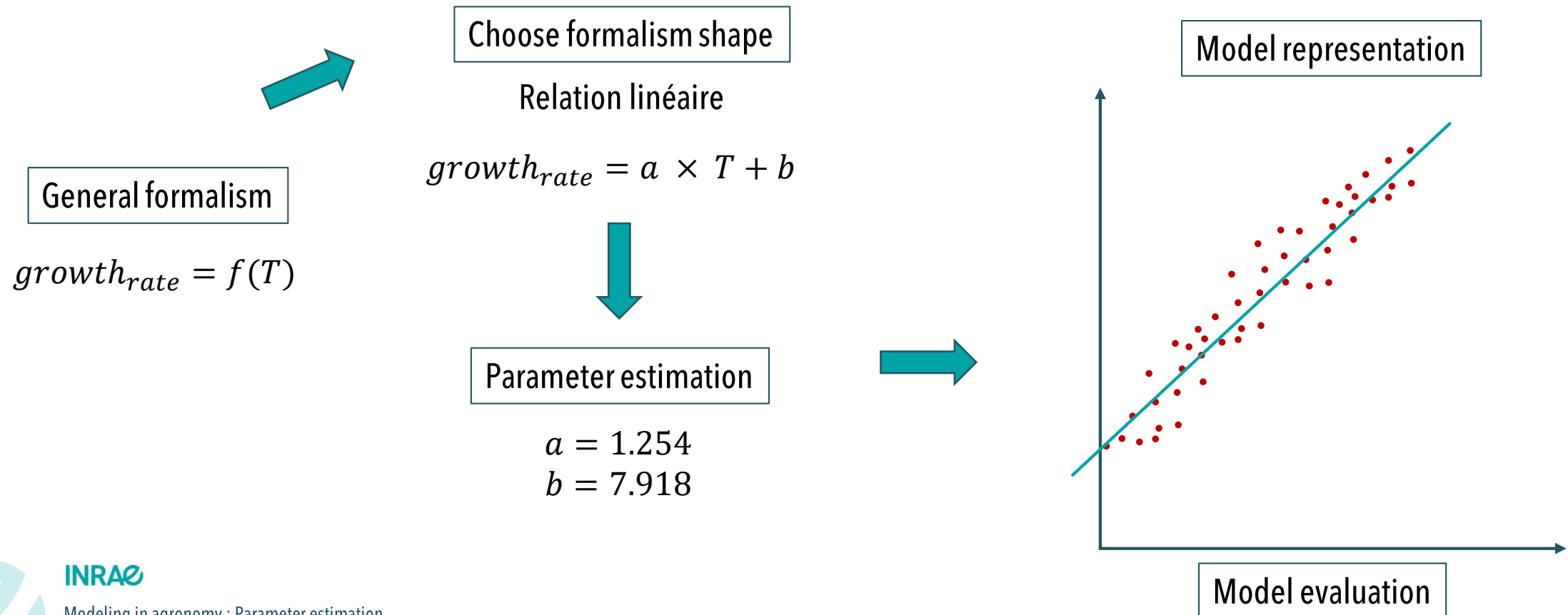
Modelisation follows the opposite reflexion pathway to **experimentation**



> Introduction

Reminders: what is a model ?

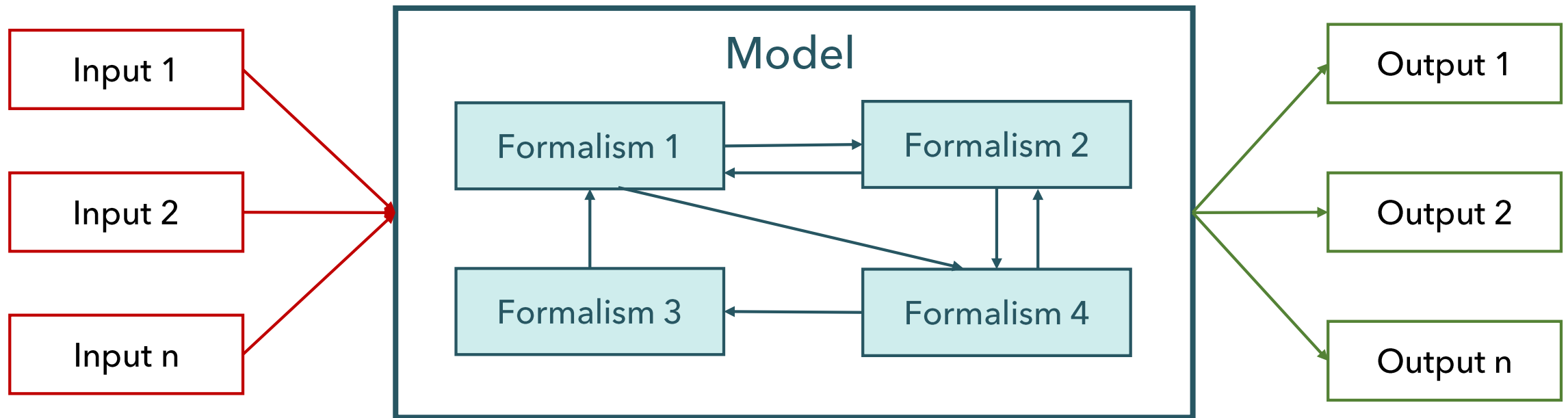
Modelisation follows the opposite reflexion pathway to **experimentation**



> Introduction

What is a model ?

Usually, models are more complex than just a simple equation and are composed of different equations that interact with each other



> Introduction

How to build a model ?

Model construction from A to Z is composed of multiple steps:

Example of phloem transport from a source to a sink within a plant (Minchin, 1993, JExpBot)

1. Model conception

The flow between **source** and **sink**:

- is generated by the difference in solute concentration between the two entities
- depends on the dimensions of vascular network



> Introduction

How to build a model ?

Model construction from A to Z is composed of multiple steps:

Example of phloem transport from a source to a sink within a plant (Minchin, 1993, JExpBot)

1. Model conception

2. Mathematical formalisation

The flow between **source** and **sink** is generated by the difference in solute concentration between the two entities and depends on the dimensions of vascular network

$$flow = \frac{(P_0 - P_1)\pi x^4}{8L\eta}$$

$$P_0 = RTs_0 \quad P_1 = RTs_1$$

P0 : Pressure of the source region

P1 : Pressure of the sink region

R : Gas constant

T : Temperature

s₀ : Source solute concentration

s₁ : Sink solute concentration

x : Tube radius

L : Tube length

η : Liquid viscosity

> Introduction

How to build a model ?

Model construction from A to Z is composed of multiple steps:

Example of phloem transport from a source to a sink within a plant (Minchin, 1993, JExpBot)

1. Model conception

The flow between **source** and **sink** is generated by the difference in solute concentration between the two entities and depends on the dimensions of vascular network

2. Mathematical formalisation

$$flow = \frac{(P_0 - P_1)\pi x^4}{8L\eta} \quad P_0 = RTs_0 \quad P_1 = RTs_1$$

3. Sensitivity analysis

Values of parameters of interest are changing within well-chosen ranges of variation to estimate the impact of each parameter on the output variables

concentration

s_0 s_1

physical properties

x η L

P0 : Pressure of the source region

P1 : Pressure of the sink region

R : Gas constant

T : Temperature

s_0 : Source solute concentration

s_1 : Sink solute concentration

x : Tube radius

L : Tube length

η : Liquid viscosity

> Introduction

How to build a model ?

Model construction from A to Z is composed of multiple steps:

Example of phloem transport from a source to a sink within a plant (Minchin, 1993, JExpBot)

1. Model conception

The flow between **source** and **sink** is generated by the difference in solute concentration between the two entities and depends on the dimensions of vascular network

2. Mathematical formalisation

$$flow = \frac{(P_0 - P_1)\pi x^4}{8L\eta} \quad P_0 = RTs_0 \quad P_1 = RTs_1$$

3. Sensitivity analysis

Values of parameters of interest are changing within well-chosen ranges of variation to estimate the impact of each parameter on the output variables

4. Model calibration

We estimate the values of the most important parameters (with the help of experimental data sets)

5. Model evaluation

Verification of formalisms and parameter values on an independent experimental data set

6. In-silico explorations

Now we play with the model, we explore scenarios that do not exist in real life.
We free ourselves from the constraints of reality

> Introduction

Different kind of models

Ref : Addiscot and Wagenet, 1985, JSolSc

Deterministic models

A system or process is considered to operate in such a way that a given set of events or parameters gives exactly the same result

Mechanistic models

Models that are as close as possible to the mechanisms. We integrate the **fundamental mechanisms** as we understand them at the moment.

Hard to reach, we will rather quantify the **degree of mechanism**.

A relation is mechanistic if the processes of a scale are represented **by emergence** using functional relations known at the lower scale

Functional models

Processes are simplified or based on empirical observations to represent them.

Stochastic models

Integration of **randomness** in the representation of certain processes. We consider that the output is uncertain. The model is structured to take into account and characterize this **uncertainty**

Mechanistic models

Same as above

Non-mechanistic models

Use of transfer functions useful for industrial processes.

Statistical models

> Introduction

Triangle of plant modeling

Ref : Kurth, 1994 ; Bornhofen 2009

Simple

Aggregated models

Models at populational scale (eg. Forest).
Relations between compartments are statistic or empirical
(and not mechanistic or fonctionnal)

The priority is the **plant structure** and the **accurate geometrical and botanic representation** of plants
The botanic representation is achieved through the use of plant **modularity** (several and independants modules that follow their own production rules).
Biological processes are not accurately represented

Structural realism

Functional realism

Crops are considered **globally**.
The priority is the representation of **biological processes**.
Plant structure or organ morphology are not considered
Many crop models (where processes are more or less represented mechanistically) are in this corner.

Complex

FSPM

Structural models

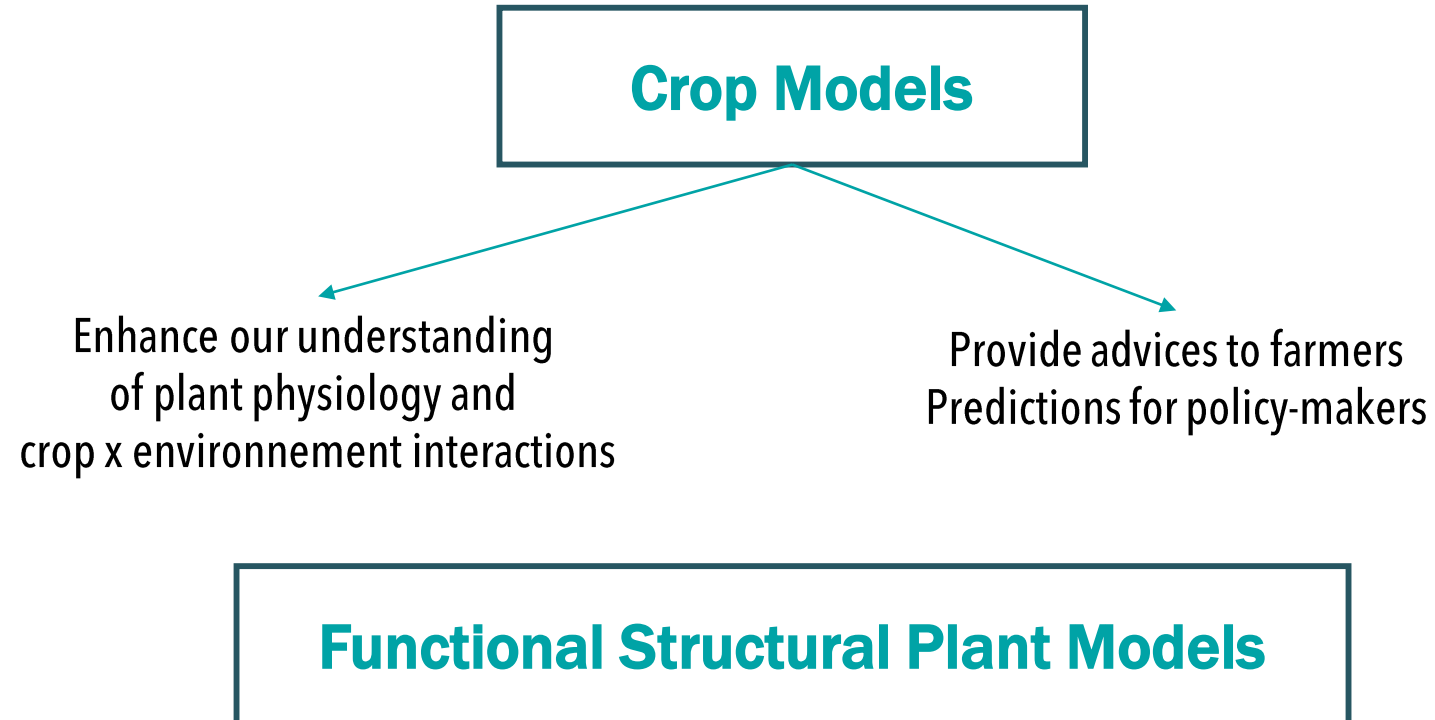
Process-based models

Compromise between the **geometric structure** and the representation of **biological processes**.
Effective tool to explore questions that need plant **3D representation** (competition for light etc.)
Huge computation time

> Introduction

Plant models in agronomy

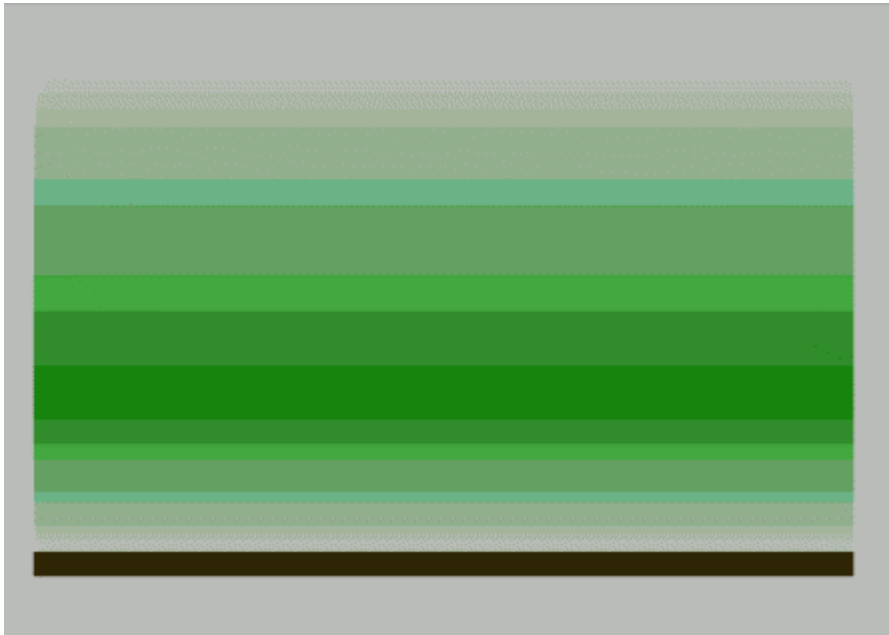
Ref : Passioura, 1996



> Introduction

Plant models in agronomy

Crop Models



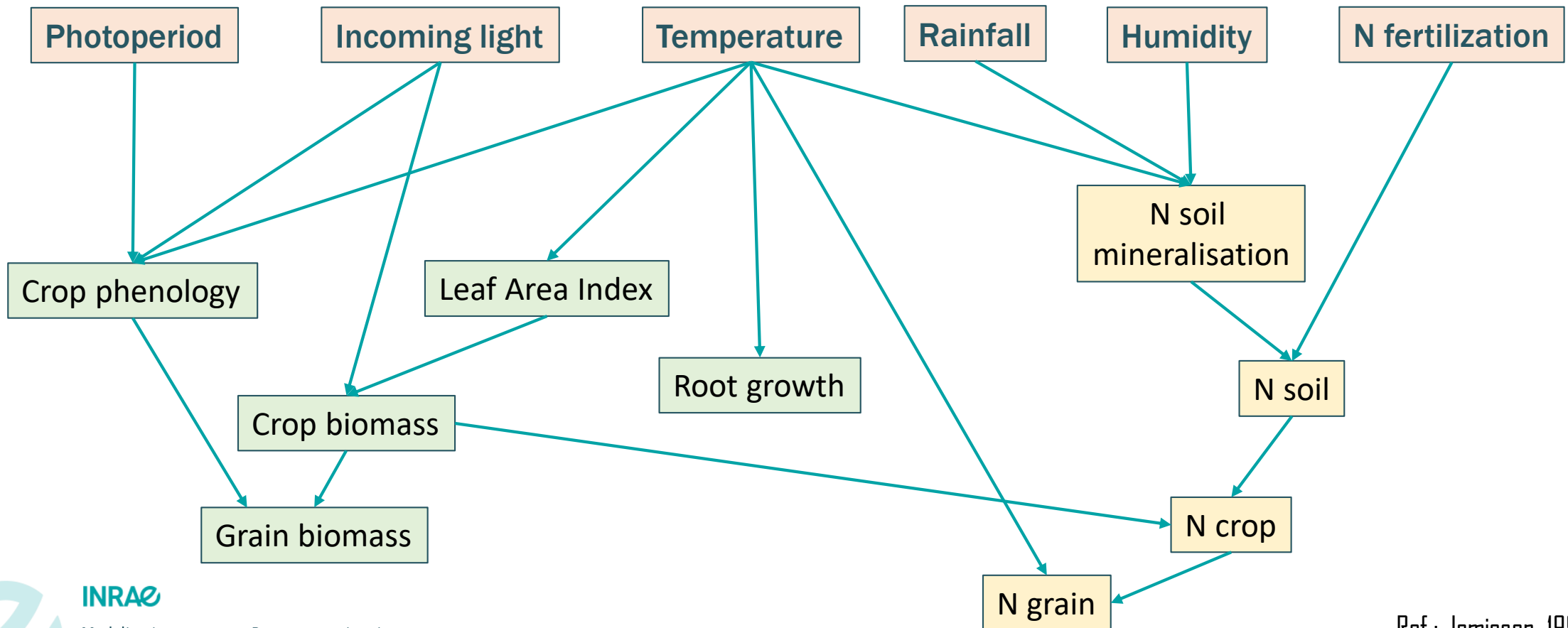
- No representation of plant structure. We consider an average plant that represent the entire crop with a random distribution of organs.
- Turbid medium : we consider only blurred layer of plant organs
- Effort for the representation of biological (and ecophysiological) processes

