#### True/False valuation of temporal logic formulae

The **True/False** valuation of temporal logic formulae is **not well adapted** to several problems :

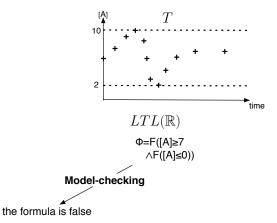
- parameter search, optimization and control of continuous models
- quantitative estimation of robustness
- sensitivity analyses

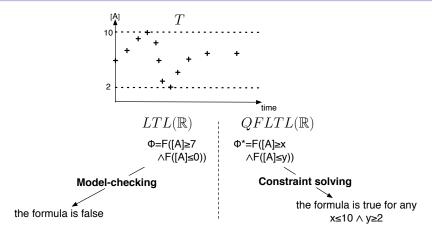
#### True/False valuation of temporal logic formulae

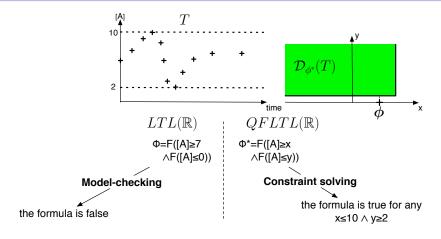
The **True/False** valuation of temporal logic formulae is **not well adapted** to several problems :

- parameter search, optimization and control of continuous models
- quantitative estimation of robustness
- sensitivity analyses
- ightarrow need for a continuous degree of satisfaction of temporal logic formulae

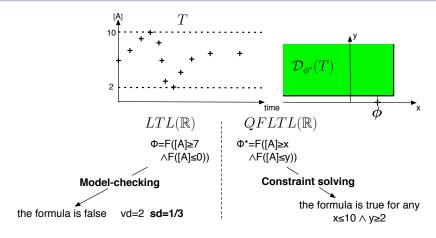
How far is the system from verifying the specification ?







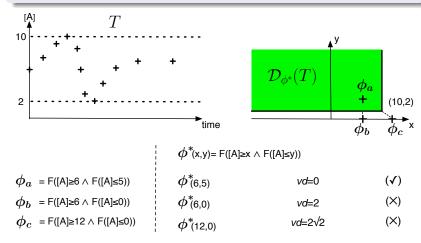
**Validity domain**  $\mathcal{D}_{\phi^*}(T)$  for the **free variables** in  $\phi^*$  [Fages Rizk CMSB'07]



Validity domain  $\mathcal{D}_{\phi^*}(T)$  for the free variables in  $\phi^*$  [Fages Rizk CMSB'07] Violation degree  $vd(T,\phi)=$  distance $(val(\phi),D_{\phi^*}(T))$  Satisfaction degree  $sd(T,\phi)=\frac{1}{1+vd(T,\phi)}\in[0,1]$ 

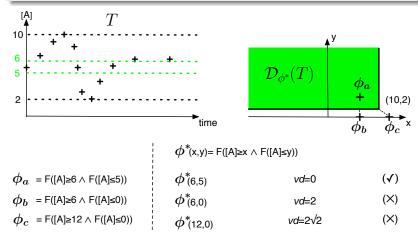
#### Definition of violation degree $vd(T,\phi)$ and satisfaction degree $sd(T,\phi)$

In the variable space of  $\phi^*$ , original formula  $\phi$  is single point  $var(\phi)$ .  $vd(T,\phi) = \min_{v \in D_{\phi^*}(T)} d(v, var(\phi)) \qquad sd(T,\phi) = \frac{1}{1+vd(T,\phi)} \in [0,1]$ 



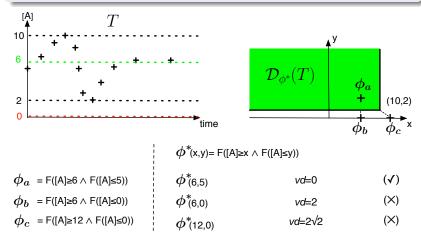
#### Definition of violation degree $vd(T,\phi)$ and satisfaction degree $sd(T,\phi)$

