# Unraveling the Complexity of MAPK Signaling Pathways

## Introduction

In the dynamic landscape of molecular biology, a pivotal and challenging task is the precise estimation of parameters that drive the behavior of intricate signaling pathways. This article delves into the heart of this challenge, focusing on the parameter estimation aspect within the context of the mitogen-activated protein kinase (MAPK) signaling pathways. These pathways, with their critical role in cellular responses, from proliferation to immunity, present an ideal backdrop for such investigations. By employing advanced techniques like gradient-based estimation and specialized tools such as PEtab and pyPESTO, this research seeks to unravel the elusive values of parameters that govern the behavior of this intricate network. Despite the hurdles encountered, including computational constraints and the intricacies of network topology, the pursuit of accurate parameter estimation using cutting-edge methods stands as a critical milestone in the field of systems biology, ultimately enhancing our understanding of the inner workings of these vital cellular processes.

The mitogen-activated protein kinases (MAPKs) are conserved signaling pathways (Figure 1a) found in organisms ranging from yeast to mammals [1]. These pathways consist of membranal and cytoplasmic signaling molecules [2], including several protein kinases that regulate critical cellular processes such as proliferation, differentiation, apoptosis, survival, inflammation, and innate immunity [1]. Among these pathways, one involving the epidermal growth factor receptor (EGFR), the tyrosine kinase receptor TrkA, and three key proteins, namely Rapidly Accelerated Fibrosarcoma (Raf), mitogen-activated protein kinase kinase (Mek or MAPKK), and Extracellular Signal-Regulated Kinase (Erk), stands out. This pathway is influenced by two stimuli, nerve growth factor (NGF) and epidermal growth factor (EGF) (Figure 1b), resulting in neuronal differentiation and cellular proliferation, respectively [3].

PC-12 cells, derived from rat adrenal pheochromocytoma, serve as a model for neuronal differentiation and have been instrumental in studying the MAPK signaling pathway [4]. The differences in how Erk is activated in response to NGF and EGF stimuli play a crucial role in shaping the cellular responses [5]. In other words, the specific timing, magnitude, and duration of Erk activation can lead to a wide array of downstream effects, as depicted in (Figure 1, b).

a. b.

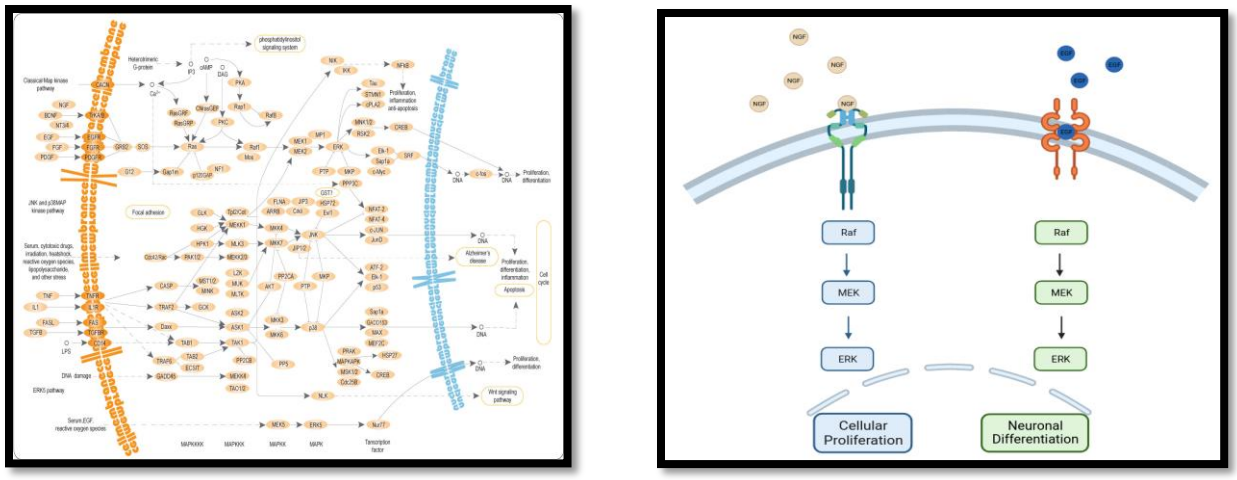


Figure 1) MAP Kinase Pathways. a) This diagram illustrates the intricate network of MAP Kinase Pathways, comprising multiple interconnected signaling pathways, each with its unique dependence on the others. The core proteins, including Paf, Mek, and Erk, play pivotal roles in most of these pathways. b) Zooming into a specific segment of the MAPK signaling pathways, we focus our attention on this select region of the complex network for our research, where key interactions and regulatory mechanisms come into play.