> Introduction

Plant models in agronomy

Example of Sirius

Crop model : estimation of **crop biomass**, **grain biomass** and **nitrogen content** from meteorological data



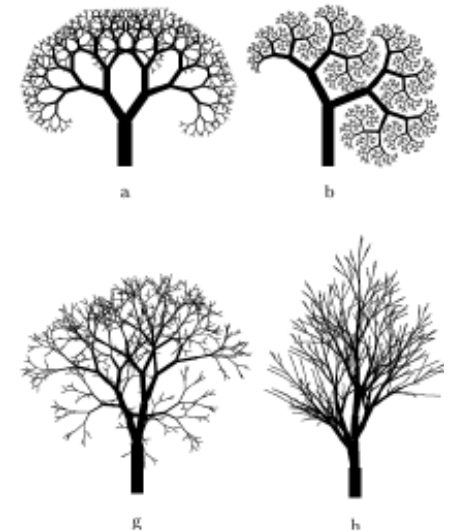
> Introduction

Plant models in agronomy

Functional Structural Plant Models

- Plant and organ morphogenesis follow a modular approach (usually L-systems).
- Organs are defined by type (meristem, leaf, internode)
- Local development rules are applied according to organ type and independantly for each organ
- Graphical 3D representation of all organs
- All informations are stored in an output string
Lstring

Figure	r_1	r_2	α_1	α_2	φ_1	φ_2	w_0	q	e	min	n
a	.75	.77	35	-35	0	0	30	.50	.40	0.0	10
b	.65	.71	27	-68	0	0	20	.53	.50	1.7	12
c	.50	.85	25	-15	180	0	20	.45	.50	0.5	9
d	.60	.85	25	-15	180	180	20	.45	.50	0.0	10
e	.58	.83	30	15	0	180	20	.40	.50	1.0	11
f	.92	.37	0	60	180	0	2	.50	.00	0.5	15
g	.80	.80	30	-30	137	137	30	.50	.50	0.0	10
h	.95	.75	5	-30	-90	90	40	.60	.45	25.0	12
i	.55	.95	-5	30	137	137	5	.40	.00	5.0	12



Ref : Bonrnhofen, 2009



INRAE

Modeling in agronomy : Parameter estimation
February 2022 / Christophe Lecarpentier

> Introduction

Plant models in agronomy

Functional Structural Plant Models

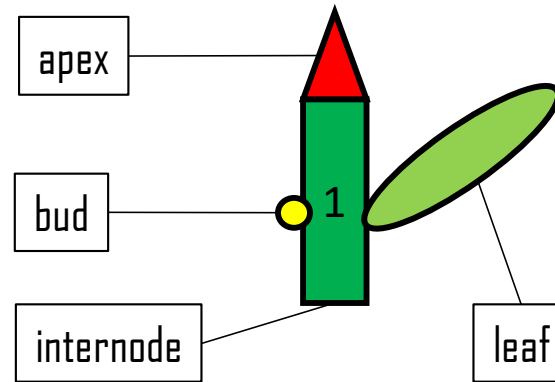
Axiom : Apex

Production rules :

Apex **produces** Leaf, Internode, Bud

Bud **becomes** Apex

Leaf **elongates** itself (2 steps)



> Introduction

Plant models in agronomy



Step 1

Functional Structural Plant Models

Axiom : Apex

Production rules :

Apex **produces** Leaf, Internode, Bud

Bud **becomes** Apex

Leaf **elongates** itself (2 steps)

Step 1 : **Apex**



INRAE

Modeling in agronomy : Parameter estimation
February 2022 / Christophe Lecarpentier

> Introduction

Plant models in agronomy

Functional Structural Plant Models

Axiom : Apex

Production rules :

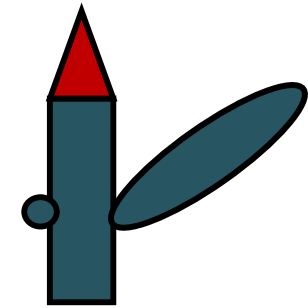
Apex **produces** Leaf, Internode, Bud

Bud **becomes** Apex

Leaf **elongates** itself (2 steps)

Step 1 : **Apex**

Step 2 : **Apex** Leaf Internode [Bud]



Step 1

Step 2

> Introduction

Plant models in agronomy

Functional Structural Plant Models

Axiom : Apex

Production rules :

Apex **produces** Leaf, Internode, Bud

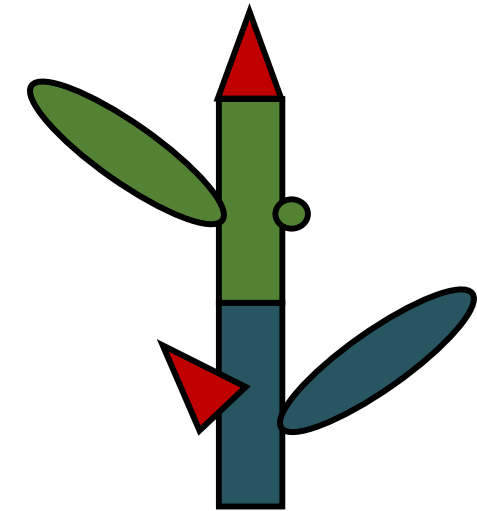
Bud **becomes** Apex

Leaf **elongates** itself (2 steps)

Step 1 : **Apex**

Step 2 : **Apex** Leaf Internode [Bud]

Step 3 : **Apex** Leaf Internode [Bud] Leaf Internode [Apex]



Step 1

Step 2

Step 3

> Introduction

Plant models in agronomy

Functional Structural Plant Models

Axiom : Apex

Production rules :

Apex **produces** Leaf, Internode, Bud

Bud **becomes** Apex

Leaf **elongates** itself (2 steps) (2 steps)

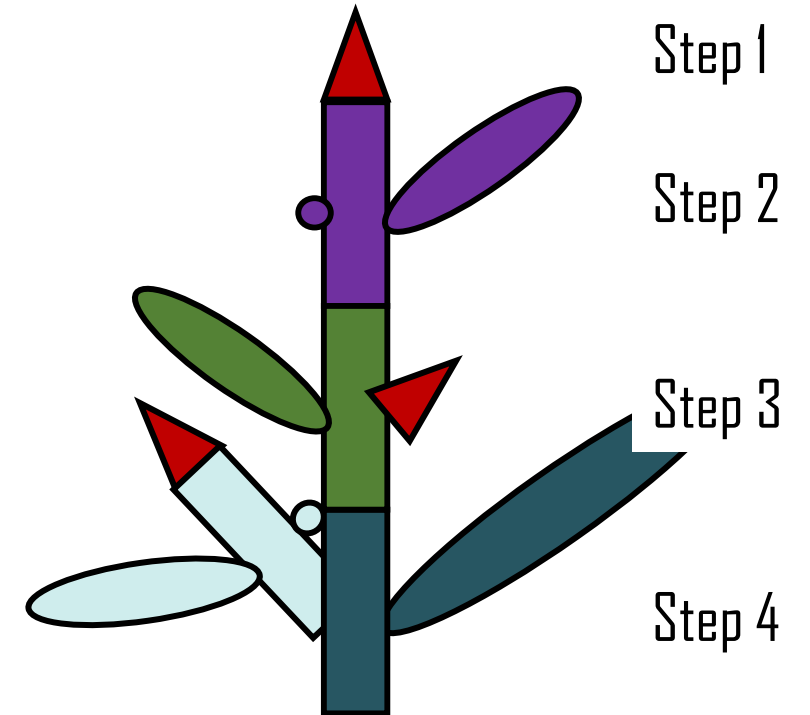
Step 1 : **Apex**

Step 2 : **Apex** Leaf Internode [Bud]

Step 3 : **Apex** Leaf Internode [Bud] Leaf Internode [Apex]

Step 4 : **Apex** Leaf Internode [Bud] Leaf Internode [Apex] Leaf Internode [Apex Leaf Internode [Bud]]

L string



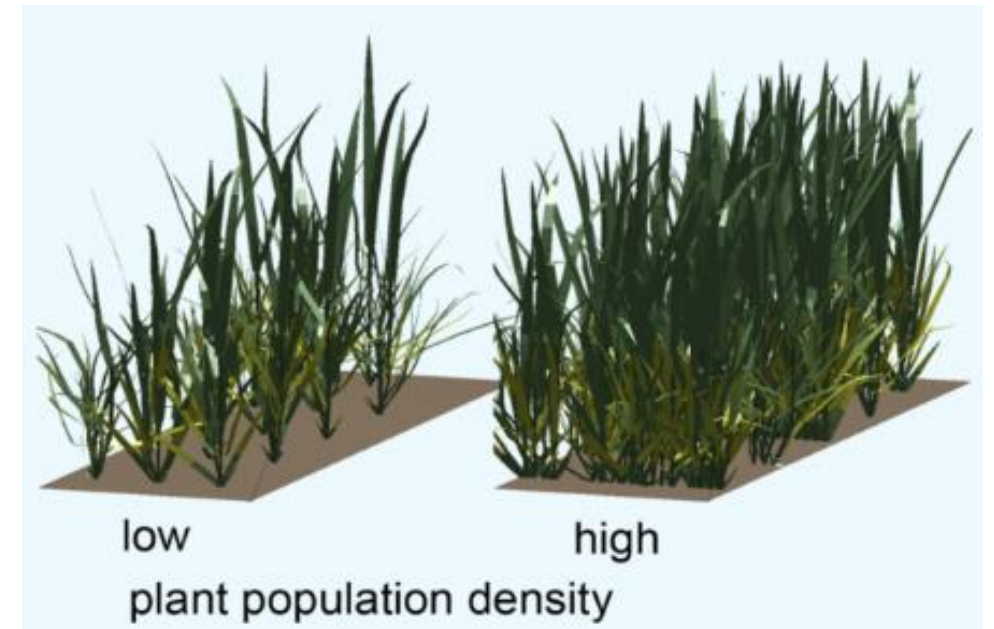
> Introduction

Plant models in agronomy

Functional Structural Plant Models

- ADEL-Wheat is considered as a wheat « drawer » as there is little rules of biological regulations
- More structural model than a fonctionnal (or FSPM)
- Many parameters to calibrate, through a huge experimental effort
- Useful model to represent plants to perform more precise measurements and measurements that are not possible experimentally
- Graphical 4D representation of all organs (all the crop cycle is simulated)

ADEL-Wheat



Ref : Fournier et al. 2003 ; Abichou et al. 2018



INRAE

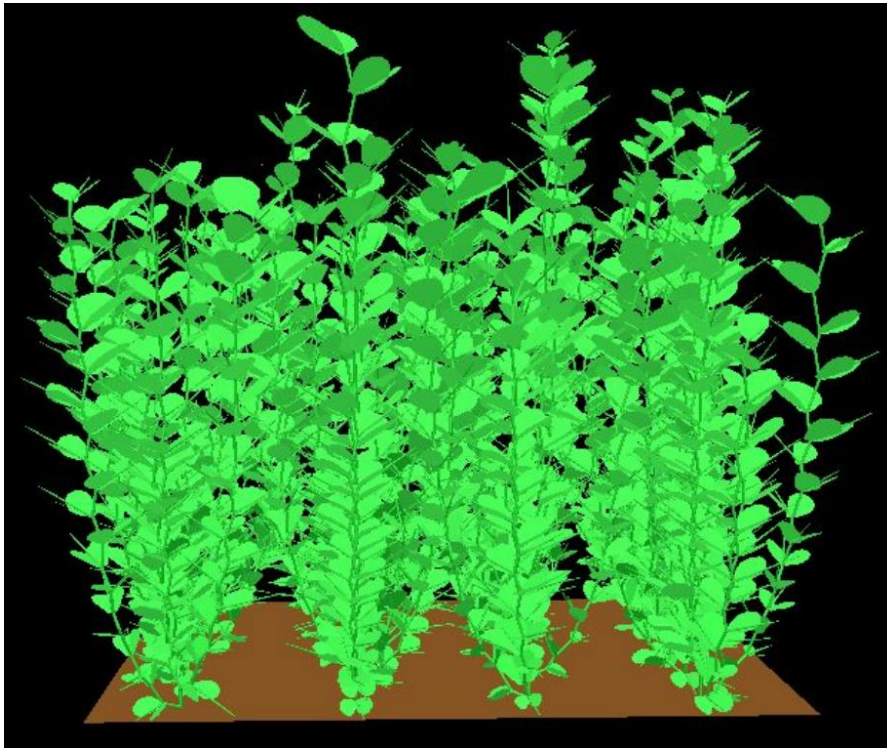
Modeling in agronomy : Parameter estimation
February 2022 / Christophe Lecarpentier

> Introduction

Plant models in agronomy

Functional Structural Plant Models

L-Pea



- Pea morphogenesis rules based on experimental datas
- L-Pea was created to quantify light partitionning among wheat / pea mixtures
- Graphical 4D representation of all organs
- Several models were coupled
 - L-Pea
 - ADEL Wheat (3D Wheat)
 - CARIBU : Light interception

Ref : Barillot, AoB, 2014

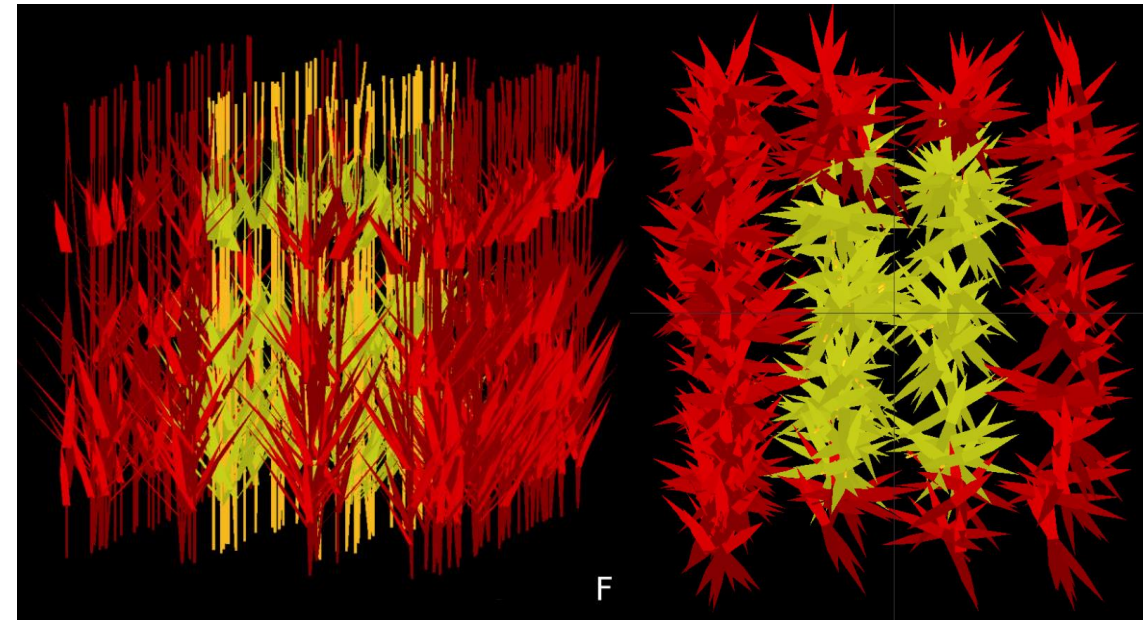
> Introduction

Plant models in agronomy

Functional Structural Plant Models

- WALTer morphogenesis is based on ADEL-Wheat and modelled using L-system
- Representation of tillering processus through simple rules
 - Emergence, and tillering death are simulated
- WALTer was developped to represent competition for light between contrasting varieties within wheat associations

WALTer



Ref : Lecarpentier et al. 2019 ; Blanc et al, 2021



INRAE

Modeling in agronomy : Parameter estimation
February 2022 / Christophe Lecarpentier

What is a parameter ?



> Introduction

What is a parameter ?

Variable

Mathematical term that is **changing** during the simulation.

State variables are variables that are calculated by the model (eg. Biomass, LAI)

Input variables are variable that are estimated experimentally

Parameter

Mathematical term that keeps its value during the whole simulation. However, its value can change between simulations (eg. phyllochron, RUE)

Constant

Mathematical term that will never change either during the simulation nor between simulations (eg. Physical constants such as constant gaz , phyllochron, RUE)



> Introduction

What is a parameter ? Examples

Variable

$$ER_n^B = \frac{L_n^B - L_{(n-1)}^S}{ED^B \times Phl}$$

Parameter

- ER_n^B : elongation rate of the blade n ($cm.degree day^{-1}$)
- L_n^B : length of the blade n (cm)
- $L_{(n-1)}^S$: length of the sheath n-1 (cm)
- ED^B : elongation duration of blades (phyll time)
- Phl : phyllochron (degree days)

> Introduction

What is a parameter ? Examples

$$T_f = \boxed{\rho} \times \left(\frac{\boxed{M_w} \times \boxed{p^*}}{\boxed{R} \times \boxed{T}} \right) \times \boxed{A_f} \times (\boxed{H_f} - \boxed{H_a})$$

Parameter

Variable

Constant

$$p^* = 0,0088048 \times e^{0,0547 (T - 273,15)}$$

- T_f : Fruit transpiration ($g.h^{-1}$)
- ρ : Permeation coefficient of fruit surface to water vapour ($cm.h^{-1}$)
- M_w : Volum mass of water ($kg.L^{-1}$)
- R : gas constant ($cm^3.bar.mol^{-1}.K^{-1}$)
- T : Temperature ($^{\circ}C$)
- A_f : Fruit surface (cm^2)
- H_f : relative humidity of air space in fruit
- H_a : ambient humidity



> Introduction

What is a parameter ?

- Parameters can be estimated through multiple ways:
 - Scientific litterature
 - Experimentation
 - Coworkers expertise
- Parameters gather many informations, some of them are considered as **genotypic** or **specific**. For example, parameters governing plant architecture (leaf angle, internode length, etc.).
- In crop modelling, the difference between **genotypic parameters** and **parameters that are constant** among all cultivar of a crop is widely used.