In the variable space of  $\phi^*$ , original formula  $\phi$  is single point  $var(\phi)$ .  $vd(T,\phi) = min_{v \in D_{\phi^*}(T)}d(v,var(\phi))$   $sd(T,\phi) = \frac{1}{1+vd(T,\phi)} \in [0,1]$ 



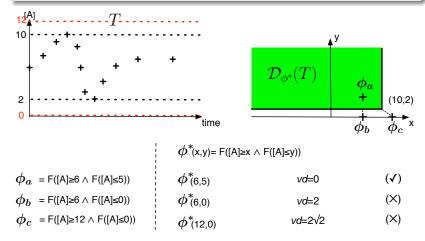
#### Definition of violation degree $vd(T,\phi)$ and satisfaction degree $sd(T,\phi)$

In the variable space of  $\phi^*$ , original formula  $\phi$  is single point  $var(\phi)$ .  $vd(T,\phi) = \min_{v \in D_{\phi^*}(T)} d(v, var(\phi)) \qquad sd(T,\phi) = \frac{1}{1+vd(T,\phi)} \in [0,1]$ 



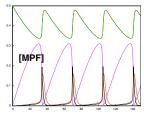
#### Definition of violation degree $vd(T, \phi)$ and satisfaction degree $sd(T, \phi)$

In the variable space of  $\phi^*$ , original formula  $\phi$  is single point  $var(\phi)$ .  $vd(T,\phi) = min_{v \in D_{\phi^*}(T)}d(v,var(\phi))$   $sd(T,\phi) = \frac{1}{1+vd(T,\phi)} \in [0,1]$ 



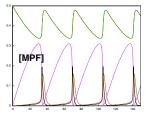
## Learning kinetic parameter values from LTL specifications

- simple model of the yeast cell cycle from [Tyson PNAS 91]
- models Cdc2 and Cyclin interactions (6 variables, 8 kinetic parameters)



#### Learning kinetic parameter values from LTL specifications

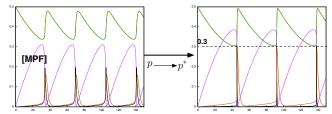
- simple model of the yeast cell cycle from [Tyson PNAS 91]
- models Cdc2 and Cyclin interactions (6 variables, 8 kinetic parameters)



• Pb : find values of 8 parameters such that amplitude is  $\geq$  0.3  $\phi^*$ : **F**( [A]>x  $\wedge$  **F**([A]<y) ) amplitude z=x-y goal : z = 0.3

## Learning kinetic parameter values from LTL specifications

- simple model of the yeast cell cycle from [Tyson PNAS 91]
- models Cdc2 and Cyclin interactions (6 variables, 8 kinetic parameters)



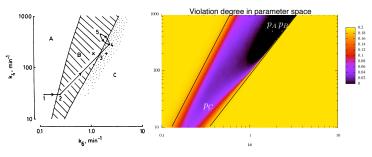
- Pb : find values of 8 parameters such that amplitude is  $\geq$  0.3  $\phi^*$ : **F**( [A]>x  $\wedge$  **F**([A]<y) ) amplitude z=x-y goal : z = 0.3
- ullet ightarrow solution found after 30s (100 calls to the fitness function)

## LTL Continuous Satisfaction Diagram

#### Example with:

- yeast cell cycle model [Tyson PNAS 91]
- oscillation of at least 0.3

$$\phi^* \colon$$
 F( [A]>x  $\land$  F([A]\geq0.3



Bifurcation diagram

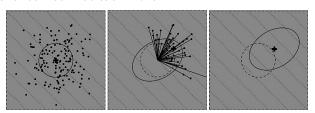
LTL satisfaction diagram

 Use existing non-linear optimization toolbox for kinetic parameter search using satisfaction degree as fitness function

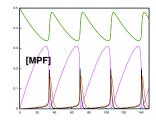
- Use existing non-linear optimization toolbox for kinetic parameter search using satisfaction degree as fitness function
- We use the state-of-the-art Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [Hansen Osermeier 01, Hansen 08]

- Use existing non-linear optimization toolbox for kinetic parameter search using satisfaction degree as fitness function
- We use the state-of-the-art Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [Hansen Osermeier 01, Hansen 08]
- CMA-ES maximizes an objective function in continuous domain in a black box scenario