The central question at hand is how variations in Erk dynamics are influenced by the processes occurring upstream in the signaling pathway. The dynamics of Erk activation are primarily governed by the interconnected components within the MAPK signaling module, configuring themselves in response to different stimuli, such as EGF or NGF. The specific arrangement of these connections can give rise to distinct patterns of Erk activation dynamics [3].

To delve into the intricacies of the MAPK signaling network and understand how variations in Erk dynamics result from upstream processes, researchers have employed the Modular Response Analysis (MRA) technique [6]. This technique involves analyzing network responses under steady-state conditions (at 5 minutes, representing a pseudo-steady state, and at 15 minutes after NGF stimulation, and at 5 minutes after EGF stimulation), following incremental perturbations introduced through small RNA interference [3]. While the measured global response coefficients (Figure 2) demonstrate how perturbations propagate through the network, inferring the underlying network topology directly from these measurements is challenging. However, by

calculating local response coefficients, which indicate the sensitivity of one module to another in isolation from the network, network connectivity maps can be generated (Figure 3).

| | EGF 5 minutes | | |
|---|---|---|---|
| | Raf | Mek | Erk |
| siRNA Raf | -0.68619036 | -0.86058334 | -1.20113717 |
| siRNA Mek | -0.12842569 | -0.87060372 | -0.95215206 |
| siRNA Erk | -0.31103388 | 1.05525896 | -0.71368895 |
| siRNA Raf | -1.19055649 | -0.29297206 | -1.59356725 |
| siRNA Mek | 0.27231917 | 0.22774869 | -0.1881955 |
| siRNA Erk | -0.96319499 | 0.22435186 | -1.66223414 |
| siRNA Raf | -0.39531907 | -0.91250734 | -0.28093267 |
| siRNA Mek | -0.09767814 | -0.46534489 | -0.36504135 |
| siRNA Erk | -0.2286492 | 0.11023622 | -1.24102056 |
| siRNA Raf | -0.8325981 | 0.13465347 | -0.18431002 |
| siRNA Mek | 0.26876685 | -0.04812966 | 0.12513944 |
| siRNA Erk | -0.15890993 | 0.41577801 | -0.46901709 |

| | NGF 5 minutes | | |
|---|---|---|---|
| | Raf | Mek | Erk |
| siRNA Raf | -0.60721868 | -0.87840609 | -0.99407429 |
| siRNA Mek | -0.27618165 | -0.47643055 | -0.8503027 |
| siRNA Erk | -0.10815451 | 0.18410351 | -0.90832514 |
| siRNA Raf | -0.5260274 | 0.08170213 | 0.06569804 |
| siRNA Mek | 0.06505771 | -0.43478261 | 0.06569804 |
| siRNA Erk | -0.1146789 | 0.43036212 | -0.46835953 |
| siRNA Raf | -0.59027266 | -0.48844585 | -0.38129029 |
| siRNA Mek | -0.22257351 | -0.27600497 | -0.38350765 |
| siRNA Erk | -0.21738582 | 0.19934426 | -0.32152198 |
| siRNA Raf | -0.51637174 | -0.09862385 | -0.2380005 |
| siRNA Mek | -0.09171975 | 0.0237581 | -0.149785 |
| siRNA Erk | -0.14629509 | 0.0867747 | -0.8399211 |

| | NGF 15 minutes | | |
|---|---|---|---|
| | Raf | Mek | Erk |
| siRNA Raf | -0.50588235 | -0.74226422 | -1.15392388 |
| siRNA Mek | 0.07239819 | -0.55260435 | -0.0539592 |
| siRNA Erk | -0.03098927 | 0.23093474 | -0.77781074 |
| siRNA Raf | -0.50107066 | 0.09848485 | -0.39281576 |
| siRNA Mek | 0.08837971 | -0.86039886 | -0.62472196 |
| siRNA Erk | 0.23564955 | 0.22363765 | -0.82296651 |
| siRNA Raf | -0.83485873 | -0.34045689 | -0.16603853 |
| siRNA Mek | 0.66501487 | -0.24947735 | -0.25768709 |
| siRNA Erk | 0.34388366 | 0.59179713 | -0.16703809 |
| siRNA Raf | -0.83485873 | -0.34045689 | -0.16637872 |
| siRNA Mek | 0.66501487 | -0.24947735 | -0.25768709 |
| siRNA Erk | 0.06159032 | 0.94167117 | -1.140721 |

Figure 2) Measured Global Response Coefficients. The tables illustrate the measured global response coefficients, quantifying changes in the activity of specific modules, including Erk1, Erk2, Mek1, Mek2, and the Raf-1 and B-Raf isoenzymes, before and after perturbation. Perturbations were induced through RNA interference (RNAi), effectively downregulating protein levels. These response coefficients were obtained from quantitative western blot experiments and are sourced from the article [3].