➤ Introduction

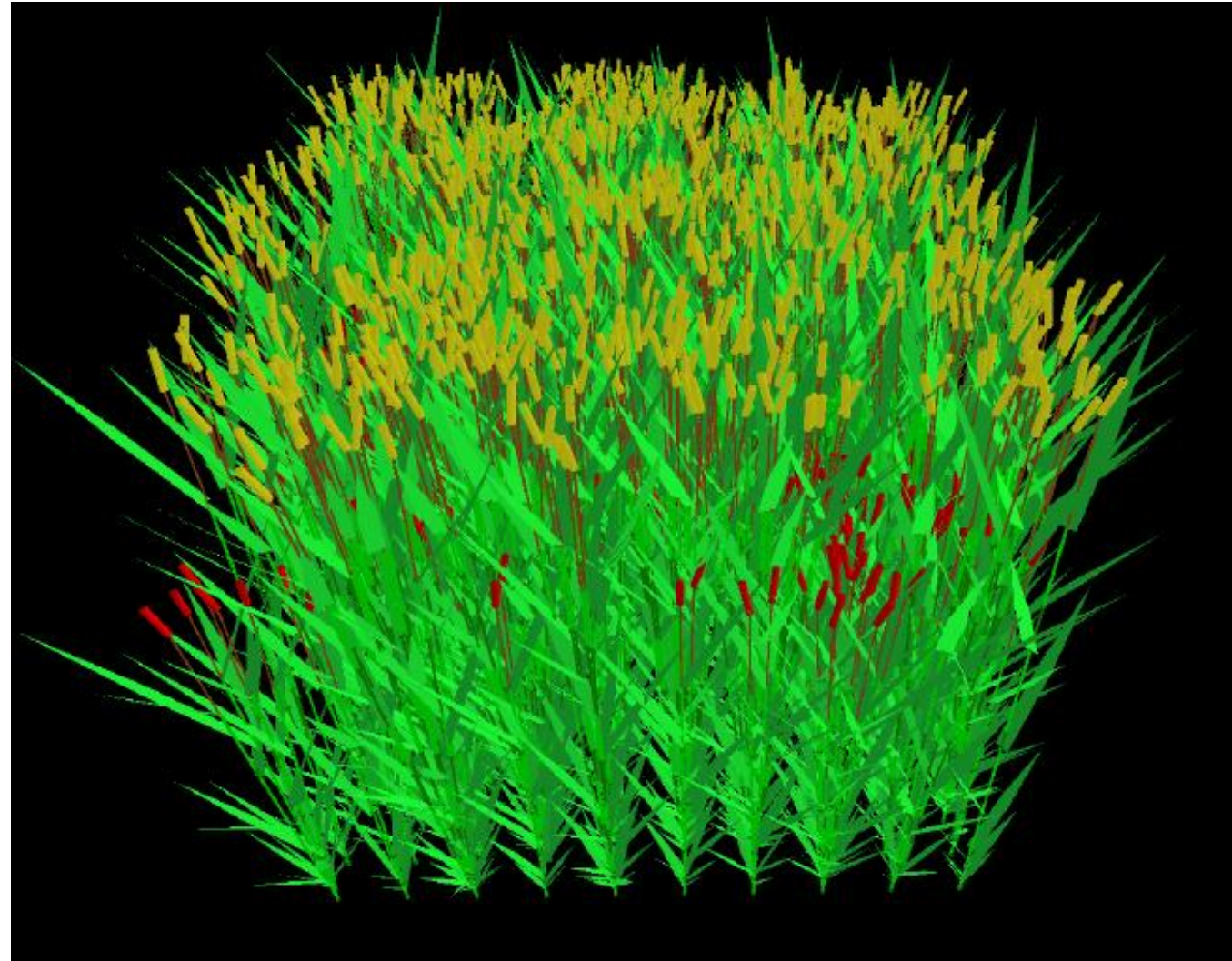
What is a parameter ?

WALTer model

$$L_n^I = (a \times n + b)^2$$

$$H = \sum_{n=1}^{n_{max}} L_n^I$$

- Modifications of the values of a and b parameters
- Creation of two populations of plants
 - Small plants (red ears)
 - Tall plants (yellow ears)



> Introduction

What is a parameter ?

ArchiSimple model

- **Dmin**: Minimal diameter
- **Dmax**: Maximal diameter
- **DIDm**: Hierarchy slope
- **IBD**: Inter-Branch Distance
- **VarD**: Variation of diameter
- **ELT**: Elongation rate per diameter
- **Tgravi** : gravitropism of root system
- **Igravi** : gravitropism of root system
- **DeIBEI** : Delay Before root Elongation

Defining contrasting architecture just using **different values** of parameters sets

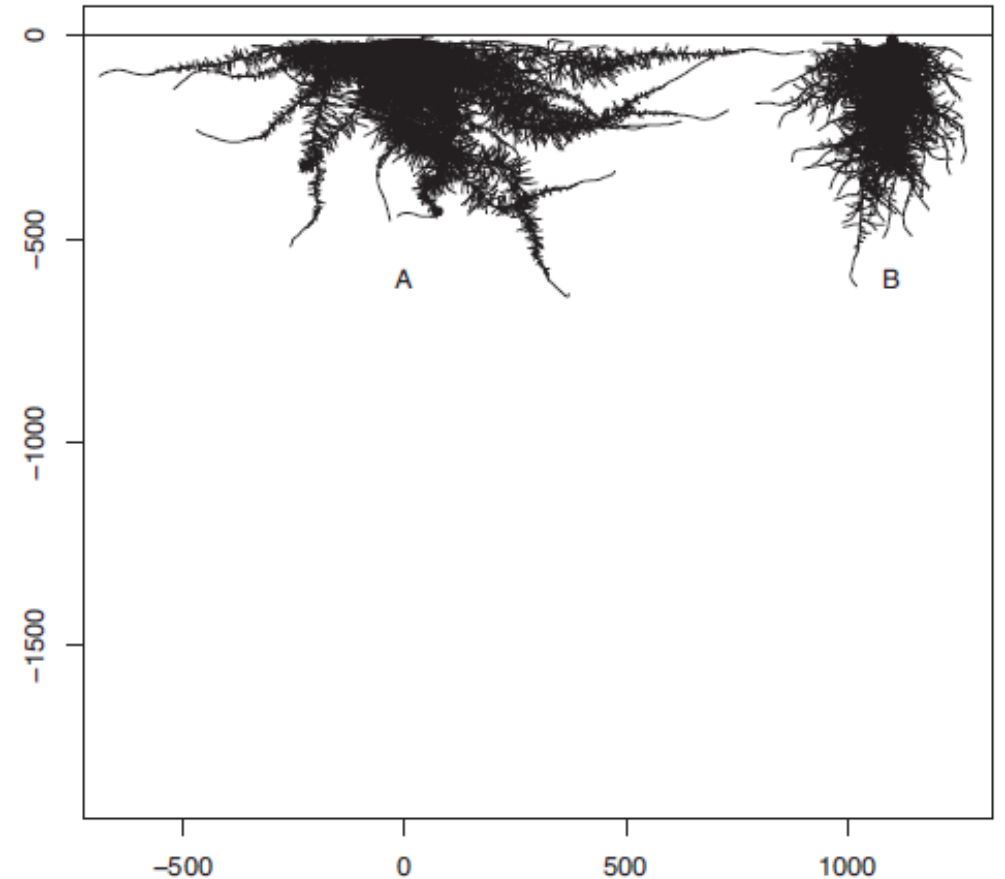


Fig. 5. Two different types of root systems simulated at day 50, illustrating *Musa* (A) and *Pisum sativum* (B). Depth is expressed in mm.

> Introduction

What is a parameter ?

- The influence of parameter values is **huge** and is generally quantified by a **sensitivity analysis**
- It is important to put effort on the estimations of **accurate values** for each parameters
 - Important to use a model for predictive simulations or for simulations that aim at enhancing our understanding of crop systems

Accurate equations and concepts are not enough to use properly a model. We also need **accurate values of parameters**



What is model calibration ?

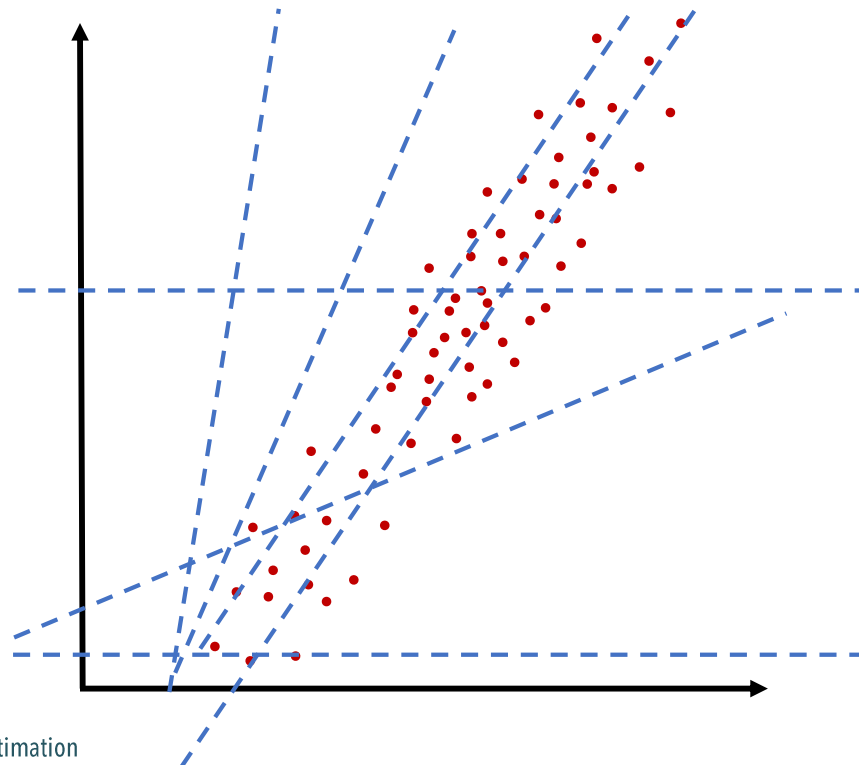


> Introduction

What is model calibration ?

« Model calibration is to adjust influential model parameters [...] within their reasonable ranges so that the modeling results [...] are comparable with observed data. »

Wang, Kemanian and Williams (2011)



Parameters values will vary according to a precise scheme (more or less complex)

> Introduction

What is model calibration ?

- Parameters values will vary according to a precise scheme (more or less complex)
- The aim is to optimise the predictions made by our model (the closest as possible to observed datas)
- Usually the calibration is taking place on integrative outputs of the model (if we have experimental data to confront our simulations with)
 - **Integrative output variable** : results of several processes that takes place in the model and actions of the environments



Why do we need to calibrate a model ?



> Introduction

Why do we need to calibrate a model ?

- Values of most parameters in models of ecology or environmental science are not well known in bibliography, so we need to estimate them numerically
- However some parameters are very specific and have no biological reality, so we need to estimate them numerically
- Models are simplifications of « **real life** ». Due to these approximations, parameter values may be different **real experimental values**.
- Genotypic or specific values represent average values of a whole group and not an individual value. parameter values may be different **real experimental values**



How do we calibrate a model ?



> Introduction

How to calibrate ?

Estimate the parameters of a
simple function



Estimate some parameters of a
systemic model

Run the entire model on a precise simulation scheme





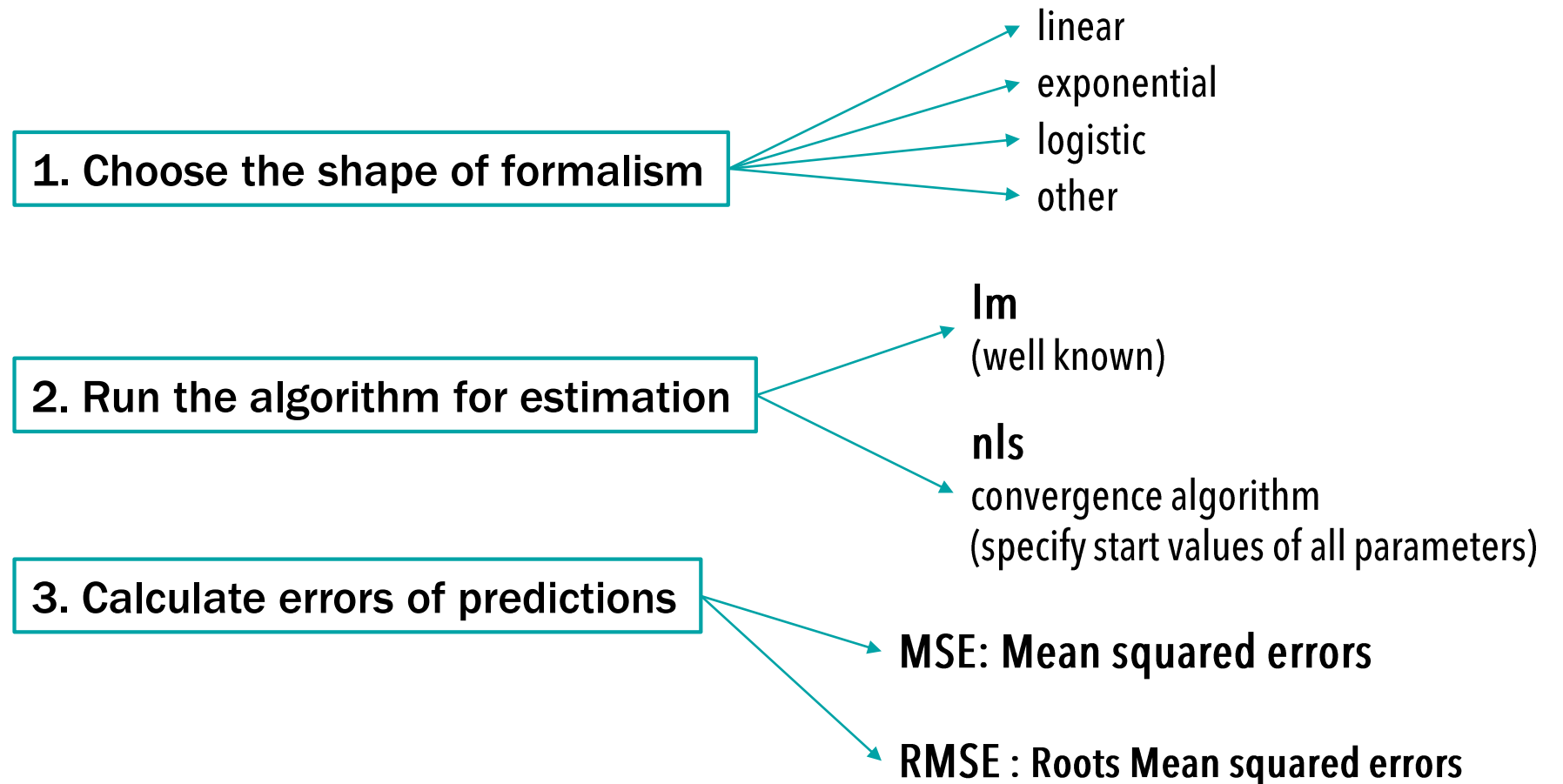
➤ Estimate the parameters of a simple function

Calibration using Ordinary Least Square method

➤ Estimate the parameters of a simple function

Steps for parameters estimation (using R)

Estimate parameters that minimize as much as possible the error of prediction.



➤ Estimate the parameters of a simple function

Estimation by Ordinary Least Square method

Errors of predictions are usually estimated as sum of squared errors

$$\text{Errors} = \sum_{i=1}^n [y_i - f(x_i; \theta)]^2$$

Observed (points to y_i)

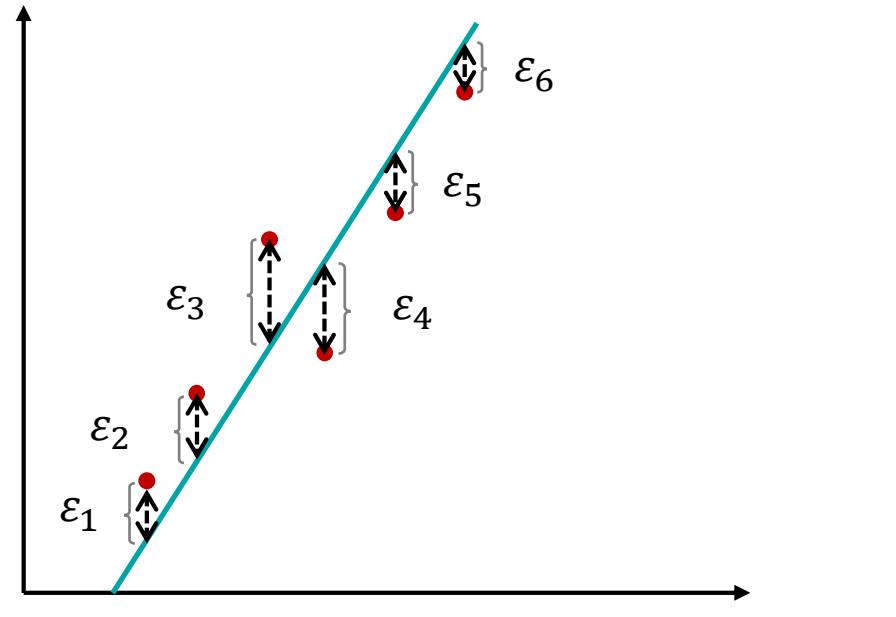
Predicted (points to $f(x_i; \theta)$)

$$MSE = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}$$

residuals (points to $Y_i - \hat{Y}_i$)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}}$$

residuals (points to $Y_i - \hat{Y}_i$)

