

X ENCONTRO ACADÊMICO

MODELAGEM COMPUTACIONAL

Modelagem da Sinalização Intracelular

Anna Claudia M. Resende

Heber L. Rocha

Regina C. Almeida

annacmr@lncc.br

heberlr@lncc.br

rcca@lncc.br

Laboratório Nacional de Computação Científica (LNCC/MCTI)

Av. Getúlio Vargas 333, 25651-075, Petrópolis, Rio de Janeiro, Brazil

Ernesto A. B. F. Lima

lima@ices.utexas.edu

Institute for Computational Engineering and Sciences (ICES), The University of Texas at Austin

201 East 24th St., 78712-1229, Austin, Texas, USA

Resumo. Os processos celulares são regulados por redes integradas de comunicação intra e inter-células, constituídas por um número grande de proteínas e moléculas, que interagem bioquimicamente de modo a produzir respostas aos estímulos do microambiente. Tais respostas são locais, individuais de cada célula, dependentes da concentração intracelular de determinadas proteínas e/ou moléculas e de estímulos extracelulares. O entendimento dos processos subcelulares que regulam individualmente as transições de fenótipos, por exemplo, é fundamental para a representação mais realista da dinâmica celular e, em última instância, da progressão de várias patologias, como o câncer. Neste trabalho estudamos a rede de sinalização do fator de crescimento epidérmico (EGF), cuja alteração está associada à diversos tipos de câncer. A partir de um modelo relativamente simples, objetivamos identificar suas principais propriedades, funções e parâmetros importantes. Avaliamos possíveis simplificações e sofisticações através da análise de sensibilidade, utilizando o método dos efeitos elementares.

Palavras-chave: Modelagem Matemática, Sinalização Celular, Análise de Sensibilidade.

1 INTRODUÇÃO

O desenvolvimento crescente das pesquisas moleculares em câncer têm apontado para a importância de redes de sinalização bioquímicas em vários processos celulares. Estímulos extracelulares são internalizados através de receptores específicos da membrana celular, desencadeando cascatas de reações bioquímicas entre moléculas, denominadas redes de sinalização, que culminam na produção de uma (ou mais) resposta regulatória (Hanahan & Weinberg, 2011 e Kholodenko *et al.*, 2010). A complexidade destes mecanismos tem desafiado o desenvolvimento de alvos quimioterápicos, cujo sucesso depende fortemente do entendimento de como a expressão gênica é regulada através das sinalizações, de modo a obter terapias eficientes e menos agressivas. Em termos de modelagem, redes lógicas e Bayesianas, modelos estequiométricos, etc. são técnicas bastante utilizadas, embora a abordagem mais comum seja fundamentada na perspectiva do *continuum* via equações diferenciais ordinárias (EDOs) (Klipp *et al.*, 2016). Neste caso, cada rede é representada por um sistema acoplado e não-linear de EDOs, construído assumindo que as moléculas estão bem misturadas, de modo que a hipótese de homogeneidade espacial é razoável. Dois grandes desafios desta abordagem consistem na definição da granularidade do modelo, que corresponde ao número de equações do sistema, e na aquisição/estimação de parâmetros para a modelagem matemática e computacional.

Particularmente interessados na modelagem da sinalização que controla a proliferação celular tumoral, neste trabalho investigamos a rede de sinalização do fator de crescimento epidérmico (EGF). EGF é um mitógeno, uma substância que estimula a proliferação celular. Via um mecanismo autócrino, é liberado e usado pelas células, e passivamente transportado no microambiente tumoral, atingindo assim também células vizinhas. A sinalização do EGF é hiperativada em cerca de 30% de todos os cânceres, incluindo os de cólon, pulmão, pâncreas, dentre outros, sendo uma das principais redes de sinalizações envolvidas na progressão de tumores. Neste trabalho, utilizamos o modelo apresentado por Wang *et al.* (2007), definido por um sistema de vinte EDOs. A solução deste sistema é obtida usando o método de Runge-Kutta de quarta ordem e a identificação das moléculas-chave é obtida através da análise de sensibilidade (AS), utilizando o método dos efeitos elementares (EE) (Morris, 1991). Como a vantagem proliferativa é definida pela taxa de variação de duas moléculas específicas (PLC_{γ} e ERK), estas foram as quantidades de interesse (QoIs) utilizadas na AS. Nas seções seguintes apresentamos, de forma sucinta, o modelo utilizado, o método EE e sua utilização para a identificação dos parâmetros mais influentes nas QoIs. Finalmente, observações finais são apresentadas na seção 5.