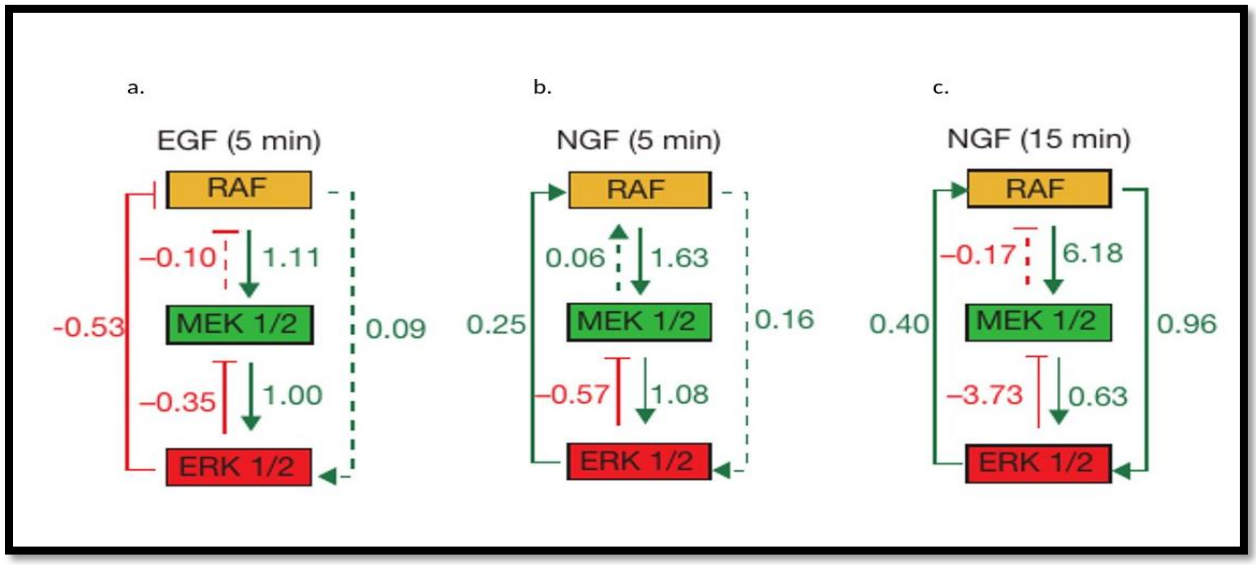


Figure 3: Wiring Diagrams of Local Response Coefficients a) Local response coefficients five minutes after stimulating the system with EGF, showcasing the network's dynamic responses under this specific condition. b) Local response coefficients five minutes after stimulating the system with NGF, representing the pseudo-steady state response. Arrows indicate coefficients, with a minus coefficient indicating an inhibitory effect, and dashed lines representing minimal effects. c) Local response coefficients fifteen minutes after stimulating the system with NGF, highlighting the network's adaptation over time. The boxes represent the proteins involved in the pathways.

In our research, we initially aimed to replicate the work[7] in the article [3], using the MRA method to estimate local response coefficients from global response coefficients obtained in experiments. Subsequently, we extended our investigation by refining the network, introducing EGF and NGF stimuli, and connecting them with the main proteins - Raf, Mek, and Erk. We then generated mathematical models for this augmented network [8]. Our next challenge was to determine unknown coefficients in these models, as well as to optimize the coefficients whose values were estimated but had confidence intervals. This was done using two different methods, both sharing a common core.

## Methods

We developed a Python program [9] to calculate local response coefficients from global response coefficients, a process rooted in the Modular Response Analysis (MRA) technique. Monte Carlo simulations [10] were then conducted for each set of coefficients under EGF 5 minutes, NGF 5 minutes, and NGF 15 minutes conditions, providing probability distributions for each coefficient [11]. During this process, we realized that MRA, with its matrix calculations (utilizing the formula: $r = -[\text{diag}(Rp^{-1})]^{-}1 * Rp^{-1}$), could be computationally intensive. To address this, we created a Python program [12] to measure computation times for different input matrix shapes.

Following a faithful replication of the original work's methods [3] to familiarize ourselves with the subject, we attempted to estimate the unknown parameters, particularly those connecting the stimulus and the Raf protein, which were not defined in the original article. To do this, we formulated five different sets of differential equations, each describing slightly different network configurations within the signaling pathway [13]. We then employed two distinct methods for parameter estimation:

## Gradient Descent Optimization Approach

We developed Python programs [14] based on the gradient descent algorithm for each set of differential equations to estimate the unknown parameters.