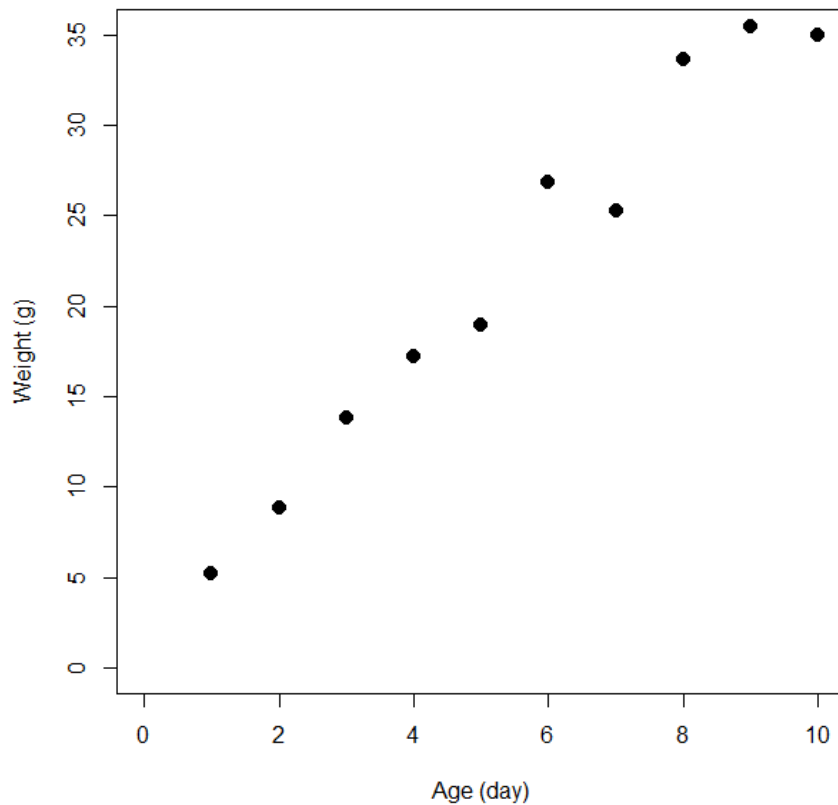


$$MSE = \frac{1}{6} \times \sum_{i=1}^6 [\epsilon_i]^2 \quad RMSE = \sqrt{\frac{1}{6} \times \sum_{i=1}^6 [\epsilon_i]^2}$$

➤ Estimate the parameters of a simple function

Estimation by Ordinary Least Square method

Example 1:



1. Choose the shape of your function → linear
2. Formalise the relation through an equation

$$\text{Weight} = a \times \text{Age} + b$$

3. Run R algorithm dedicated to linear functions

`lm(Y ~ X)`

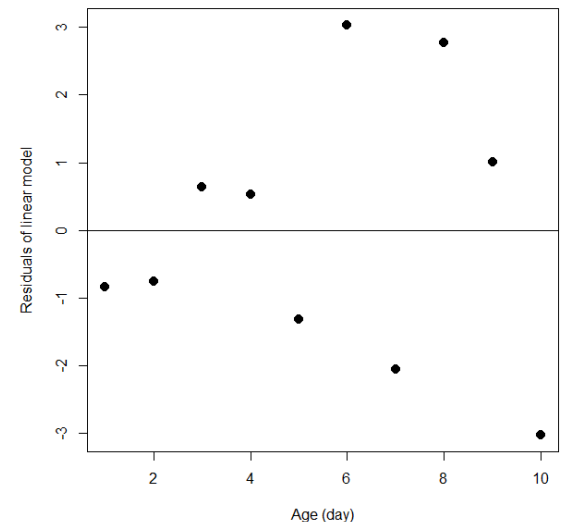
```
Call:
lm(formula = Y ~ X)

Coefficients:
(Intercept)      X 
      2.496      3.552
```

$a = 3.552$

$b = 2.496$

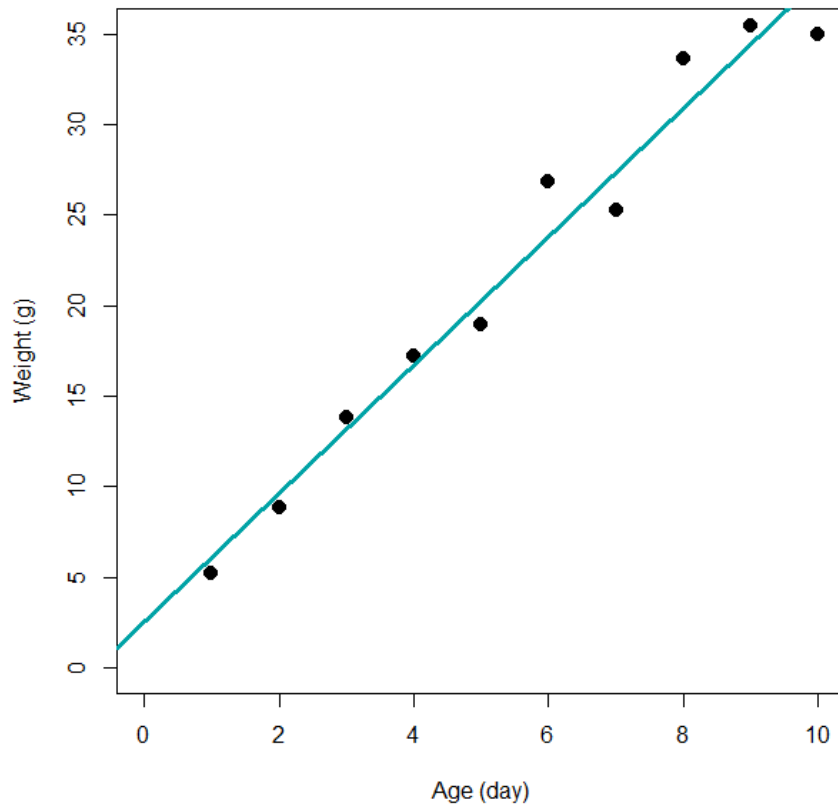
4. Draw the residuals versus time



➤ Estimate the parameters of a simple function

Estimation by Ordinary Least Square method

Example 1:



1. Choose the shape of your function → linear
2. Formalise the relation through an equation

$$\text{Weight} = a \times \text{Age} + b$$

3. Run R algorithm dedicated to linear functions

`lm(Y ~ X)`

```
Call:
lm(formula = Y ~ X)

Coefficients:
(Intercept)      X 
      2.496      3.552
```

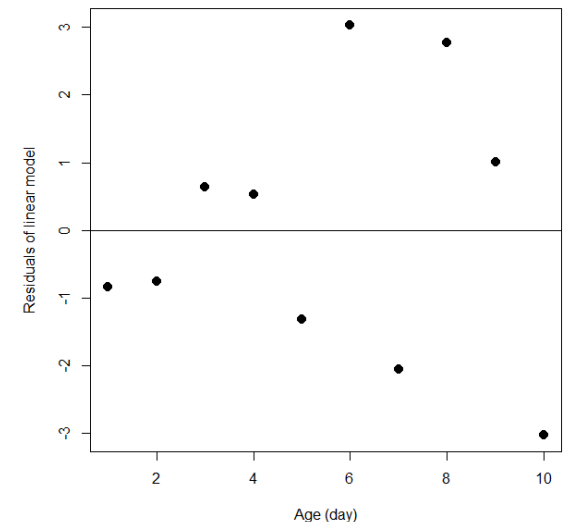
$a = 3.552$

$b = 2.496$

4. Draw the residuals versus time
5. Draw the predicted model using estimated parameters

$MSE = 3.48$

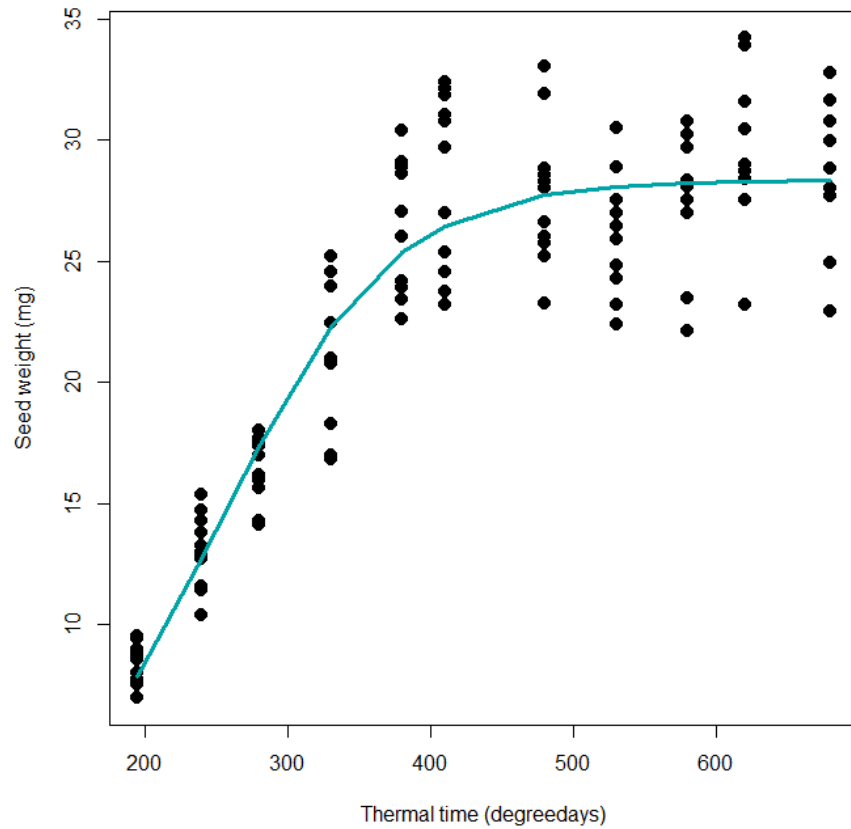
$RMSE = 1.87 \text{ g}$



➤ Estimate the parameters of a simple function

Estimation by Ordinary Least Square method

Example 2:



1. Choose the shape of your function → **logistic**
2. Formalise the relation through an equation

$$\text{Weight} = \frac{W}{1 + e^{B - C \times \text{ThermalTime}}}$$

3. Run R algorithm dedicated to linear functions

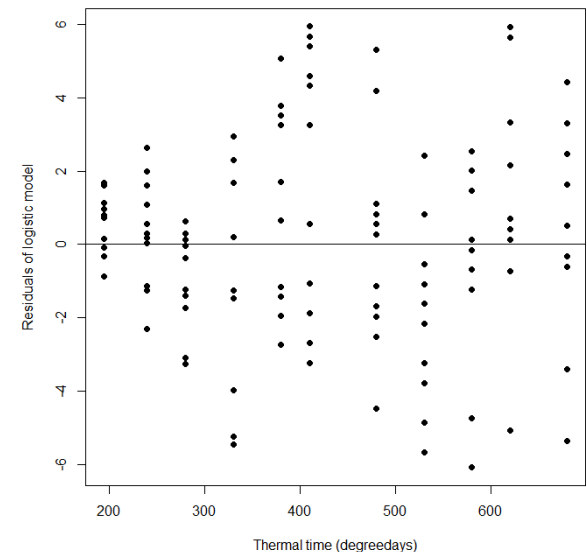
`nls(Weight ~ W/(1+exp(B-C*ThermalTime)), start = list(W=30, B = 5, C = 0.02))`

```
Nonlinear regression model
model: weight ~ W/(1 + exp(B - C * ThermalTime))
data: seed
      W      B      C
28.35090 4.22364 0.01673
residual sum-of-squares: 957.9
```

4. Draw the residuals versus time
5. Draw the predicted model using estimated parameters

MSE = 95.78

RMSE = 9.78 mg



INRAE

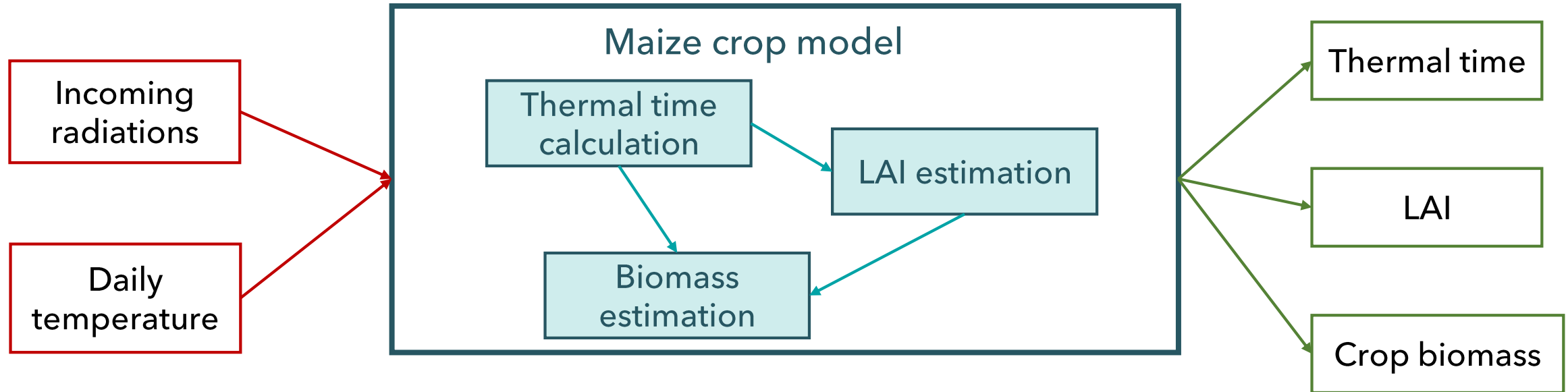
➤ **Estimate the parameters of a complex model**

Calibration using Ordinary Least Square method

➤ Estimate the parameters of a complex model

Presentation of the maize crop model

Dynamic crop model of maize that simulated the development of a maize crop



$$TT_{day+1} = TT_{day} + \Delta TT_{day}$$

$$B_{day+1} = B_{day} + \Delta B_{day}$$

$$LAI_{day+1} = LAI_{day} + \Delta LAI_{day}$$

Integration of many key concepts that are used in many crop models

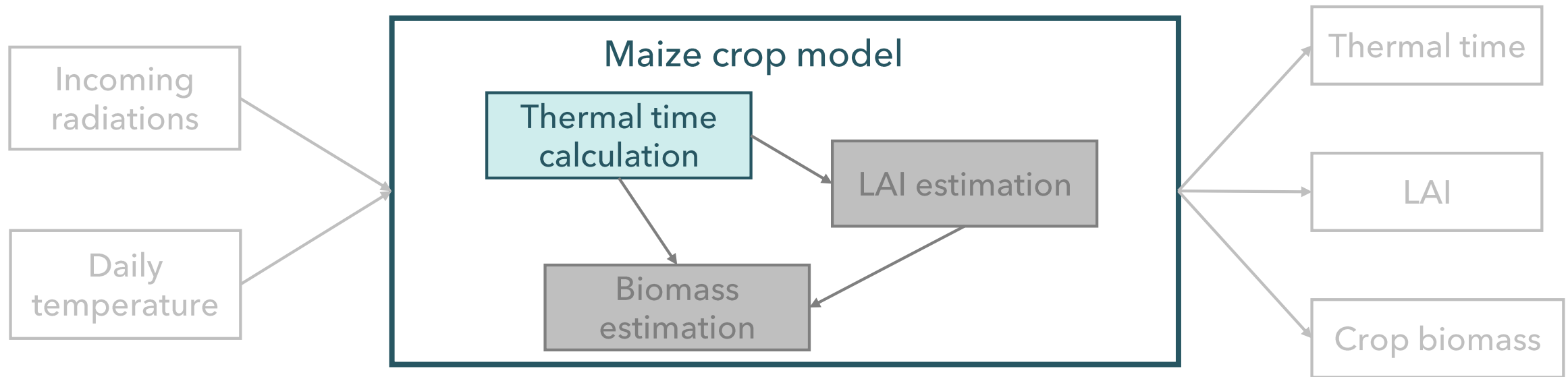
- Daily time step
- Calculation of a increment each day (thermal time, biomass and LAI)
- Crop growth is driven by thermal time



➤ Estimate the parameters of a complex model

Presentation of the maize crop model

Dynamic crop model of maize that simulated the development of a maize crop



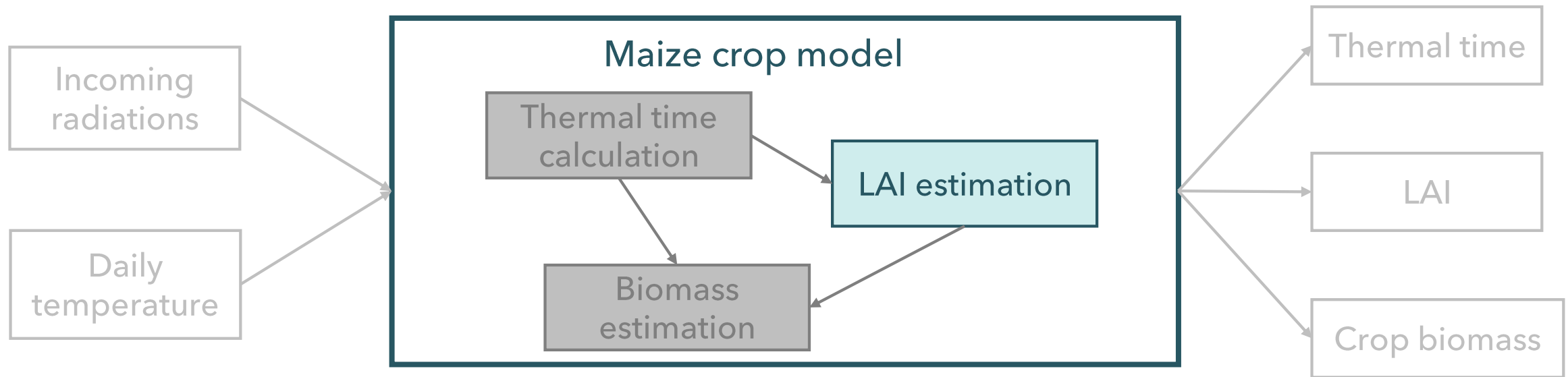
$$\Delta TT_{day} = \max \left(\frac{T_{day}^{min} + T_{day}^{max}}{2} - T_{base}; 0 \right)$$

Thermal time increment is the difference between daily mean temperature and basis temperature