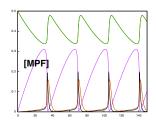
- Use existing non-linear optimization toolbox for kinetic parameter search using satisfaction degree as fitness function
- We use the state-of-the-art Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [Hansen Osermeier 01, Hansen 08]
- CMA-ES maximizes an objective function in continuous domain in a black box scenario
- CMA-ES uses a probabilistic neighborhood and updates information in covariance matrix at each move



## Learning Parameter Values from Period Constraints in LTL

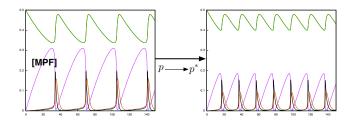


## Learning Parameter Values from Period Constraints in LTL



• Pb : find values of 8 parameters such that period is 20  $\phi^* : \mathbf{F}(\mathsf{MPF}_{localmaximum} \land \mathsf{Time} = t1 \land \ \mathbf{F}(\mathsf{MPF}_{localmaximum} \land \mathsf{Time} = t2) \ )$  (with  $\mathsf{MPF}_{localmaximum} : \mathsf{d}([\mathsf{MPF}])/\mathsf{dt} > 0 \land \ \mathbf{X}(\mathsf{d}([\mathsf{MPF}])/\mathsf{dt} < 0) \ )$  period z = t2 - t1 goal z = 20

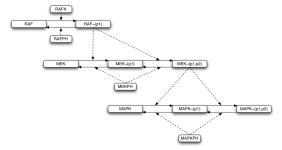
#### Learning Parameter Values from Period Constraints in LTL



- Pb : find values of 8 parameters such that period is 20  $\phi^*: \textbf{F}(\mathsf{MPF}_{localmaximum} \land \mathsf{Time} = t1 \land \ \textbf{F}(\mathsf{MPF}_{localmaximum} \land \mathsf{Time} = t2) \ )$  (with  $\mathsf{MPF}_{localmaximum} : \mathsf{d}([\mathsf{MPF}])/\mathsf{dt} > 0 \land \ \textbf{X}(\mathsf{d}([\mathsf{MPF}])/\mathsf{dt} < 0) \ )$  period z = t2 t1 goal z = 20
- → Solution found after 60s (200 calls to the fitness function)

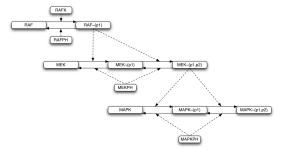
# Oscillations in MAPK signal transduction cascade

• MAPK signaling model [Huang Ferrel PNAS 96]



## Oscillations in MAPK signal transduction cascade

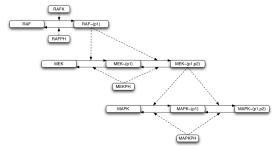
MAPK signaling model [Huang Ferrel PNAS 96]



- search for oscillations in **37 dimensions** (30 parameters and 7 initial conditions)
  - $\rightarrow$  solution found after 3 min (200 calls to the fitness function) Oscillations already observed by simulation [Qiao et al. 07]

## Oscillations in MAPK signal transduction cascade

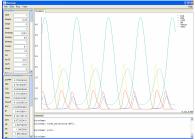
MAPK signaling model [Huang Ferrel PNAS 96]



- **search for oscillations** in **37 dimensions** (30 parameters and 7 initial conditions)
  - $\rightarrow$  solution found after 3 min (200 calls to the fitness function) Oscillations already observed by simulation [Qiao et al. 07]
- No negative feedback in the reaction graph, but negative circuits in the influence graph [Fages Soliman FMSB'08, CMSB'06]

## Coupled Models of Cell Cycle, Circadian Clock, DNA repair

- Context of colorectal cancer chronotherapies
   EU FP6 TEMPO, EraSysBio C5Sys, coord. F. Lévi INSERM Villejuif
- Coupled model of the cell cycle [Tyson Novak 04][Gerard Goldbeter 09] and the circadian clock [Leloup Goldbeter 99] with condition of entrainment in period [Calzone Soliman 06]