## PEtab and pyPESTO

Here, we leveraged specialized tools, namely PEtab [15] and the Python Parameter Estimation Toolbox (pyPESTO) [16], to perform the same parameter estimation. We preprocessed information, prepared PEtab tables [17] in the required format, and used pyPESTO to estimate the unknown parameters for each simulation [18].

## Results

Despite our diligent efforts, we encountered challenges that prevented us from obtaining the desired results. These challenges included insufficient information about the system's topology and structure, as well as a lack of raw data in a suitable format (the ppErk concentration was normalized by an unknown method). Another hurdle was the multitude of ways to formulate differential equations for such systems, resulting in uncertainty regarding which equations best describe the system. Lastly, computational limitations posed a significant challenge. The computational demands of parameter estimation can be considerable, and the constrained parameter ranges in our study hindered our ability to estimate the unknown parameters effectively.

## Conclusion

Parameter estimation methods play a pivotal role in systems biology research [19]. Choosing the appropriate method is a critical decision in this field. For future research in this area, we recommend considering the use of powerful computational servers, if feasible, to overcome computational limitations. Gaining a more comprehensive understanding of the system's structure and acquiring additional raw data, potentially in a format amenable to your analysis, can help resolve the challenges we encountered. Furthermore, exploring a wider range for each parameter could enhance the effectiveness of parameter estimation.

## References


1) Kim EK, Choi EJ. Compromised MAPK signaling in human diseases: an update. Arch Toxicol. 2015 Jun;89(6):867-82. doi: 10.1007/s00204-015-1472-2. Epub 2015 Feb 18. PMID: 25690731.

2) Seger, Rony, and Edwin G. Krebs. "The MAPK signaling cascade." *The FASEB journal* 9.9 (1995): 726-735.

3) Santos, S., Verveer, P. & Bastiaens, P. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. *Nat Cell Biol* **9**, 324–330 (2007).

4) Greene, L. A. & Tischler, A. S. Establishment of a noradrenergic clonal line of rat

adrenalpheochromocytoma cells which respond to nerve growth factor. Proc. Natl Acad. Sci.

USA 73, 2424–2428 (1976)

5) Specificity of Receptor Tyrosine Kinase Signaling: Transient versus Sustained Extracellular

Signal-Regulated Kinase Activation, C. J. Marshall. Cell, Vol, 80, 179-185, January27, 1995, Copyright© 1995by Cell Press

6) Kholodenko, B. N. et al. Untangling the wires: a strategy to trace functional interac-

tions in signaling and gene networks. Proc. Natl Acad. Sci. USA 99, 12841–12846 (2002).

7) https://github.com/LoqmanSamani/MAPK_pathway/tree/systembiology/ReplicaWork

8) https://github.com/LoqmanSamani/MAPK_pathway/blob/systembiology/differential_equations

9) https://github.com/LoqmanSamani/MAPK_pathway/tree/systembiology/ReplicaWork/LRC

10) Bonate, P.L. A Brief Introduction to Monte Carlo Simulation. *Clin Pharmacokinet* **40**, 15–22 (2001).

11) https://github.com/LoqmanSamani/MAPK_pathway/tree/systembiology/ReplicaWork/MC

12) https://github.com/LoqmanSamani/MAPK_pathway/tree/systembiology/ReplicaWork/CI

13) https://github.com/LoqmanSamani/MAPK_pathway/blob/systembiology/differential_equations

14) https://github.com/LoqmanSamani/MAPK_pathway/tree/systembiology/params_estimation

15) Schmiester L, Schälte Y, Bergmann FT, Camba T, Dudkin E, Egert J, et al. (2021) PEtab—Interoperable specification of parameter estimation problems in systems biology. PLoS Comput Biol 17(1): e1008646.

16) Leonard Schmiester, Daniel Weindl, Jan Hasenauer, Efficient gradient-based parameter estimation for dynamic models using qualitative data, *Bioinformatics*, Volume 37, Issue 23, December 2021, Pages 4493–4500.

17) https://github.com/LoqmanSamani/MAPK_pathway/tree/systembiology/petab_%26_pypesto/petab

18) https://github.com/LoqmanSamani/MAPK_pathway/tree/systembiology/petab_%26_pypesto/pypesto

19) Lillacci G, Khammash M (2010) Parameter Estimation and Model Selection in Computational Biology. PLoS Comput Biol 6(3): e1000696.