> Estimate the parameters of a complex model

Presentation of the maize crop model

Dynamic crop model of maize that simulated the development of a maize crop



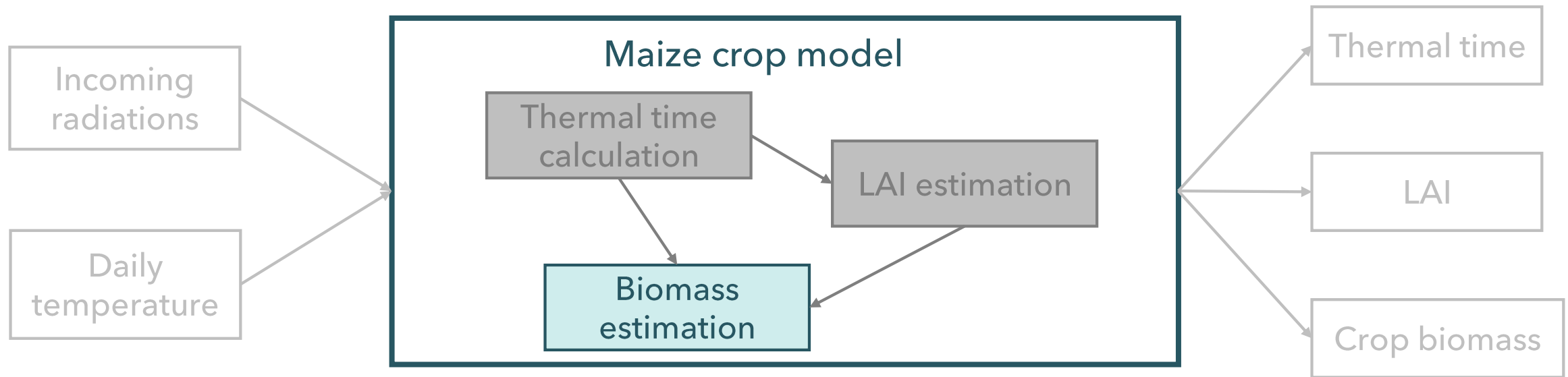
$$\Delta LAI_{day} = \alpha \times \Delta TT_{day} \times LAI_{day} \times \max(LAI_{max} - LAI_{day}; 0) \quad \text{if } TT_{day} \leq TTL$$
$$\Delta LAI_{day} = 0 \quad \text{if } TT_{day} > TTL$$

- LAI increment depends on:
 - Alpha
 - Thermal time
 - The difference between LAI max and current value of LAI
 - TTL

➤ Estimate the parameters of a complex model

Presentation of the maize crop model

Dynamic crop model of maize that simulated the development of a maize crop

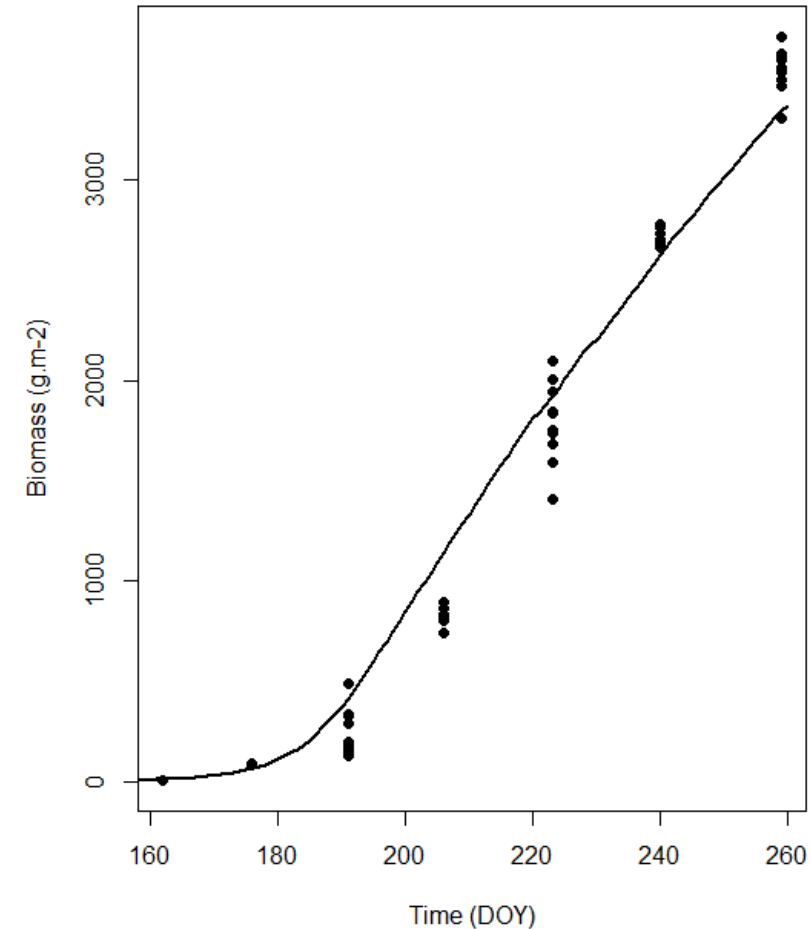
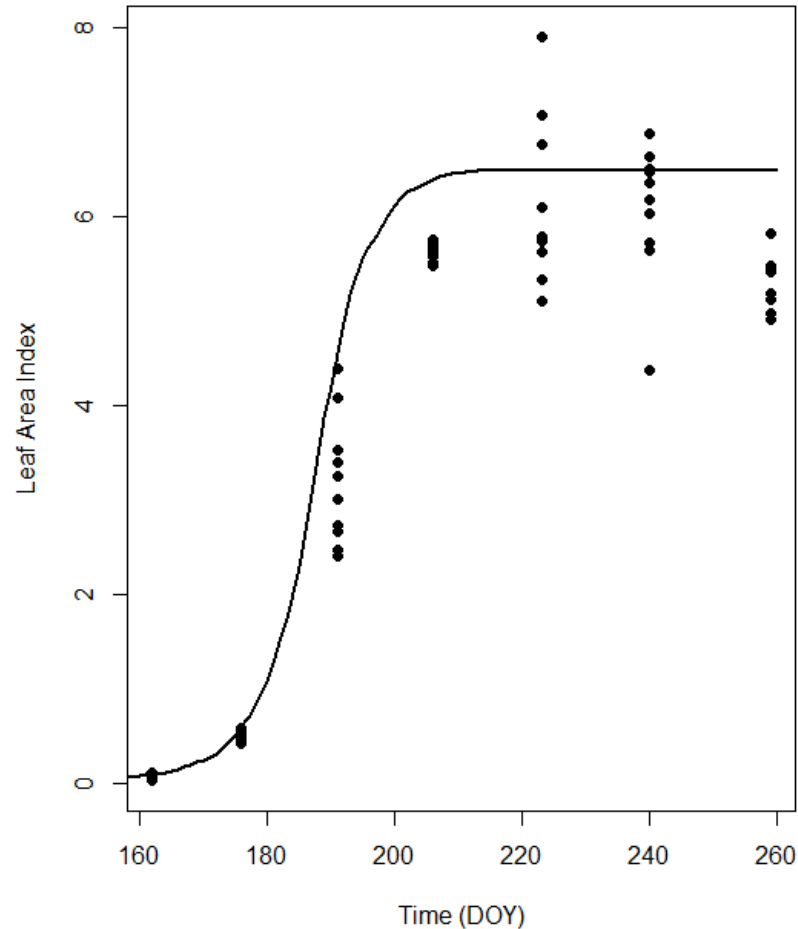


$$\Delta B_{day} = \boxed{RUE} \times \left(1 - e^{-\boxed{K} \times LAI_{day} \times I_{day}}\right) \text{ if } TT_{day} \leq \boxed{TTM}$$
$$\Delta B_{day} = 0 \text{ if } TT_{day} > TTM$$

- Biomass increment depends on:
 - RUE and K
 - Current value of LAI
 - Incoming radiations
 - TTM

➤ Estimate the parameters of a complex model

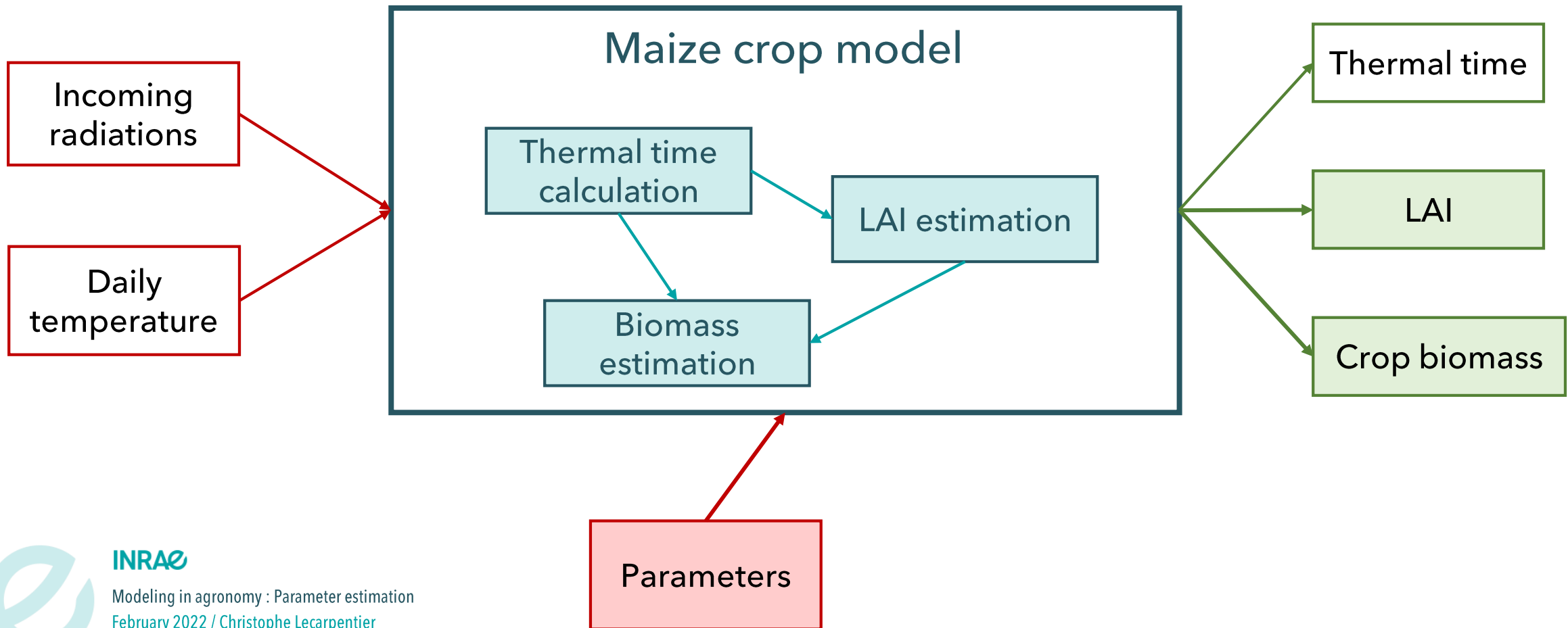
Presentation of the maize crop model



➤ Estimate the parameters of a complex model

Presentation of the maize crop model

Define values of parameters that produce predictions of **LAI** and **Biomass** close to **experimental ones**



➤ Estimate the parameters of a complex model

Steps to follow

1. Select **which** parameters to estimate
2. Design a **simulation scheme**
 1. Define range of variation of each parameters
 2. Define the shape of the scheme (regular grid, latin hyper cube)
3. **Run n** time the model
4. Calculate **RMSE** and/or **MSE**
5. Select the **best parameter set**



> Estimate the parameters of a complex model

1 - Select which parameters to estimate

No **sensitivity analysis**. We need to determine which parameters to calibrate according to our knowledge and expertise

- **Tbase** : Temperature below which the growth is stopeed ($^{\circ}\text{C}$)
- **LAI_{max}** : Maximum LAI that can be reached by the crop (*no dim*)
- **TTL** : End of the increase of LAI (*degreeday*)
- **TTM** : Maturity of the crop (*degreeday*)
- **RUE** : Radiation Use efficiency (conversion of quen quantity of interception radiations to crop biomass (g.MJ^{-1}))
- **K** : Extinction coefficient (*no dim*)
- **Alpha** : parameter the relative rate of leaf area index increase for small values of leaf area index (degreeday^{-1})



> Estimate the parameters of a complex model

1 - Select which parameters to estimate

No **sensitivity analysis**. We need to determine which parameters to calibrate according to our knowledge and expertise

- **Tbase** : Temperature below which the growth is stopeed ($^{\circ}\text{C}$)
- **LAI_{max}** : Maximum LAI that can be reached by the crop (*no dim*)
- **TTL** : End of the increase of LAI (*degreeday*)
- **TTM** : Maturity of the crop (*degreeday*)
- **RUE** : Radiation Use efficiency (conversion of quen quantity of interception radiations to crop biomass (g.MJ^{-1}))
- **K** : Extinction coefficient (*no dim*)
- **Alpha** : parameter the relative rate of leaf area index increase for small values of leaf area index (degreeday^{-1})



> Estimate the parameters of a complex model

2 – Design a relevant simulation scheme

Define range of variations of parameters

→ $RUE = A, \text{ range} = [A_{\min}; A_{\max}]$

- If we have **an idea of parameters** value (coworker expertise, scientific litterature, etc.)
 - If you have no idea of the impact of the parameters
 - Borders of range of variations : $[0.5 \times BaseValue ; 2 \times BaseValue]$ → **For the first try**
 - Borders of range of variations : $[0.75 \times BaseValue ; 1.5 \times BaseValue]$ → **If high impact**
 - Borders of range of variations : $[0.1 \times BaseValue ; 10 \times BaseValue]$ → **If low impact**
 - If you have an idea of the impact of the parameters
 - Lucky you
- If we have **no idea of parameters value**
 - Read scientific litterature or contact model developper
 - Run with contrasting values : **[0.01; 0.1; 1; 10; 100; 1000]**

Use sensitivity analysis

> Estimate the parameters of a complex model

2 – Design a relevant simulation scheme

Define the shape of the simulation scheme :

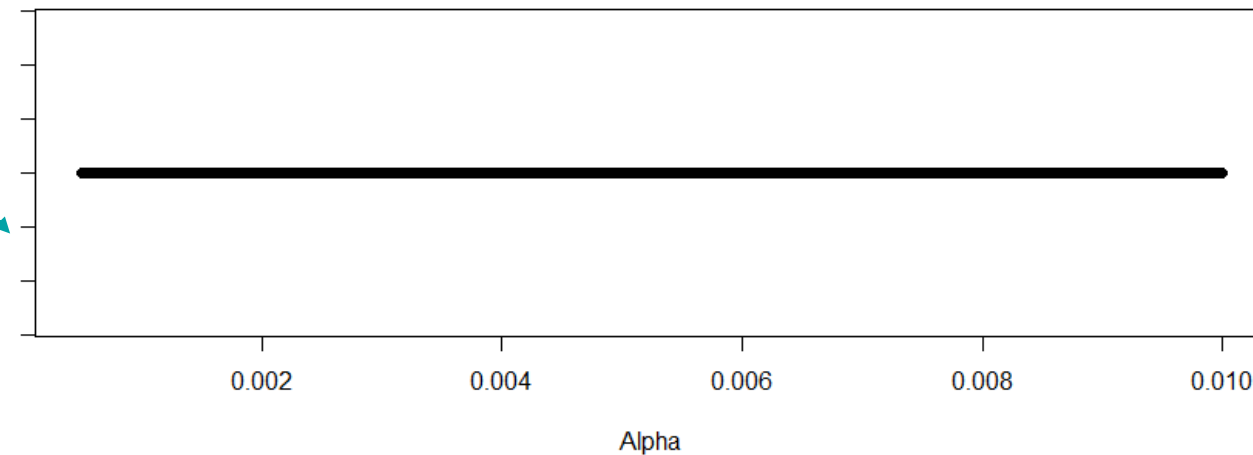
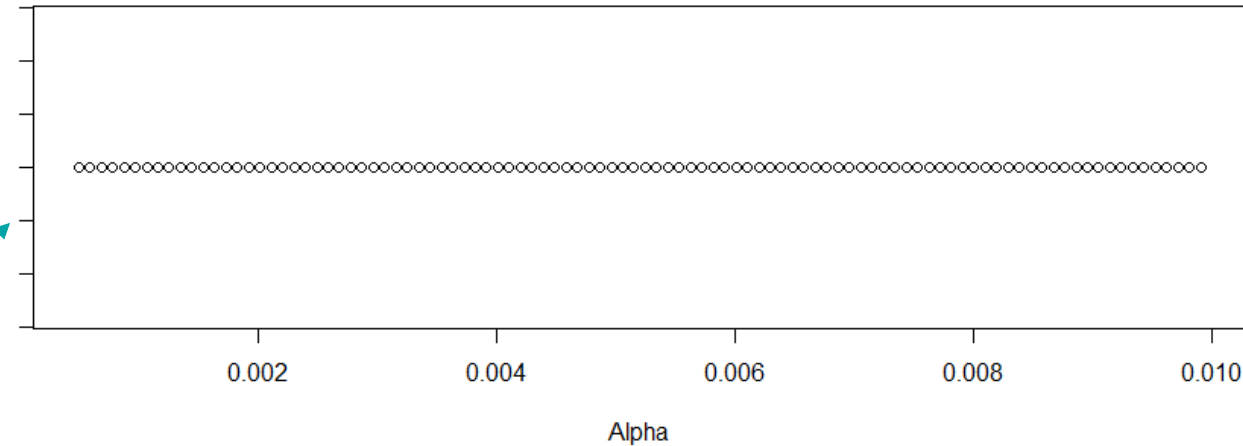
1 parameter to calibrate

- $\alpha : [0.005 : 0.01]$

Define
Number of simulations

100

1000



➤ Estimate the parameters of a complex model

2 – Design a relevant simulation scheme

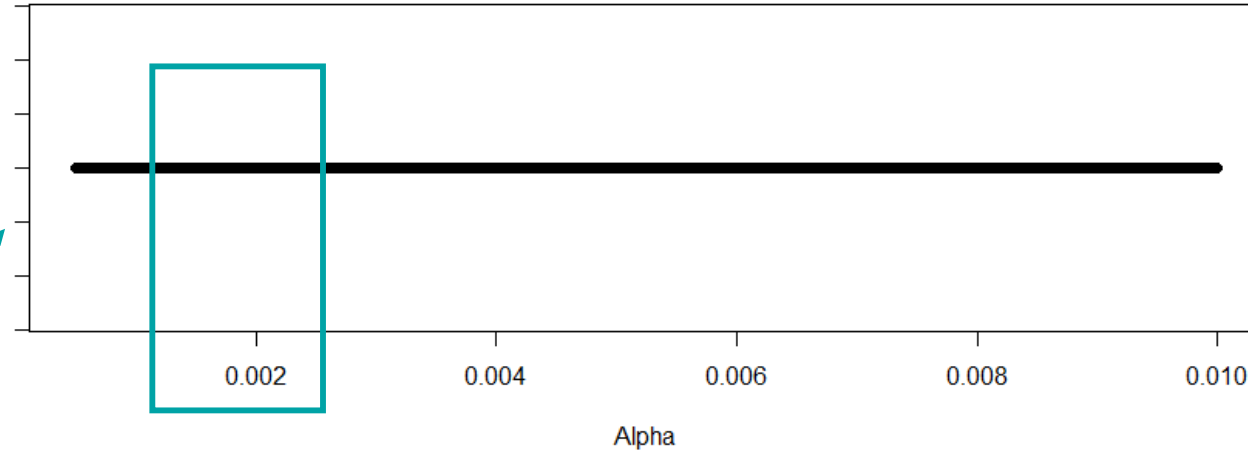
Define the shape of the simulation scheme :