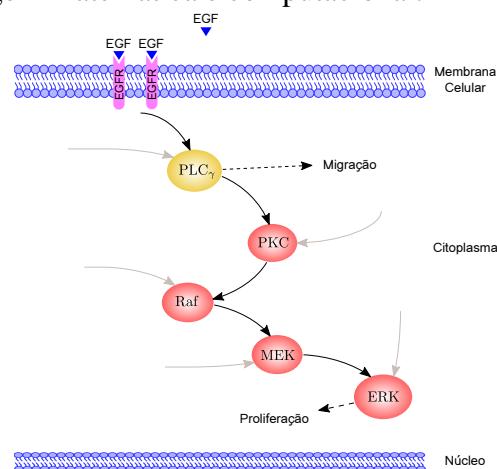


Figura 1: Representação esquemática do modelo apresentado em Wang *et al.* (2007). Os triângulos azuis indicam os fatores de crescimento (EGF), em rosa estão indicados os receptores do ligante (EGFR), em amarelo a enzima PLC_{γ} e as proteínas quinase PKC, Raf, MEK e ERK em vermelho. As setas indicam os trajetos de sinalização (em preto), as confluências com outras vias de sinalizações biológicas (em cinza) e vantagens regulatórias (tracejadas).

2 MODELO CINÉTICO

A sinalização de EGF tem papel crucial na divisão e diferenciação celular. A sinalização anormal ocorre quando há hiperexpressão de receptores, desencadeando a amplificação de moléculas-chave implicadas na divisão celular e também em outros processos regulatórios, como a migração e a apoptose celular. O modelo apresentado em Wang *et al.* (2007) é um modelo cinético formado por um sistema de EDOs que descreve a sinalização entre 20 moléculas/complexos distintos. Além do ligante EGF e do receptor de membrana EGFR, 18 outras variáveis estão interligadas na cascata de sinalização, incluídas a enzima PLC γ e as proteínas Raf, MEK e ERK. A descrição de cada molécula e associada reação cinética, que representa o balanço entre a produção e o consumo das espécies moleculares, são apresentadas nas Tabelas 1 e 2. A Fig. 1 apresenta um esquema simplificado desta cascata, estando também indicadas as confluências com outras trajetórias de sinalização (*cross-talks*). Conforme Wang *et al.* (2007), as ativações da PLC γ e da ERK estão associadas à respostas celulares específicas de migração e proliferação celular (ver Fig. 1). Assim, as taxas das variáveis PLC γ -P (X_9) e ERK-PP (X_{20}) são as quantidades de interesse da análise de sensibilidade apresentada a seguir.

3 ANÁLISE DE SENSIBILIDADE

A análise de sensibilidade é uma ferramenta importante para determinar como pequenas mudanças nos parâmetros de entrada de um modelo influenciam na variação de uma determinada quantidade de interesse (QoI) (Saltelli *et al.*, 2008). Neste trabalho, utilizando o método EE, avaliamos as taxas de variações de ERK-PP (X_{20}) e PLC γ -P (X_9) por terem um papel crucial na mudança de fenótipos celulares. Tais QoIs são representadas por uma função $Y(\boldsymbol{\theta})$, em que $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)$ é um vetor composto por d parâmetros de entrada independentes.