 Coupled model with DNA repair system p53/Mdm2 [Cilberto et al.04], metabolism of irinotecan, and drug administration optimization [De Maria Soliman Fages 09 CMSB]

## Basis of Operators of LTL(R)

Atomic propositions: arithmetic expressions with  $<, \leq, =, \geq, >$  over the state variables (closed by negation)

Duality: 
$$\neg \mathbf{X}\phi = \mathbf{X}\neg \phi$$
,  $\neg \mathbf{F}\phi = \mathbf{G}\neg \phi$ ,  $\neg \mathbf{G}\phi = \mathbf{F}\neg \phi$ ,  $\neg (\phi \mathbf{U} \psi) = (\neg \psi \mathbf{W} \neg \phi)$ ,  $\neg (\phi \mathbf{W} \psi) = (\neg \psi \mathbf{U} \neg \phi)$ ,

Properties: 
$$\mathbf{F}\phi = \text{true } \mathbf{U} \ \phi$$
,  $\mathbf{G}\phi = \phi \ \mathbf{W} \ \text{false}$ ,  $\phi \mathbf{W}\psi = \mathbf{G}\phi \lor (\phi \mathbf{U}(\phi \land \psi))$ 

Negation free formulae: expressed with  $\land$ ,  $\lor$ , F, G, U, X with negations eliminated down to atomic propositions.

## LTL(R) model-checking

Given a finite trace T and an LTL(R) formula  $\phi$ 

- lacktriangled label each state with the atomic sub-formulae of  $\phi$  that are true at this state;
- ② add sub-formulae of the form  $\phi_1$  U  $\phi_2$  to the states labeled by  $\phi_2$  and to the predecessors of states labeled with  $\phi_2$  as long as they are labeled by  $\phi_1$ ;
- **3** add sub-formulae of the form  $\phi_1$  **W**  $\phi_2$  to the last state if it is labeled by  $\phi_1$ , and to the states labeled by  $\phi_1$  and  $\phi_2$ , and to their predecessors as long as they are labeled by  $\phi_1$ ;
- add sub-formulae of the form  $\mathbf{X}\phi$  to the last state if it is labeled by  $\phi$  and to the immediate predecessors of states labeled by  $\phi$ ;
- **5** return the vertices labeled by  $\phi$ .

## QFLTL(R) Formulae with Variables

Quantifier free LTL formulae, noted  $\phi(\mathbf{y})$  with free variables  $\mathbf{y}$ 

The satisfaction domain of  $\phi(\mathbf{y})$  in a trace T is the set of  $\mathbf{y}$  values for which  $\phi(\mathbf{y})$  holds:

$$\mathcal{D}_{\mathcal{T},\phi(\mathbf{y})} = \{ \mathbf{y} \in \mathbb{R}^q \mid \mathcal{T} \models \phi(\mathbf{y}) \}$$
 (1)

For linear constraints over R, satisfaction domains can be computed with polyhedral libraries.

Biocham uses the Parma Polyhedral Library PPL

# QFLTL(R) constraint solving

The satisfaction domains of QFLTL formulae satisfy the equations:

• 
$$\mathcal{D}_{T,\phi(\mathbf{y})} = \mathcal{D}_{s_0,\phi(\mathbf{y})}$$
,

• 
$$\mathcal{D}_{s_i,\pi(\mathbf{y})} = \{\mathbf{y} \in \mathbb{R}^m \mid s_i \models_{\mathcal{R}} \pi(\mathbf{y})\},$$

$$\bullet \ \mathcal{D}_{s_i,\phi(\mathbf{y})\wedge\psi(\mathbf{y})} = \mathcal{D}_{s_i,\phi(\mathbf{y})}\cap \mathcal{D}_{s_i,\psi(\mathbf{y})},$$

• 
$$\mathcal{D}_{s_i,\phi(\mathbf{y})\vee\psi(\mathbf{y})}=\mathcal{D}_{s_i,\phi(\mathbf{y})}\cup\mathcal{D}_{s_i,\psi(\mathbf{y})}$$
,