1 parameter to calibrate

- $\alpha : [0.005 : 0.01]$

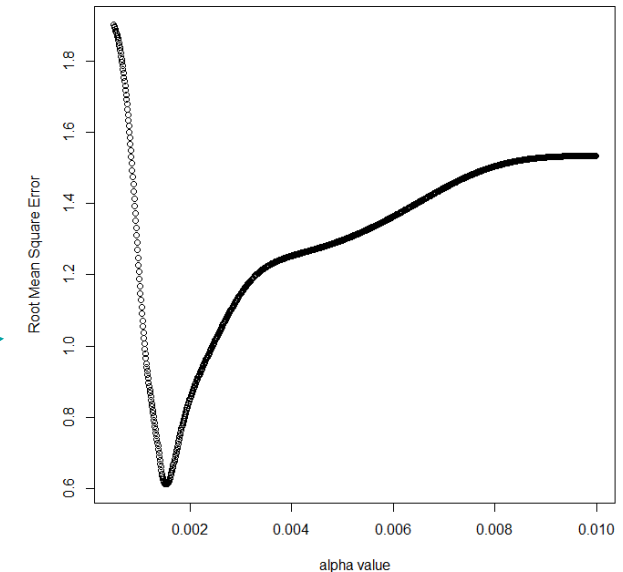
Define
Number of simulations

1000



First run

- Plot Error versus parameter values to identify **local** and **global** minima



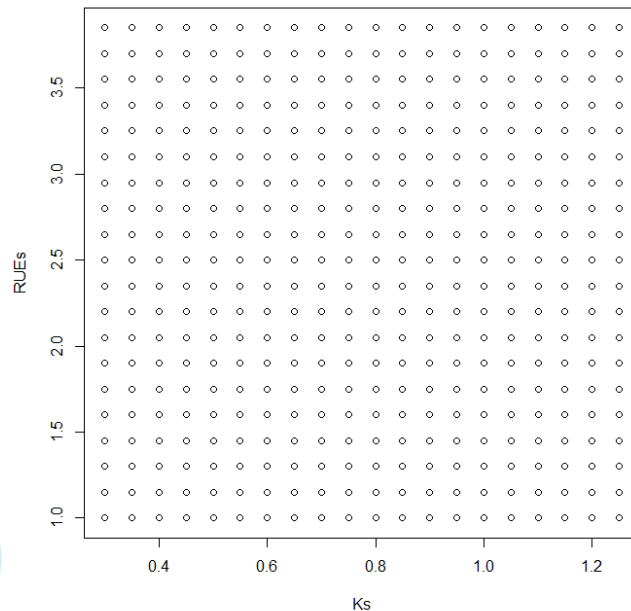
➤ Estimate the parameters of a complex model

2 – Design a relevant simulation scheme

Define the shape of the simulation scheme :

Regular grid

- 20 values for each parameter
- 20^p simulations for calibration (20, 400, 8000)
- Huge number of simulations for more than 2 parameters
- Redundancy of parameter values



8000 simulations
20 different values

2 parameters to calibrate

- RUE : [1 : 3.7]
- K : [0.3 : 1.2]

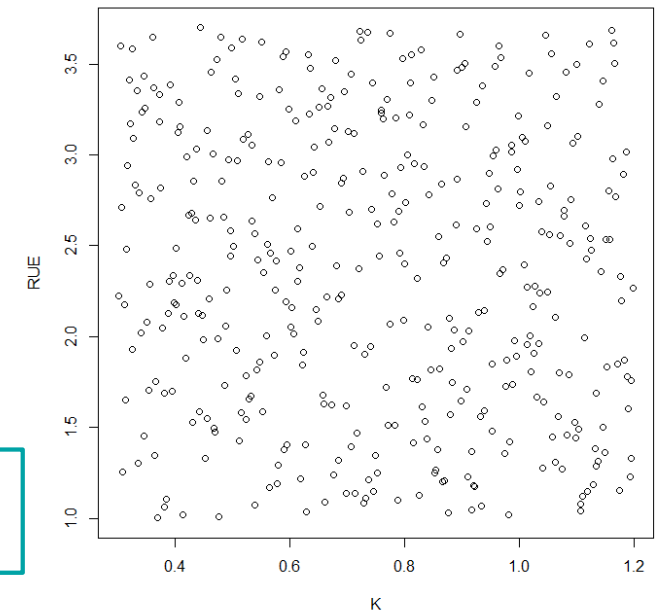
3 parameters to calibrate

- RUE : [1 : 3.7]
- K : [0.3 : 1.2]
- alpha : [0.005 : 0.01]

3000 simulations
3000 different values

Latin hypercube

- Random values within range for each parameter
 - 1 parameter = 1 dimension of the hypercube
- Better exploration of the space: **maximin algorithm**
 - Define total number of simulations
- Same number of simulations depending of the dimensions
- Be carefull to the resolution if you have several parameters



➤ Estimate the parameters of a complex model

3 – Run n time the model

- Essentially a programming step
- Use a wrapper to run automatically (in batch) n times the model
- Construct a « parameter matrix »



N simul	alpha	RUE	K
1	0.0005	1	0.3
2	0.0005095	1	0.3
3	0.0005190	1	0.3
4	0.0005285	1	0.3
...
...
n	0.01	3.7	1.2

➤ Estimate the parameters of a complex model

3 – Run n time the model

- Essentially a programming step
- Use a wrapper to run automatically (in batch) n times the model
- Construct a « parameter matrix »



N simul	alpha	RUE	K
1	0.0005	1	0.3
2	0.0005095	1	0.3
3	0.0005190	1	0.3
4	0.0005285	1	0.3
...
...
n	0.01	3.7	1.2

➤ Estimate the parameters of a complex model

3 – Run n time the model

- Essentially a programming step
- Use a wrapper to run automatically (in batch) n times the model
- Construct a « parameter matrix »




N simul	alpha	RUE	K
1	0.0005	1	0.3
2	0.0005095	1	0.3
3	0.0005190	1	0.3
4	0.0005285	1	0.3
...
...
n	0.01	3.7	1.2

➤ Estimate the parameters of a complex model

3 – Run n time the model

- Essentially a programming step
- Use a wrapper to run automatically (in batch) n times the model
- Construct a « parameter matrix »



N simul	alpha	RUE	K
1	0.0005	1	0.3
2	0.0005095	1	0.3
3	0.0005190	1	0.3
4	0.0005285	1	0.3
...
...
n	0.01	3.7	1.2

- Create a loop to run n times the model

➤ Estimate the parameters of a complex model

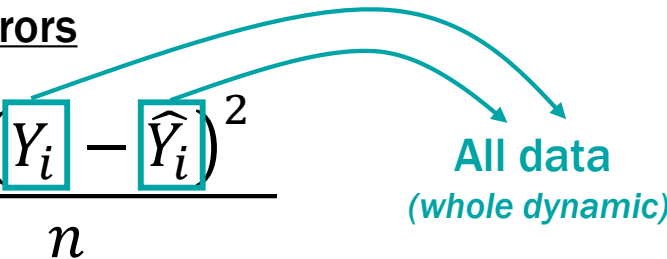
4 – Evaluate Parametrisation: calculate indicators (MSE, RMSE, etc.)

Several estimators of goodness of fit

Mean Squared Errors

$$MSE = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}$$

All data
(whole dynamic)



- Average of the squared differences between predictions and observations
- Sensitive to outliers (few simulations that are far from observations)

Root Mean Squared Errors

$$RMSE = \sqrt{MSE}$$

- Same unit than Y

Relative Root Mean Squared Errors

$$RRMSE = \frac{RMSE}{\bar{Y}}$$

- Dimensionless
- Error as a fraction of the average value
- Relevant to compare errors based on different datasets

Model Efficiency

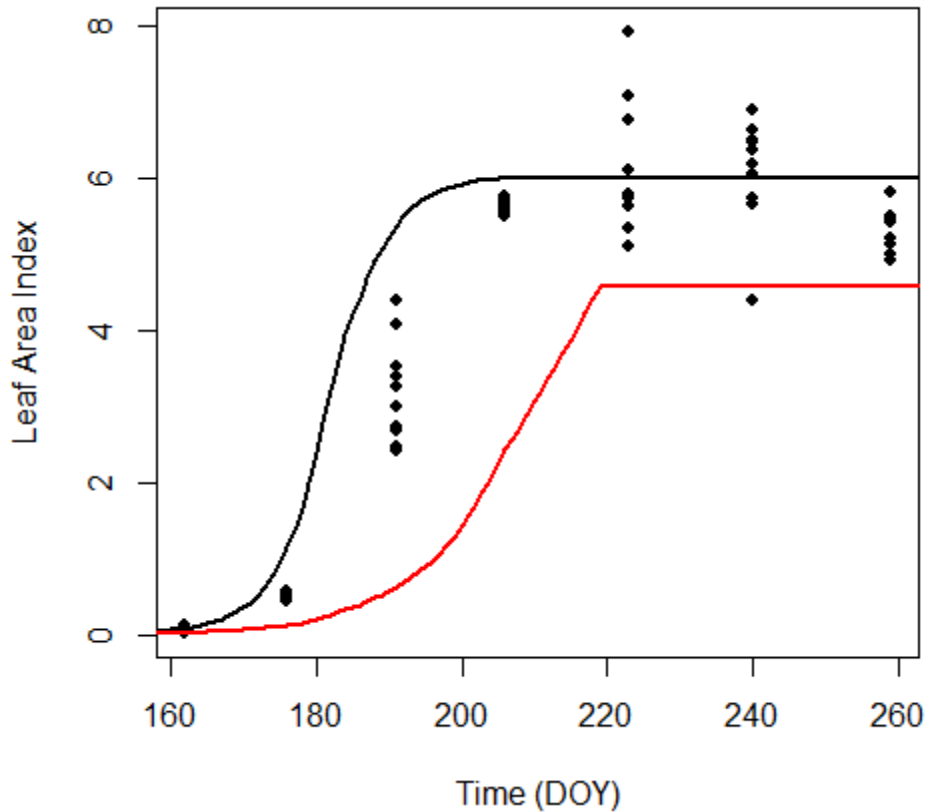
$$EF = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

- EF = 1 : Perfect model
- EF = 0 : Not a good model. The prediction is not better than the average value (no explanation of the variability)

➤ Estimate the parameters of a complex model

4 – Evaluate Parametrisation: calculate MSE, RMSE, etc.

Example of two simulations



- Calculate RMSE for each output variable (RMSE_Biomass, RMSE_LAI)

$$MSE = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n} \quad RMSE = \sqrt{MSE}$$

$$MSE = 1.056$$

$$MSE = 3.383$$

$$RMSE = 1.028$$

$$RMSE = 1.839$$

$$RRMSE = 0.267$$

$$RRMSE = 0.478$$

$$EF = 0.82$$

$$EF = 0.44$$

➤ Estimate the parameters of a complex model

3 – Run n time the model

- Calculate indicators previously calculated for each simulation for each variables
- Add them to the « parameter matrix »

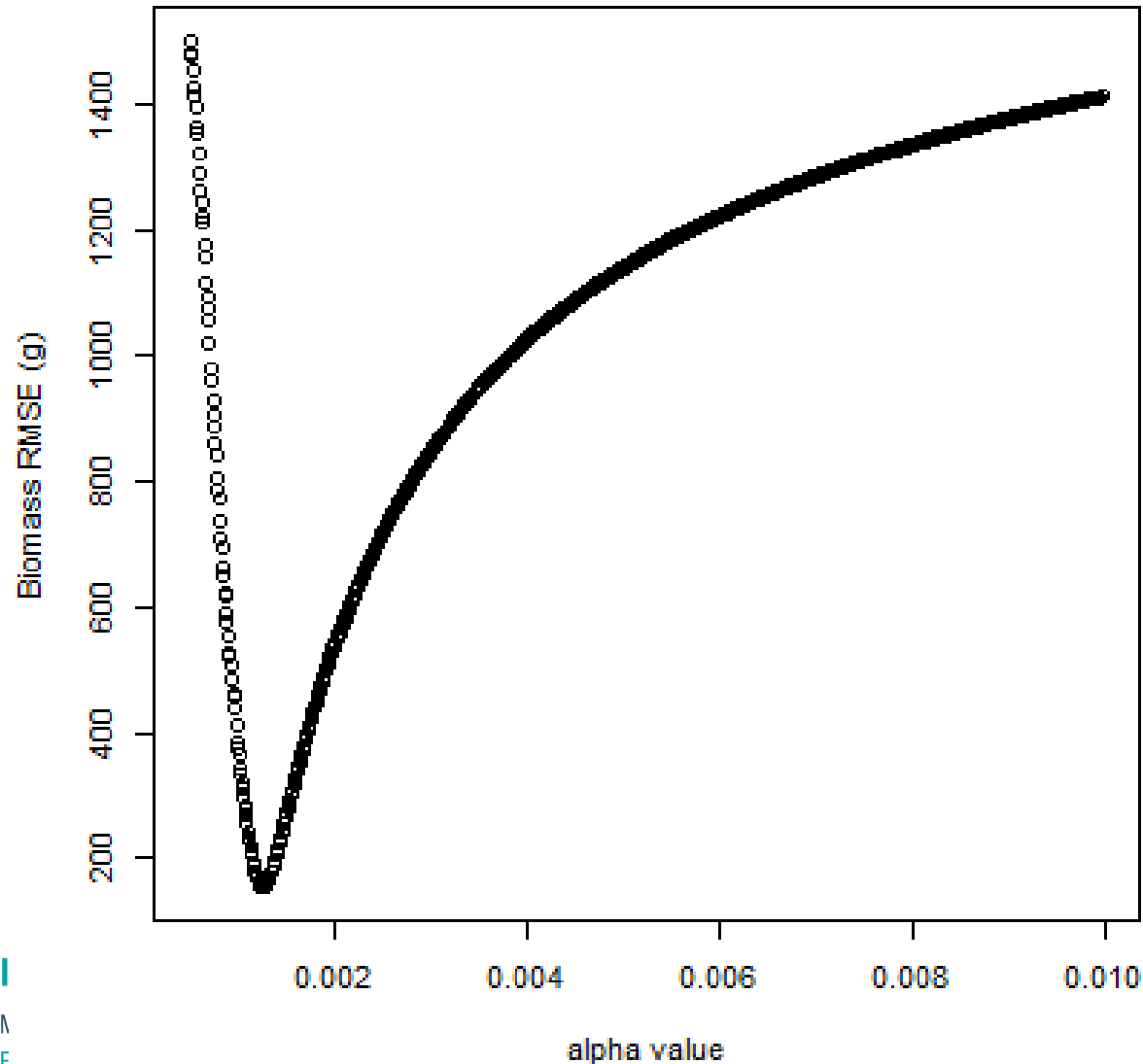
*Values of errors indicators
for each simulation*

N simul	alpha	RUE	K	MSE_B	RMSE_B	RRMSE_B	EF_B	MSE_LAI	RMSE_LAI	RRMSE_LAI	EF_LAI
1	0.0005	1	0.3	X	X	X	X	X	X	X	X
2	0.0005095	1	0.3	X	X	X	X	X	X	X	X
3	0.0005190	1	0.3	X	X	X	X	X	X	X	X
4	0.0005285	1	0.3	X	X	X	X	X	X	X	X
...	X	X	X	X	X	X	X	X
...	X	X	X	X	X	X	X	X
n	0.01	3.7	1.2	X	X	X	X	X	X	X	X