Considerando que todas as variáveis do espaço paramétrico são transformadas em variáveis adimensionais em um hipercubo unitário ($\boldsymbol{\theta} \in [0, 1]^d$), o EE do i -ésimo parâmetro de entrada é definido como:

$$EE_i = \frac{Y(\theta_1, \dots, \theta_i + \Delta, \dots, \theta_d) - Y(\boldsymbol{\theta})}{\Delta} = \frac{Y(\boldsymbol{\theta} + \mathbf{e}_i \Delta) - Y(\boldsymbol{\theta})}{\Delta}, \quad (1)$$

em que $\Delta \in \{1/(p-1), \dots, 1-1/(p-1)\}$ e p representa o número de níveis de discretização do hipercubo. Sendo \mathbf{e}_i um vetor de zeros com uma unidade em sua i -ésima componente, observa-se que $(\boldsymbol{\theta} + \mathbf{e}_i \Delta)$ difere de $\boldsymbol{\theta}$ apenas em sua i -ésima componente, que foi incrementada em

Tabela 1: Equações cinéticas e concentrações iniciais [nM] para cada variável molecular X_i . Para complementar, veja a Tabela 2.

Reagente	Variável Molecular	Concentração Inicial	EDO
X_1	EGF	θ_{39}	$d(X_1)/dt = -v_1$
X_2	EGFR	θ_{40}	$d(X_2)/dt = -v_1$
X_3	EGF-EGFR	-	$d(X_3)/dt = v_1 - 2v_2$
X_4	(EGF-EGFR)2	-	$d(X_4)/dt = v_2 + v_4 - v_3$
X_5	EGF-EGFR-P	-	$d(X_5)/dt = v_3 + v_7 - v_4 - v_5$
X_6	PLC γ	θ_{41}	$d(X_6)/dt = v_8 - v_5$
X_7	EGF-EGFR-PLC γ	-	$d(X_7)/dt = v_5 - v_6$
X_8	EGF-EGFR-PLC γ -P	-	$d(X_8)/dt = v_6 - v_7$
X_9	PLC γ -P	-	$d(X_9)/dt = v_7 - v_8 - v_9 - v_{10}$
X_{10}	PLC γ -P-I	-	$d(X_{10})/dt = v_9$
X_{11}	PKC	θ_{42}	$d(X_{11})/dt = -v_{10}$
X_{12}	PKC*	-	$d(X_{12})/dt = v_{10} - v_{11}$
X_{13}	Raf	θ_{43}	$d(X_{13})/dt = -v_{11}$
X_{14}	Raf*	-	$d(X_{14})/dt = v_{11} - v_{12} - v_{14}$
X_{15}	MEK	θ_{44}	$d(X_{15})/dt = v_{13} - v_{12}$
X_{16}	MEK-P	-	$d(X_{16})/dt = v_{12} + v_{15} - v_{13} - v_{14}$
X_{17}	MEK-PP	-	$d(X_{17})/dt = v_{14} - v_{15} - v_{16} - v_{18}$
X_{18}	ERK	θ_{45}	$d(X_{18})/dt = v_{17} - v_{16}$
X_{19}	ERK-P	-	$d(X_{19})/dt = v_{16} + v_{19} - v_{17} - v_{18}$
X_{20}	ERK-PP	-	$d(X_{20})/dt = v_{18} - v_{19}$

Δ (Saltelli *et al.*, 2008). Medidas de sensibilidade globais são calculadas a partir dos valores obtidos para cada EE, utilizando como estimativas a média absoluta μ_i^* e o desvio padrão σ_i definidas como:

$$\mu_i^* = \frac{1}{r} \sum_{j=1}^r |EE_i^j|; \quad \sigma_i^2 = \frac{1}{r-1} \sum_{j=1}^r (EE_i^j - \mu_i)^2, \quad (2)$$

em que r representa o número de trajetórias utilizadas para “percorrer” o espaço paramétrico. A média μ_i^* avalia o impacto do i -ésimo parâmetro sobre a QoI analisada e o desvio padrão σ_i determina a interação linear/não-linear existente entre os parâmetros e a QoI. Consideramos um total de $d = 45$ parâmetros, identificados nas Tabelas 1 e 2, cada um tendo uma distribuição uniforme contínua $\mathcal{U}(0.8\bar{\theta}_i; 1.2\bar{\theta}_i)$, sendo $\bar{\theta}_i$ o valor médio listado nas Tabelas 1 e 2. Maiores detalhes de implementação podem ser obtidos em Resende (2016).