• 
$$\mathcal{D}_{s_i, \mathbf{F}\phi(\mathbf{y})} = \cup_{j \in [i, n]} \mathcal{D}_{s_i, \phi(\mathbf{y})}$$
,

$$\bullet \ \mathcal{D}_{s_i,\mathbf{G}\phi(\mathbf{y})} = \cap_{j\in[i,n]} \mathcal{D}_{s_j,\phi(\mathbf{y})},$$

$$\bullet \ \mathcal{D}_{s_i,\phi(\mathbf{y})\mathsf{U}\psi(\mathbf{y})} = \cup_{j\in[i,n]} (\mathcal{D}_{s_j,\psi(\mathbf{y})} \cap \cap_{k\in[i,j-1]} \mathcal{D}_{s_k,\phi(\mathbf{y})}),$$

$$\bullet \ \mathcal{D}_{s_i, \mathbf{X} \phi(\mathbf{y})} = \left\{ \begin{array}{ll} \mathcal{D}_{s_{i+1}, \phi(\mathbf{y})}, & \text{ if } i < n, \\ \mathcal{D}_{s_i, \phi(\mathbf{y})}, & \text{ if } i = n, \end{array} \right.$$

## Complexity with bound constraints x > b, x < b

Bound constraints define boxes  $\mathcal{R}_i \in \mathbb{R}^{\nu}$ .

Let the size of a union of boxes be the least integer k such that  $\mathcal{D} = \bigcup_{i=1}^k \mathcal{R}_i$ .

#### Proposition (complexity of the solution domain)

The validity domain of a QFLTL formula of size f containing v variables on a trace of length n is a union of boxes of size less than  $(nf)^{2v}$ .

The maximum number of bounds for a variable x is  $n \times f$  E.g;  $\mathbf{F}([A] = u \vee [A] + 1 = u \vee \cdots \vee [A] + f = u)$ 

**F**([
$$A_1$$
] =  $X_1 \lor [A_1] + 1 = X_1 \lor ... \lor [A_1] + f = X_1) \land ...$   
 ∧ **F**([ $A_v$ ] =  $X_v \lor [A_v] + 1 = X_v \lor ... \lor [A_v] + f = X_v)$   
 has a solution domain of size  $(nf)^v$  on a trace of  $n$  values with  $[A_i] + k$  all different for  $1 \le i \le v$ ,  $0 \le k \le f$ .

#### Robustness Measure Definition

Robustness defined with respect to:

- a biological system
- a functionality property  $D_a$
- a set *P* of perturbations
- General notion of robustness proposed in [Kitano MSB 07]:

$$\mathcal{R}_{a,P} = \int_{p \in P} D_a(p) \ prob(p) \ dp$$

#### Robustness Measure Definition

Robustness defined with respect to:

- a biological system
- a functionality property  $D_a$
- a set P of perturbations
- General notion of robustness proposed in [Kitano MSB 07]:

$$\mathcal{R}_{a,P} = \int_{p \in P} D_a(p) \; prob(p) \; dp$$

• Computational measure of robustness w.r.t. LTL( $\mathbb{R}$ ) spec:

$$\mathcal{R}_{\phi,P} = \int_{p \in P} sd(T(p),\phi) \ prob(p) \ dp$$

where T(p) is the trace obtained by numerical integration of the ODE for perturbation p

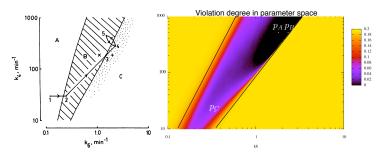
## Robustness analysis w.r.t parameter perturbations

#### Example with:

- cell cycle model [Tyson PNAS 91]
- oscillation of amplitude at least 0.2

$$\phi^*$$
: **F**( [A]>x  $\wedge$  **F**([A]\geq0.2

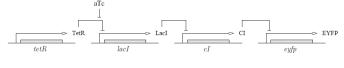
• parameters normally distributed,  $\mu = p_{ref}$ , CV=0.2

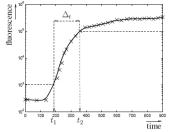


$$\mathcal{R}_{\phi,p_A} = 0.83$$
,  $\mathcal{R}_{\phi,p_B} = 0.43$ ,  $\mathcal{R}_{\phi,p_C} = 0.49$ 

## Application to Synthetic Biology in E. Coli

Cascade of transcriptional inhibitions added to *E.coli* [Weiss et al PNAS 05] **input** small molecule aTc **output** protein EYFP





**Specification**: EYFP has to remain below  $10^3$  for at least 150 min., then exceeds  $10^5$  after at most 450 min., and switches from low to high levels in less than 150 min.