➤ Estimate the parameters of a complex model

4 – Evaluate Parametrisation: visual evaluation

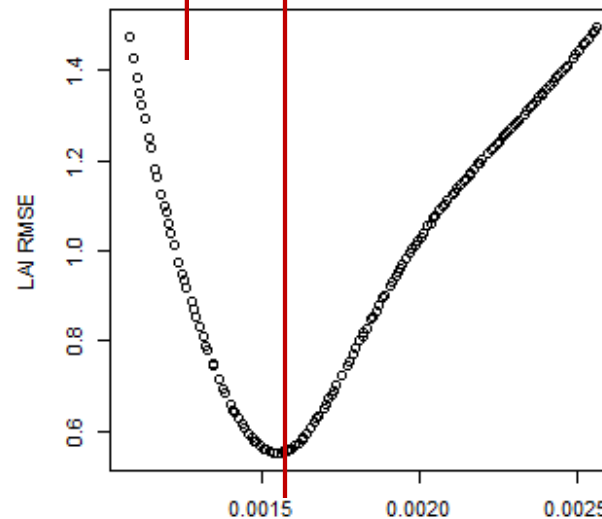
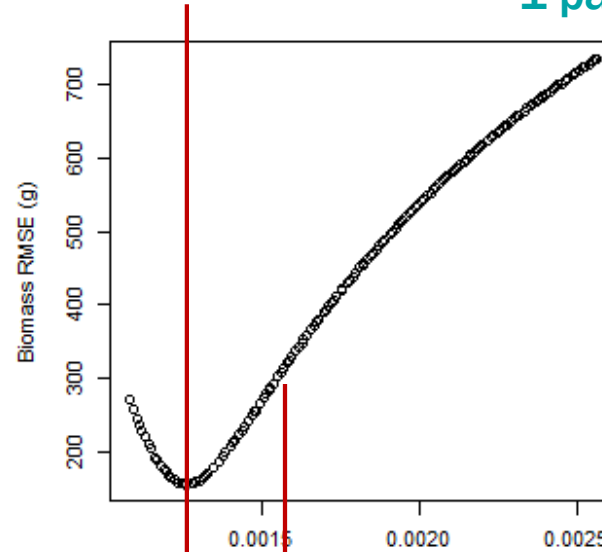
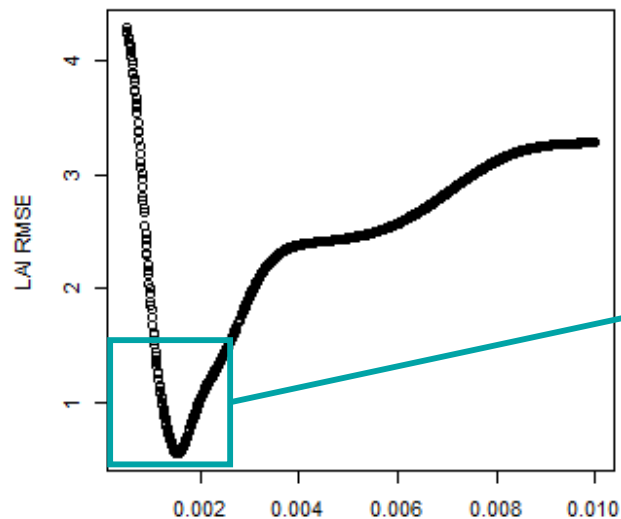
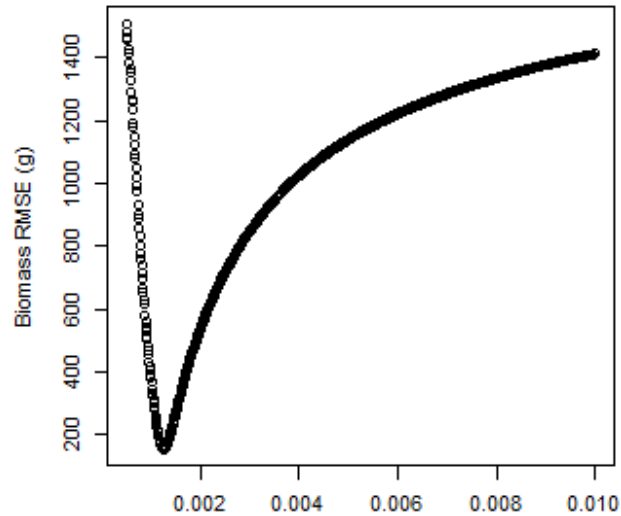
1 parameter to calibrate : alpha



➤ Estimate the parameters of a complex model

4 – Evaluate Parametrisation: visual evaluation

1 parameter to calibrate : alpha

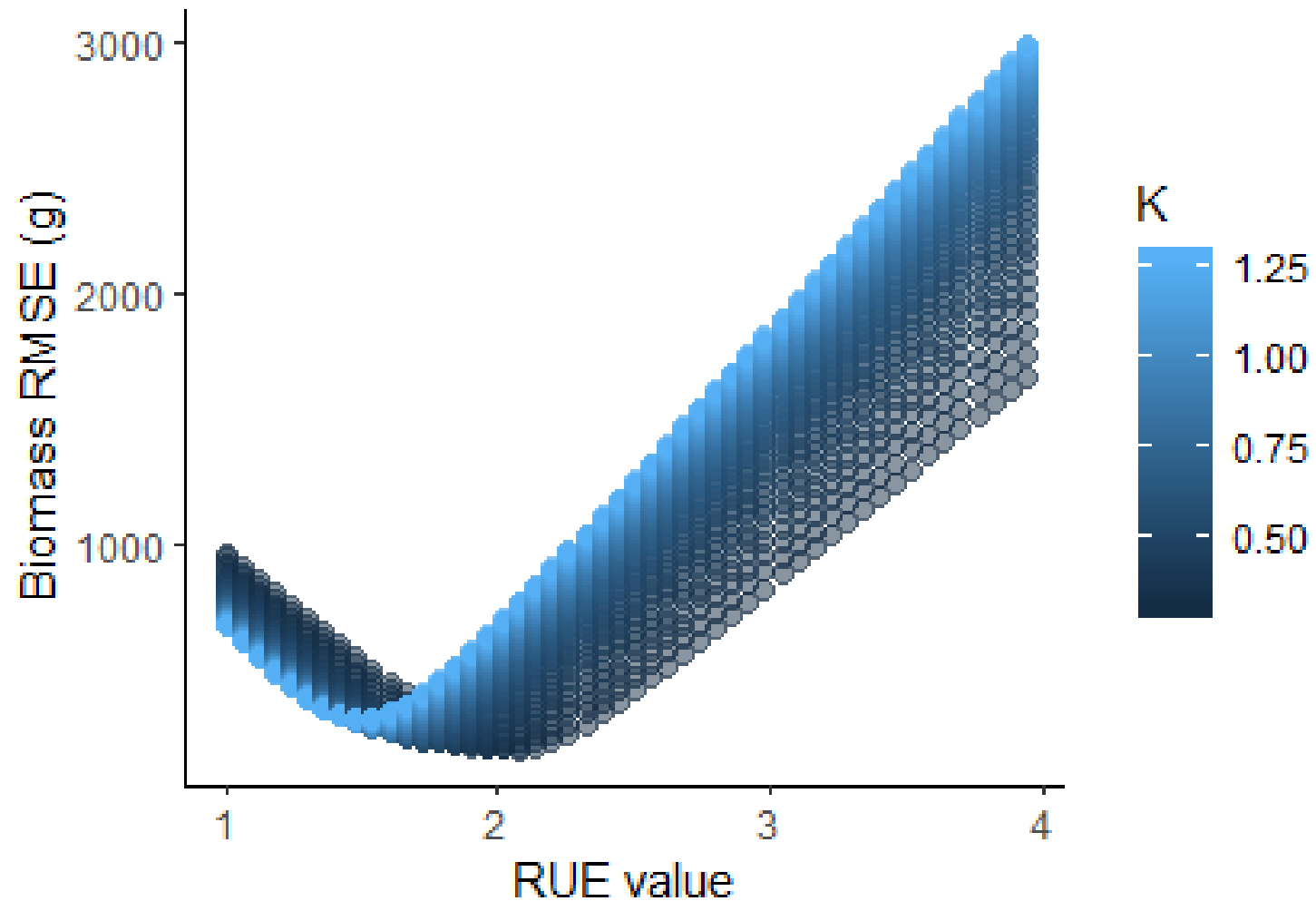


- Best alpha value is different depending on the output variable we calibrate on
- Choice to make between the variable
- Make a compromise putting a weight on calculated RMSE

➤ Estimate the parameters of a complex model

4 – Evaluate Parametrisation: visual evaluation

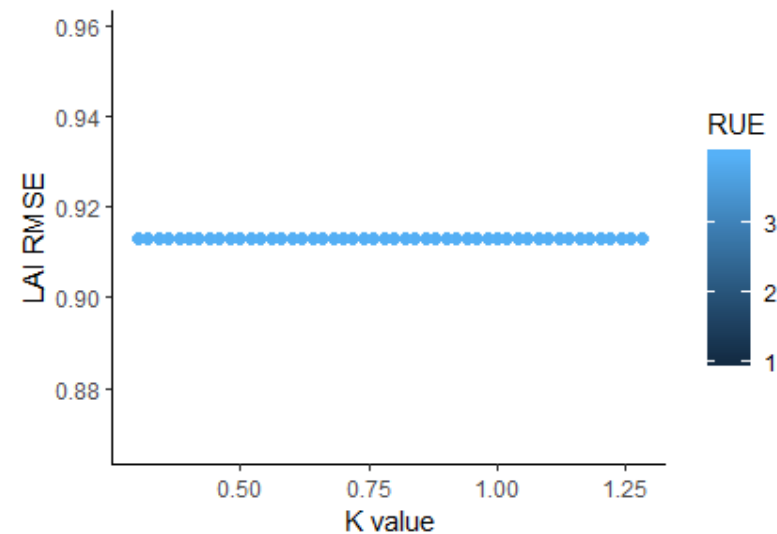
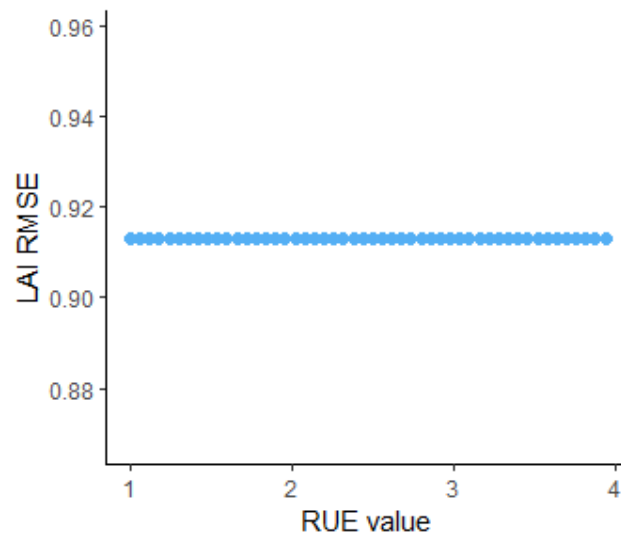
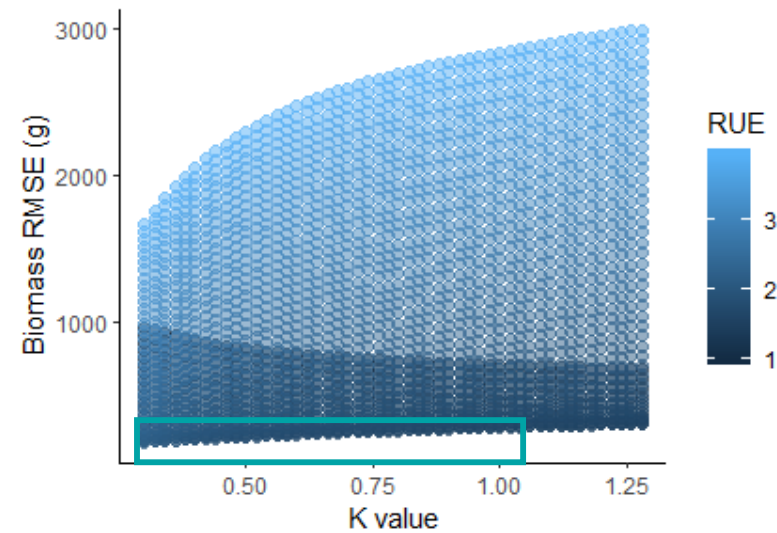
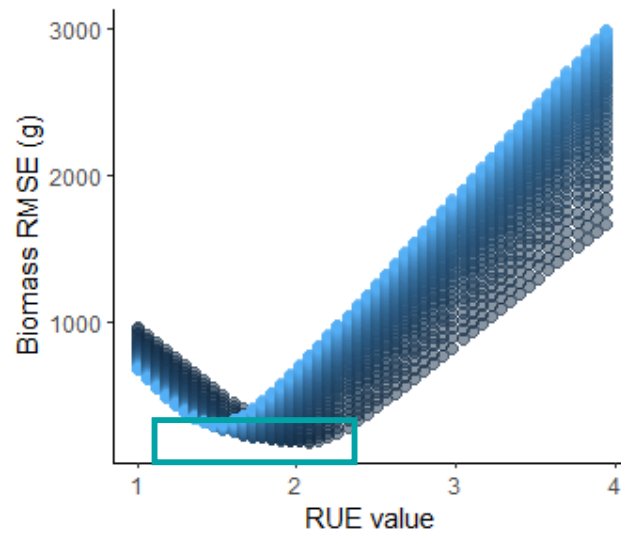
2 parameters to calibrate : K and RUE



➤ Estimate the parameters of a complex model

4 – Evaluate Parametrisation: visual evaluation

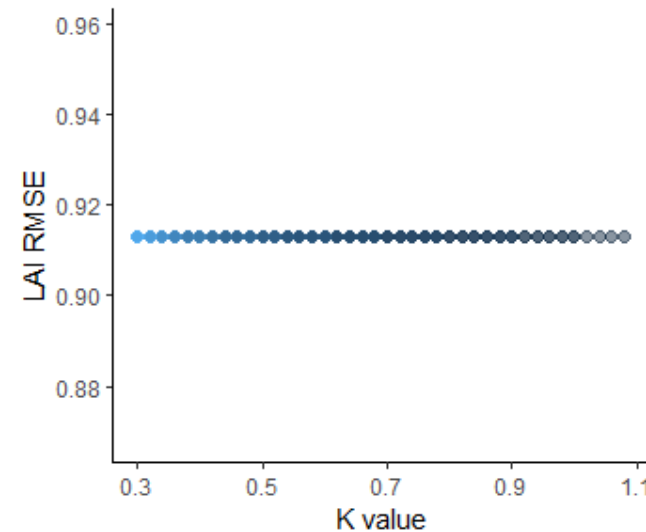
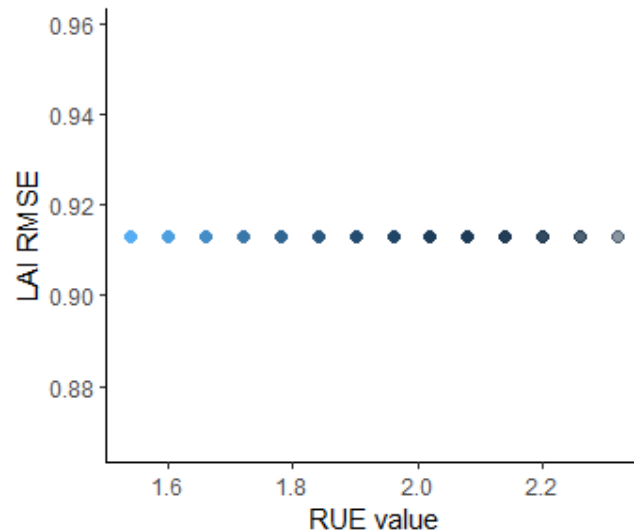
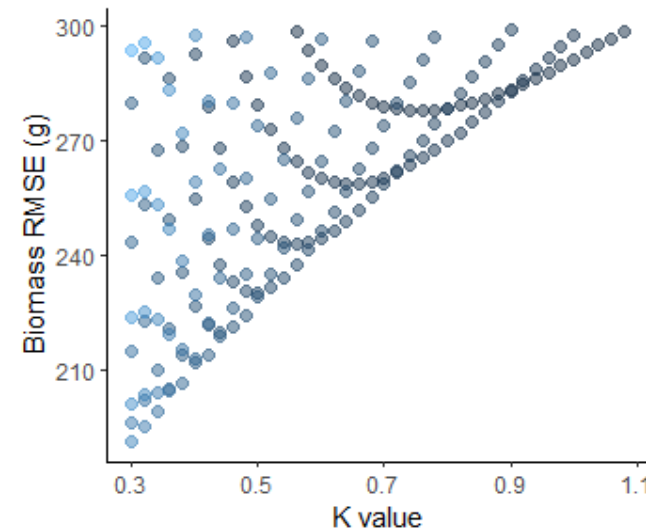
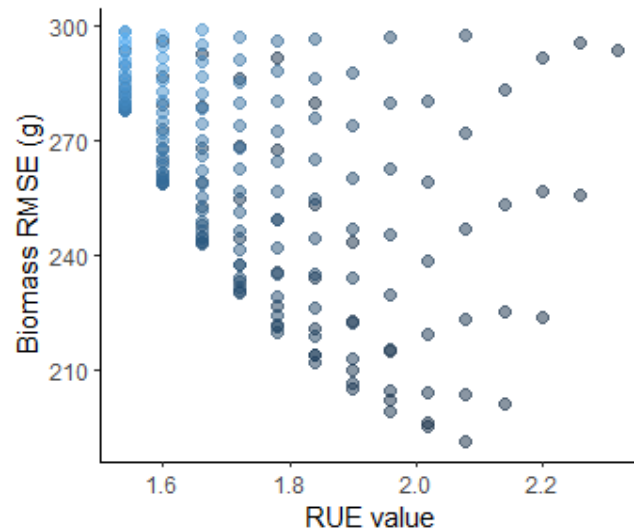
2 parameters to calibrate : K and RUE