4 EXPERIMENTOS

Na ausência de dados experimentais, assumimos que os $d = 45$ parâmetros θ_i (Tabelas 1 e 2) são independentes e estão distribuídos uniformemente segundo $\mathcal{U}(0.8\bar{\theta}_i; 1.2\bar{\theta}_i)$. O hipercubo paramétrico foi dividido em $p = 4$ níveis, com $\Delta = 1/3$, e utilizamos $r = 20$ trajetórias para o cálculo dos EEs.

A AS é apresentada nas Figs. 2 e 3 (referentes às taxas de variação do ERK-PP e do PLC_γ -P, respectivamente) para três diferentes tempos de simulação: $T_1 = 10\text{ s}$, $T_2 = 80\text{ s}$ e $T_3 = 150\text{ s}$. Para compatibilizar a comparação entre os estimadores, todos os valores obtidos para μ_i^* e σ_i foram norma-

Tabela 2: Reações (R) e parâmetros cinéticos. As concentrações e as constantes de Michaelis-Menten (M) são dadas em [nM]. Taxas constantes de primeira e segunda ordem são dadas em [s^{-1}] e [$nM^{-1}s^{-1}$], respectivamente. As velocidades V são expressas em [nMs^{-1}] (ver Wang *et al.* (2007)).

R.	Equação	Parâmetros Cinéticos			
v_1	$k_1 \cdot X_1 \cdot X_2 - k_{r1} \cdot X_3$	θ_1	$k_1 = 3.0 \times 10^{-3}$	θ_2	$k_{r1} = 0.06$
v_2	$k_2 \cdot X_3 \cdot X_3 - k_{r2} \cdot X_4$	θ_3	$k_2 = 0.01$	θ_4	$k_{r2} = 0.1$
v_3	$k_3 \cdot X_4 - k_{r3} \cdot X_5$	θ_5	$k_3 = 1.0$	θ_6	$k_{r3} = 0.01$
v_4	$V_4 \cdot X_5 / (M_4 + X_5)$	θ_7	$V_4 = 450.0$	θ_8	$M_4 = 50.0$
v_5	$k_5 \cdot X_5 \cdot X_6 - k_{r5} \cdot X_7$	θ_9	$k_5 = 0.06$	θ_{10}	$k_{r5} = 0.2$
v_6	$k_6 \cdot X_7 - k_{r6} \cdot X_8$	θ_{11}	$k_6 = 1.0$	θ_{12}	$k_{r6} = 0.05$