# Specifying the expected behavior in $\mathsf{QFLTL}(\mathbb{R})$

The timing specifications can be formalized in temporal logic as follows:

$$\begin{array}{ll} \phi(t_1,t_2) = & \quad \textbf{G}(\textit{time} < t_1 \to [\texttt{EYFP}] < 10^3) \\ & \wedge \quad \textbf{G}(\textit{time} > t_2 \to [\texttt{EYFP}] > 10^5) \\ & \wedge \quad t_1 > 150 \wedge t_2 < 450 \wedge t_2 - t_1 < 150 \end{array}$$

which is abstracted into

$$\phi(t_1, t_2, b_1, b_2, b_3) = egin{array}{ccc} {f G}(\textit{time} < t_1 
ightarrow [{ t EYFP}] < 10^3) \ & {f G}(\textit{time} > t_2 
ightarrow [{ t EYFP}] > 10^5) \ & {f \Lambda} & t_1 > b1 {f \Lambda} & t_2 < b_2 {f \Lambda} & t_2 - t_1 < b_3 \end{array}$$

for computing validity domains for  $b_1, b_2, b_3$ 

with the objective  $b_1 = 150$ ,  $b_2 = 450$ ,  $b_3 = 150$  for computing the satisfaction degree in a given trace.

Variance-based global sensitivity indices

$S_i = rac{Var(E(R P_i))}{Var(R)} \in [0,1]$	$S_{\gamma}$	20.2 %	$S_{\kappa_{eyfp},\gamma}$	8.7 %
	$S_{\kappa_{evfp}}$	7.4 %	$S_{\kappa_{cl},\gamma}$	6.2 %
	$S_{\kappa_{eyfp}} \ S_{\kappa_{cl}}$	6.1%	$S_{\kappa^0_{cl},\gamma}$	5.0 %
	$S_{\kappa 0}$	3.3 %	$S_{\kappa_{cl}^{0},\kappa_{eyfp}}^{cl}$	2.8 %
	$S_{\kappa cl}^{lacl}$	2.0%	$S_{\kappa_{cl},\kappa_{eyfp}}$	1.8 %
$Var(R) = L^{2}$	$S_{\kappa_{lacl}}^{cl}$	1.5 %	$S_{\kappa_{eyfp}^0,\gamma}$	1.5 %
	$S_{\kappa_{eyfp}^0}$	0.9%	$S_{\kappa_{cl}^{0},\kappa_{cl}}^{\epsilon_{ll},\kappa_{cl}}$	1.1 %
	S <sub>uaTc</sub>	0.4%	$S_{\kappa_{cl}^{0},\kappa_{lacl}}^{0}$	0.5 %
	total first order	40.7 %	total second order	31.2 %

degradation factor  $\gamma$  has the strongest impact on the cascade.

Variance-based global sensitivity indices

$S_i = rac{Var(E(R P_i))}{Var(R)} \in [0,1]$	$S_{\gamma}$	20.2 %	$S_{\kappa_{eyfp},\gamma}$	8.7 %
	$S_{\kappa_{\it eyfp}}$	7.4 %	$S_{\kappa_{cl},\gamma}^{s,r}$	6.2 %
	$S_{\kappa_{cl}}$	6.1 %	$S_{\kappa^0_{cl},\gamma}$	5.0 %
	$S_{\kappa_{lacl}}^{0}$	3.3 %	$S_{\kappa_{cl},\kappa_{eyfp}}^{cl}$	2.8 %
	$S_{\kappa_{cl}^{0}}^{lacl}$	2.0 %	$S_{\kappa_{cl},\kappa_{eyfp}}^{\kappa_{cl},\kappa_{eyfp}}$	1.8 %
Var(R) = [0, -1]	$S_{\kappa_{lacl}}^{cl}$	1.5 %	$S_{\kappa_{eyfp}^0,\gamma}$	1.5 %
	$S_{\kappa_{eyfp}^0}$	0.9%	$S_{\kappa_{cl}^{0},\kappa_{cl}}^{\text{eyrp}}$	1.1%
	S <sub>uaTc</sub>	0.4%	$S_{\kappa_{cl}^{0},\kappa_{lacl}}^{0}$	0.5 %
	total first order	40.7 %	total second order	31.2 %
	·			

degradation factor  $\boldsymbol{\gamma}$  has the strongest impact on the cascade.