➤ Estimate the parameters of a complex model

4 – Evaluate Parametrisation: visual evaluation

2 parameters to calibrate : K and RUE

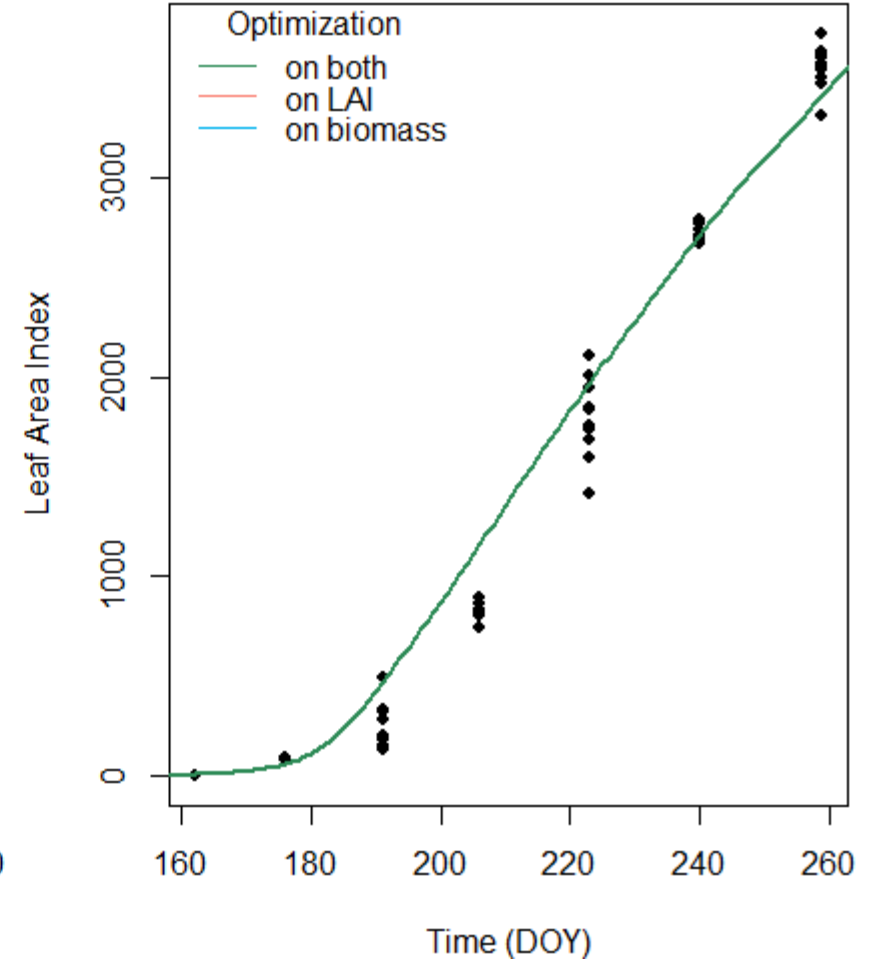
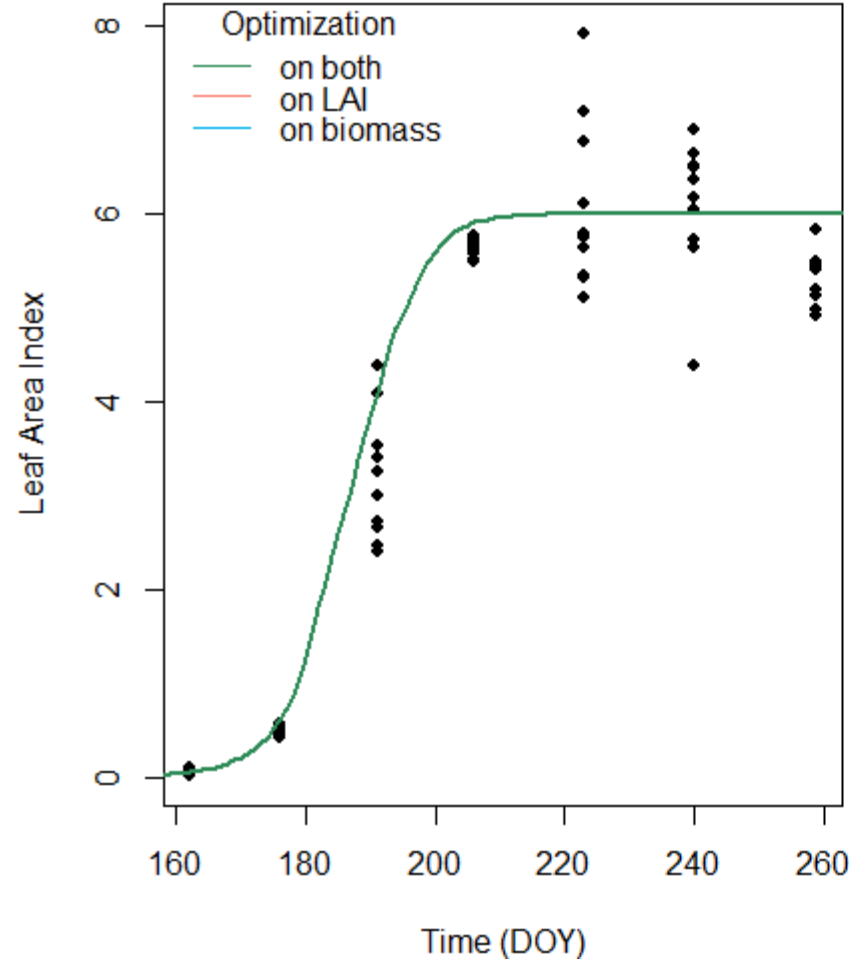


- No effect of K and RUE Choice to make between the variable
- Best parameter set :
K = 0.3 (lowest value)
RUE = 2.1

➤ Estimate the parameters of a complex model

5 – Evaluate Parametrisation: visual evaluation

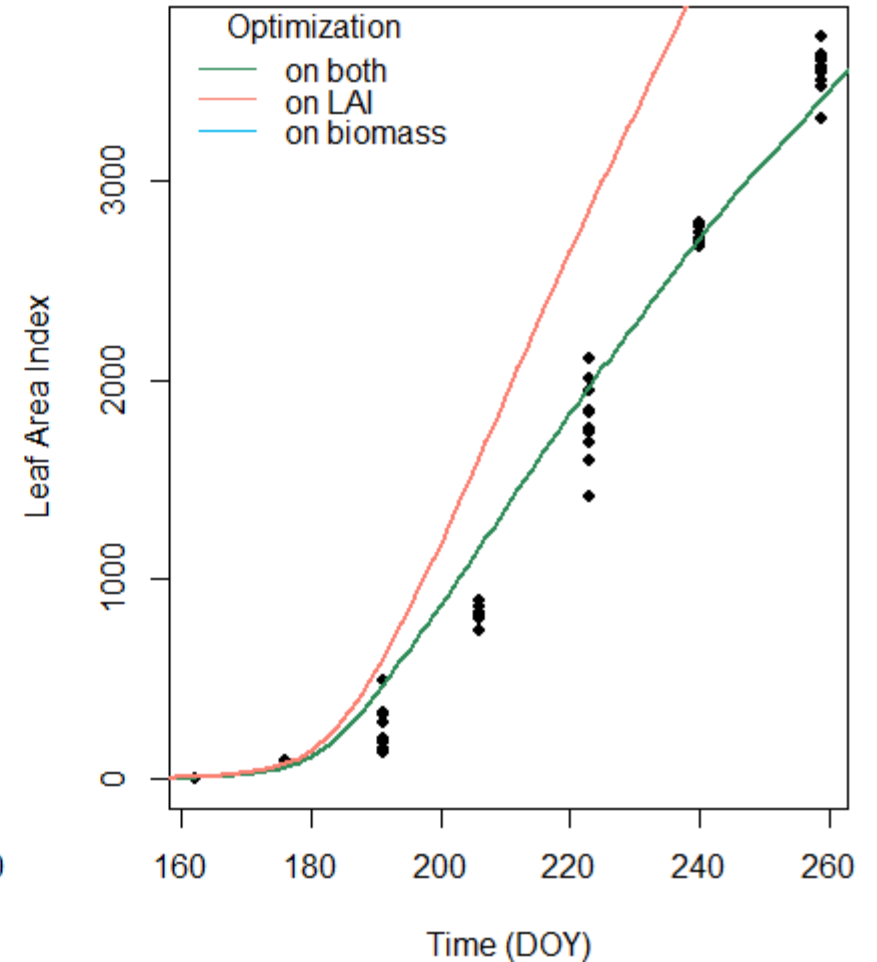
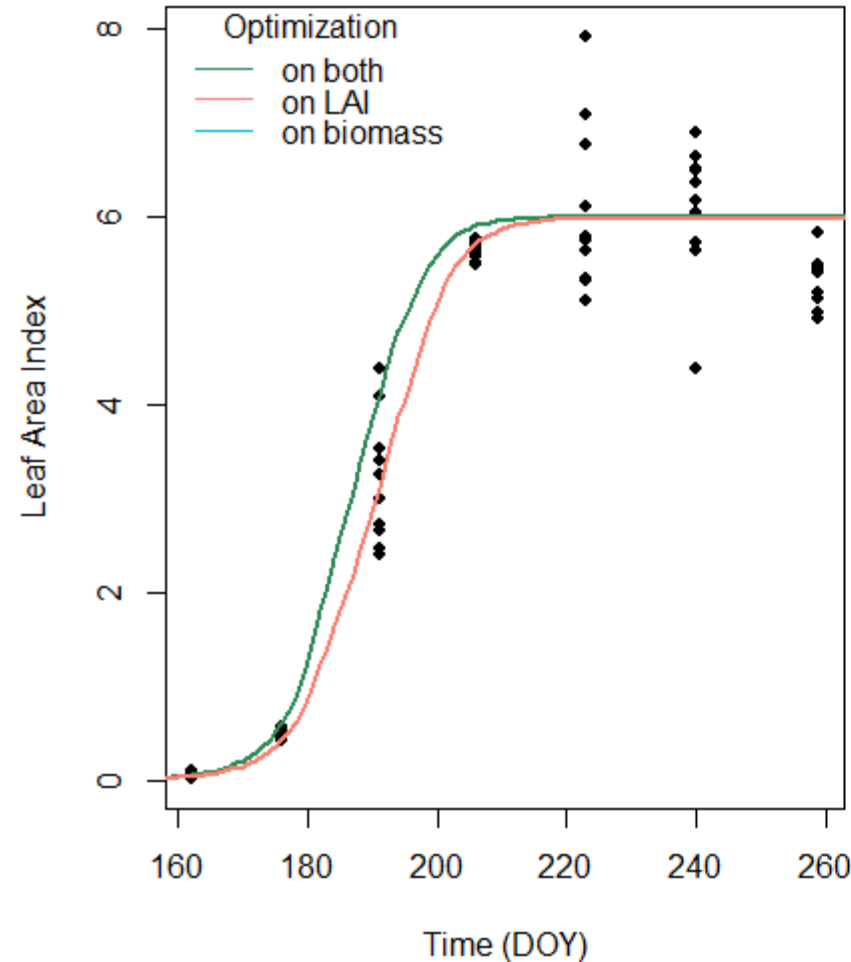
- Select the parameter set which give the **lowest RMSE**
- Which RMSE to choose?
Biomass RMSE?
LAI RMSE?
A mix ?



➤ Estimate the parameters of a complex model

5 – Evaluate Parametrisation: visual evaluation

- Select the parameter set which give the **lowest RMSE**
- Which RMSE to choose?
Biomass RMSE?
LAI RMSE?
A mix ?

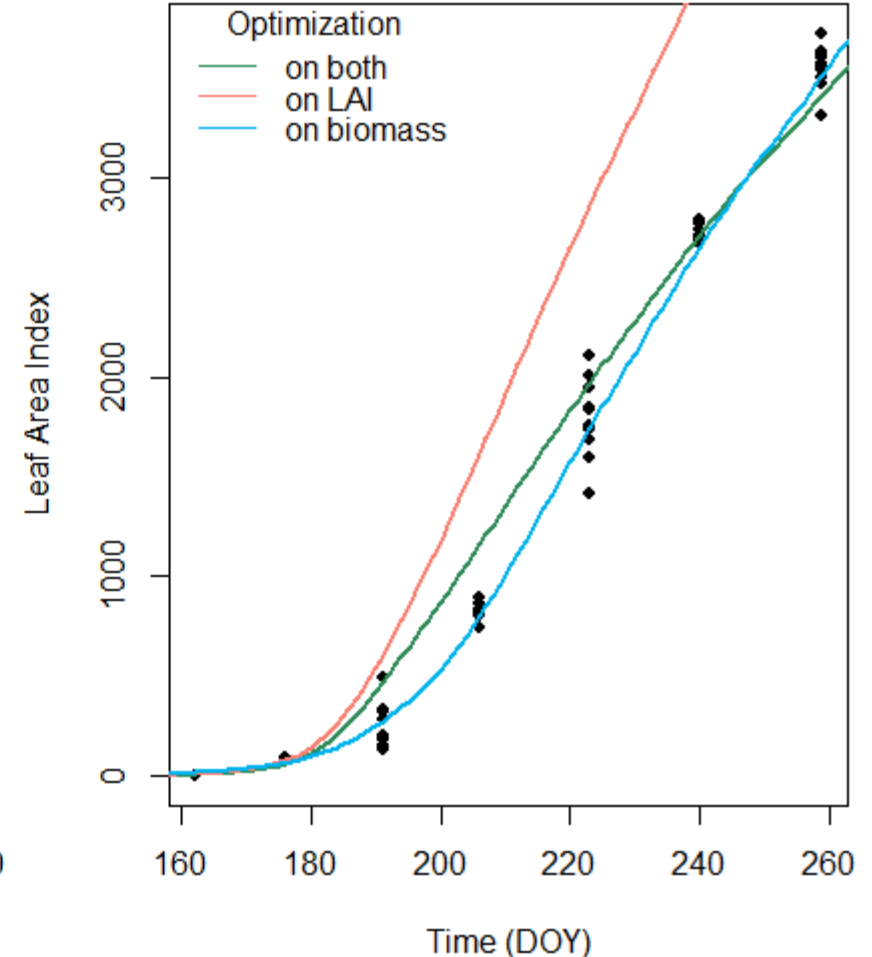
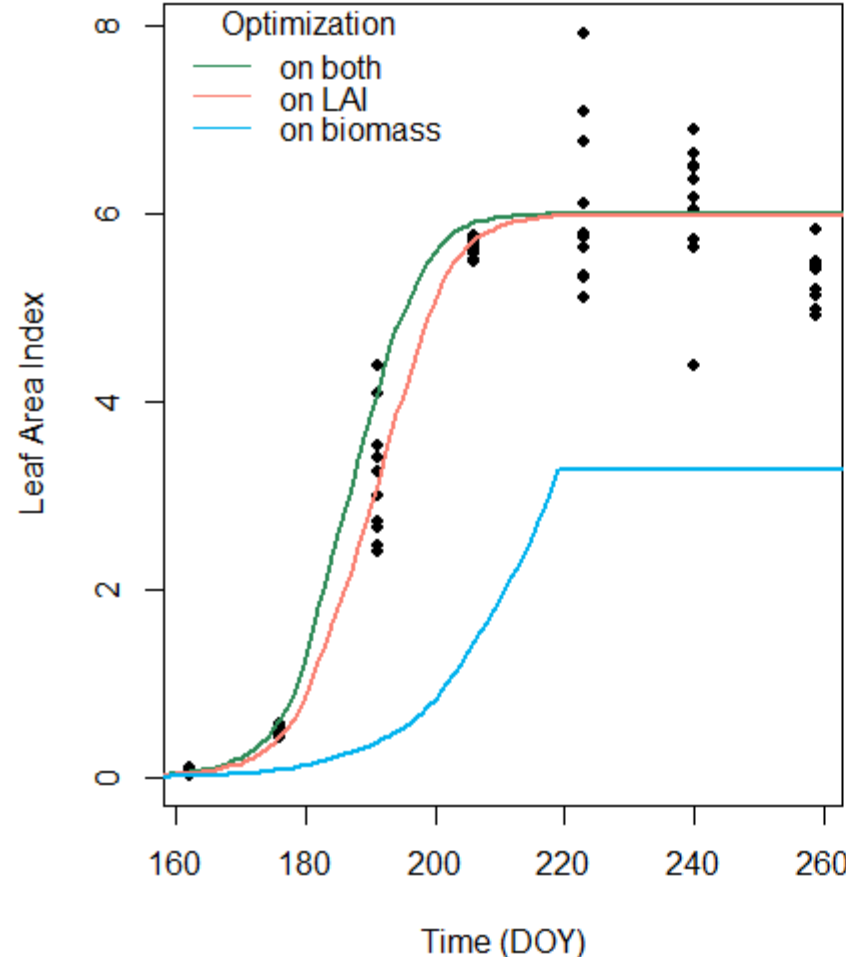


➤ Estimate the parameters of a complex model

5 – Evaluate Parametrisation: visual evaluation

- Select the parameter set which give the **lowest RMSE**
- Which RMSE to choose ?
Biomass RMSE?
LAI RMSE ?
A mix ?

Choose the way you
calibrate your model



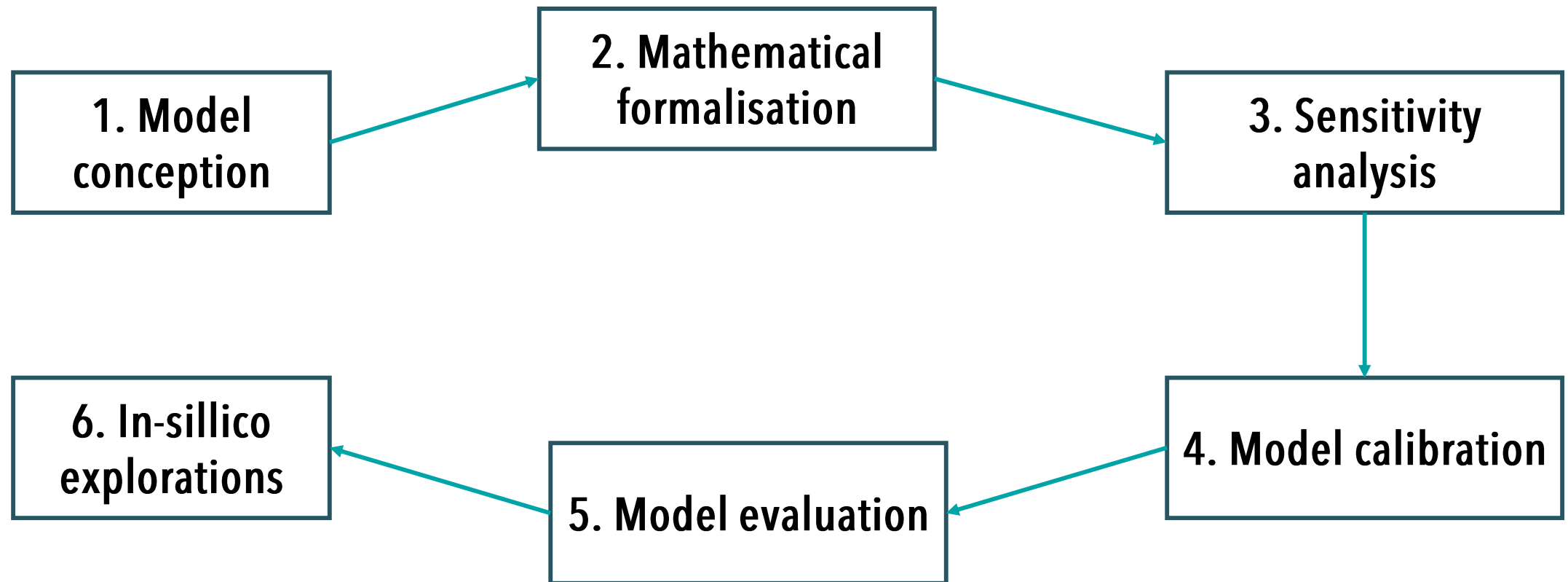
Just two reminders



> Introduction

How to build a model ?

Model construction from A to Z is composed of multiple steps:



➤ Estimate the parameters of a complex model

Steps to follow

1. Select **which** parameters to estimate

2. Design a **simulation scheme**
range of variation of each parameters
shape of the scheme

3. **Run n times** the model

4. Calculate **errors indicators**

5. Select the **best parameter set**