v_7	$k_7 \cdot X_8 - k_{r7} \cdot X_5 \cdot X_9$	θ_{13}	$k_7 = 0.3$	θ_{14}	$k_{r7} = 6.0 \times 10^{-3}$
v_8	$V_8 \cdot X_9 / (M_8 + X_9)$	θ_{15}	$V_8 = 1.0$	θ_{16}	$M_8 = 100.0$
v_9	$k_9 \cdot X_9 - k_{r9} \cdot X_{10}$	θ_{17}	$k_9 = 1.0$	θ_{18}	$k_{r9} = 0.03$
v_{10}	$k_{10} \cdot X_9 \cdot X_{11} - k_{r10} \cdot X_{12}$	θ_{19}	$k_{10} = 0.214$	θ_{20}	$k_{r10} = 5.25$
v_{11}	$V_{11} \cdot X_{12} \cdot X_{13} / (M_{11} + X_{13})$	θ_{21}	$V_{11} = 4.0$	θ_{22}	$M_{11} = 64.0$
v_{12}	$V_{12} \cdot X_{14} \cdot X_{15} / [M_{12}(1 + X_{16}/M_{14}) + X_{15}]$	θ_{23}	$V_{12} = 3.5$	θ_{24}	$M_{12} = 317.0$
v_{13}	$V_{13} \cdot X_{16} / [M_{13}(1 + X_{17}/M_{15}) + X_{16}]$	θ_{25}	$V_{13} = 0.058$	θ_{26}	$M_{13} = 2.2 \times 10^3$
v_{14}	$V_{14} \cdot X_{14} \cdot X_{16} / [M_{14}(1 + X_{15}/M_{12}) + X_{16}]$	θ_{27}	$V_{14} = 2.9$	θ_{28}	$M_{14} = 317.0$
v_{15}	$V_{15} \cdot X_{17} / [M_{15}(1 + X_{16}/M_{13}) + X_{17}]$	θ_{29}	$V_{15} = 0.058$	θ_{30}	$M_{15} = 60.0$
v_{16}	$V_{16} \cdot X_{17} \cdot X_{18} / [M_{16}(1 + X_{19}/M_{18}) + X_{18}]$	θ_{31}	$V_{16} = 9.5$	θ_{32}	$M_{16} = 1.46 \times 10^5$
v_{17}	$V_{17} \cdot X_{19} / [M_{17}(1 + X_{20}/M_{19}) + X_{19}]$	θ_{33}	$V_{17} = 0.3$	θ_{34}	$M_{17} = 160.0$
v_{18}	$V_{18} \cdot X_{17} \cdot X_{19} / [M_{18}(1 + X_{18}/M_{16}) + X_{19}]$	θ_{35}	$V_{18} = 16.0$	θ_{36}	$M_{18} = 1.46 \times 10^5$
v_{19}	$V_{19} \cdot X_{20} / [M_{19}(1 + X_{19}/M_{17}) + X_{20}]$	θ_{37}	$V_{19} = 0.27$	θ_{38}	$M_{19} = 60.0$