aTc variations have a very low impact

Variance-based global sensitivity indices

	$S_{\gamma}$	20.2 %	$S_{\kappa_{\it eyfp},\gamma}$	8.7 %
	$S_{\kappa_{\it eyfp}}$	7.4%	$S_{\kappa_{cl},\gamma}$	6.2 %
	$S_{\kappa_{cl}}$	6.1 %	$S_{\kappa^0_{cl},\gamma}$	5.0 %
	$S_{\kappa_{lacl}}^{0}$	3.3 %	$S_{\kappa_{cl}^{0},\kappa_{eyfp}}^{cl}$	2.8 %
$S_i = \frac{Var(E(R P_i))}{Var(R)} \in [0,1]$	$S_{\kappa_{cl}^0}^{lacl}$	2.0 %	$S_{\kappa_{cl},\kappa_{eyfp}}$	1.8 %
$S_I = \frac{1}{Var(R)} = [0, 1]$	$S_{\kappa_{lacl}}^{cl}$	1.5 %	$S_{\kappa_{eyfp}^0,\gamma}$	1.5 %
	$S_{\kappa_{eyfp}^0}$	0.9%	$S_{\kappa_{cl}^{0},\kappa_{cl}}^{eyip}$	1.1%
	S <sub>uaTc</sub>	0.4%	$S_{\kappa_{cl}^{0},\kappa_{lacl}}^{0}$	0.5 %
	total first order	40.7 %	total second order	31.2 %
			•	

degradation factor  $\gamma$  has the strongest impact on the cascade.

aTc variations have a very low impact

different importance of the basal  $\kappa_{\it eyfp}^0$  and regulated  $\kappa_{\it eyfp}$  EYFP production rates

Variance-based global sensitivity indices

$S_i = rac{Var(E(R P_i))}{Var(R)} \in [0,1]$	$S_{\gamma}$	20.2 %	$S_{\kappa_{eyfp},\gamma}$	8.7 %
	$S_{\kappa_{evfp}}$	7.4 %	$S_{\kappa_{cl},\gamma}$	6.2 %
	$S_{\kappa_{eyfp}} \ S_{\kappa_{cl}}$	6.1%	$S_{\kappa^0_{cl},\gamma}$	5.0 %
	$S_{\kappa 0}$	3.3 %	$S_{\kappa_{cl}^{0},\kappa_{eyfp}}^{cl}$	2.8 %
	$S_{\kappa cl}^{lacl}$	2.0%	$S_{\kappa_{cl},\kappa_{eyfp}}$	1.8 %
$Var(R) = L^{2}$	$S_{\kappa_{lacl}}^{cl}$	1.5 %	$S_{\kappa_{eyfp}^0,\gamma}$	1.5 %
	$S_{\kappa_{eyfp}^0}$	0.9%	$S_{\kappa_{cl}^{0},\kappa_{cl}}^{\epsilon_{ll},\kappa_{cl}}$	1.1 %
	S <sub>uaTc</sub>	0.4%	$S_{\kappa_{cl}^{0},\kappa_{lacl}}^{0}$	0.5 %
	total first order	40.7 %	total second order	31.2 %

degradation factor  $\gamma$  has the strongest impact on the cascade.

aTc variations have a very low impact

different importance of the basal  $\kappa_{eyfp}^0$  and regulated  $\kappa_{eyfp}$  EYFP production rates

the basal production of EYFP is due to an incomplete repression of the promoter by CI (high effect of  $\kappa_{cl}$ ) rather than a constitutive leakage of the promoter (low effect of  $\kappa_{eyfp}^0$ ).