lizados a partir dos valores máximos de cada experimento. Estipulamos um limiar de 40% (para exemplificar) acima do qual identificamos os parâmetros θ_i que impactam a dinâmica das QoIs. Como esperado, a quantidade do estímulo extracelular (EGF) e de receptores de membrana (EGFR) são importantes na dinâmica inicial e na manutenção da resposta celular. A Fig. 2 indica que, para os tempos T_1 e T_2 , os parâmetros que mais influenciam a dinâmica da taxa de variação do ERK-P (X_{20}), em ordem decrescente de importância, são $\{\theta_9, \theta_{21}, \theta_{40}, \theta_{41}\}$ e $\{\theta_9, \theta_{41}, \theta_{39}, \theta_{21}\}$, respectivamente. Note a importância da comunicação transversa introduzida pela condição inicial de PLC_γ , que, em conjunto com as reações v_5 e v_{11} , atuam sobre a dinâmica das moléculas-chave PLC_γ e Raf.

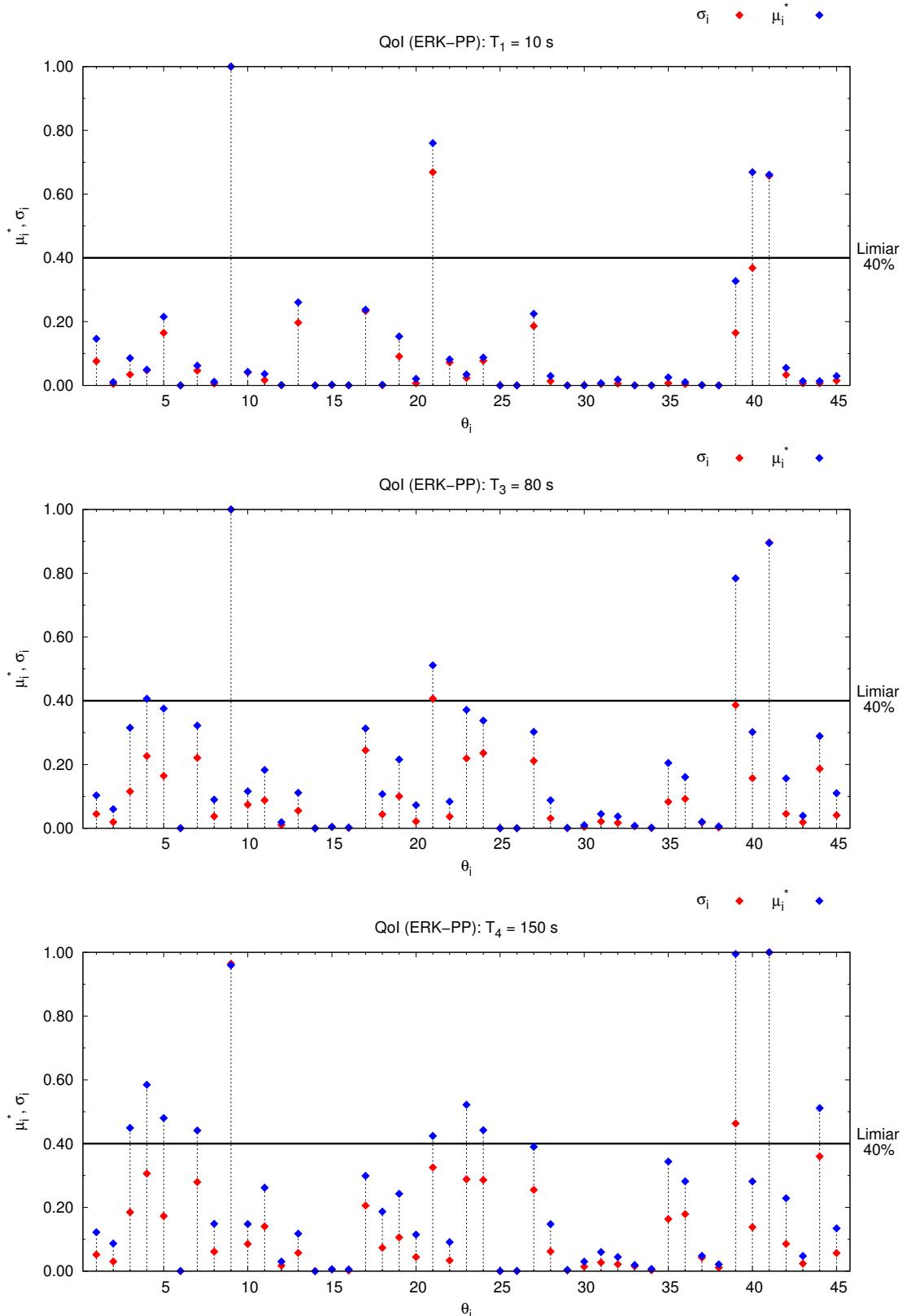


Figura 2: AS dos parâmetros com relação à taxa de variação de ERK-PP em três diferentes tempos de simulação: $T_1 = 10\text{ s}$, $T_2 = 80\text{ s}$ e $T_3 = 150\text{ s}$. Os parâmetros θ_i ($i = 1, \dots, 38$) referem-se aos parâmetros cinéticos listados na Tabela 2 e os parâmetros restantes às condições iniciais listadas na Tabela 1. As retas traçadas em $\mu_i^* = 0.4$ representam o limiar de importância estipulado para nossos experimentos (40%).

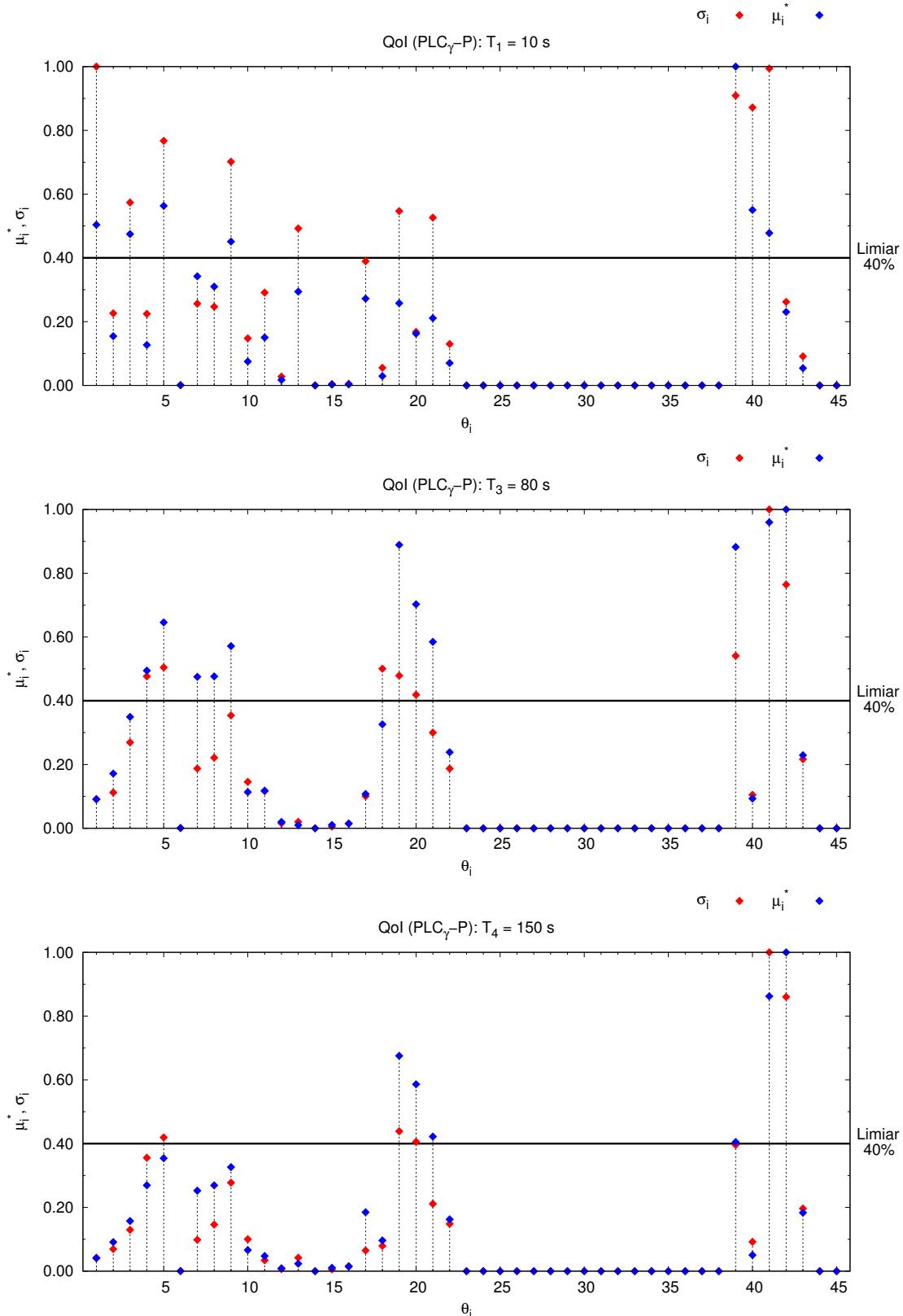


Figura 3: AS dos parâmetros com relação à taxa de variação de $\text{PLC}_\gamma\text{-P}$ em três diferentes tempos de simulação: $T_1 = 10 \text{ s}$, $T_2 = 80 \text{ s}$ e $T_3 = 150 \text{ s}$. Os parâmetros θ_i ($i = 1, \dots, 38$) referem-se aos parâmetros cinéticos listados na Tabela 2 e os parâmetros restantes às condições iniciais listadas na Tabela 1. As retas traçadas em $\mu_i^* = 0.4$ representam o limiar de importância estipulado nos experimentos (40%).

Para tempos maiores, mais moléculas passam a influenciar na sustentação do sinal regulatório. Parâmetros relacionados às reações v_2 , v_3 , v_5 , v_{11} e v_{12} tornam-se importantes, assim como a condição inicial especificada para a variável molecular MEK. Com relação à taxa de variação de PLC $_{\gamma}$ -P (X_9), observamos que inicialmente as medidas de μ_i^* indicam que os parâmetros $\{\theta_{39}, \theta_5, \theta_{40}, \theta_1, \theta_3, \theta_{41}, \theta_9\}$ são os mais influentes, destacando-se a intensidade de seus efeitos não-lineares sobre a QoI. Além disso, os parâmetros $\{\theta_{19}, \theta_{21}, \theta_{13}\}$ apresentam efeito não-linear muito mais intenso do que o linear. Este cenário altera-se bastante ao longo do tempo, embora as reações envolvendo $\{\theta_{23} - \theta_{39}\}$, que estão a jusante de PLC $_{\gamma}$ -P na cascata bioquímica, mantenham-se não influentes.

5 CONCLUSÕES

De modo geral a análise de sensibilidade realizada no modelo apresentado em Wang *et al.* (2007) confirmou a importância de alguns *cross-talks* na cascata de sinalização e o papel das denominadas moléculas-chave. Simplificações e extensões deste modelo podem ser realizadas com o intuito de auxiliar a busca de modelos de baixa granularidade para representação efetiva dos processos intracelulares de regulação do ciclo celular. Um primeiro passo para reduzir a granularidade consiste em agrupar reações que estão envolvidas num mesmo processo bioquímico e então modelar o processo agregando novas hipóteses simplificadoras. Neste passo, o modelo de Michaelis-Menten é frequentemente utilizado (ver Aldridge *et al.*, 2006 e Hornberg *et al.*, 2005). Os resultados obtidos também permitem avaliar e comparar outros modelos de vias de sinalização de EGF disponíveis na literatura (Juty *et al.*, 2015).

REFERÊNCIAS

- Aldridge, B. B., Burke, J. M., Lauffenburger, D. A., & Sorger, P. K. 2006. Physicochemical modelling of cell signalling pathways. *Nature Cell Biology*, **8(11)**, 1195.
- Hanahan, D., & Weinberg, R. A. 2011. Hallmarks of Cancer: The Next Generation. *Cell*, **144(5)**, 646–674.
- Hornberg, J. J., Bruggeman, F. J., Binder, B., Geest, C. R., de Vaate, A. J., Lankelma, J., Heinrich, Reinhart, & Westerhoff, H. V. 2005. Principles behind the multifarious control of signal transduction. ERK phosphorylation and kinase/phosphatase control. *The FEBS Journal*, **272(1)**, 244–258.
- Juty, N., Ali, R., Glont, M., Keating, S., Rodriguez, N.s, Swat, M. J., Wimalaratne, S. M., Hermjakob, H., Le Novère, N., Laibe, C., & Chelliah, V. 2015. BioModels: Content, Features, Functionality and Use. *CPT: Pharmacometrics & Systems Pharmacology*.
- Kholodenko, B. N., Hancock, J. F., & Kolch, W. 2010. Signalling ballet in space and time. *Nature Reviews: Molecular Cell Biology*, **6(11)**, 414–426.
- Klipp, E., Liebermeister, W., Wierling, C., & Kowald, A. 2016. *Systems Biology*. Wiley-VCH.
- Morris, M. D. 1991. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*, **33(2)**, 161–174.
- Resende, A. C. M. 2016. *Sensitivity Analysis as a Tool for Tumor Growth Modeling*. Dissertação de Mestrado, Laboratório Nacional de Computação Científica (LNCC/MCTI), Petrópolis/RJ, Brazil.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. 2008. *Global Sensitivity Analysis: The Primer*. John Wiley & Sons.
- Wang, Z., Zhang, L., Sagotsky, J., & Deisboeck, T. S. 2007. Simulating non-small cell lung cancer with a multiscale agent-based model. *Theoretical Biology & Medical Modelling*, **50(4)**.