

# Of (Biological) models and simulations

## Abstract

Modeling and simulation are recognized as important aspects of the scientific method for more than 70 years but its adoption in biology has been slow. Debates on its representativeness, usefulness, and whether the effort spent on such endeavors is worthwhile, exist to this day. Here, I argue that most of learning is modeling; hence, arriving at a contradiction if models are not useful. Representing biological systems through mathematical models can be difficult but the modeling procedure is a process in itself that follows a semi-formal set of rules. Although seldom reported, failure in modeling is not a rare event but I argue that this is usually a result of erroneous underlying knowledge or mis-application of a model beyond its intended purpose. I argue that in many biological studies, simulation is the only experimental tool. In others, simulation is a means of reducing possible combinations of experimental work; thereby, presenting an economical case for simulation; thus, worthwhile to engage in this endeavor. The representativeness of simulation depends on the validation, verification, assumptions, and limitations of the underlying model. This will be illustrated using the inter-relationship between population, samples, probability theory, and statistics.

**Keywords:** model, simulation, simulacrum, philosophy, biological modeling and simulation

Volume 3 Issue 4 - 2016

Maurice HT Ling<sup>1,2</sup>

<sup>1</sup>Colossus Technologies LLP, Republic of Singapore

<sup>2</sup>School of BioSciences, the University of Melbourne, Australia

**Correspondence:** Maurice HT Ling, Colossus Technologies LLP, 8 Burns Road, Trivex, Singapore, 369977, Republic of Singapore, Tel +6596669233, Email mauriceling@colossus-tech.com

**Received:** April 08, 2016 | **Published:** May 03, 2016

## Models and simulations

All models are wrong, but some are useful.

Box and Draper.<sup>1</sup>

All models are right... Most are useless.

Tarpey<sup>2</sup>

The role of models has been recognized as an important aspect to the scientific method<sup>3</sup> for more than 70 years.<sup>4</sup> In recent years, there are increasing interests in metabolic modeling in biology.<sup>5,6</sup> The process of building a model is known as modeling. Generally speaking, models can be classified as either material models or formal models. Rosenblueth & Wiener<sup>7</sup> consider material models to be a physical representation of the object under study whereas formal models are symbolic or logical representations of the object under study. For example, a scale down model of an airplane for testing its aerodynamic properties in a wind tunnel is a material model. The main purpose of a scaled-down material model is the ease of study, which had been clearly demonstrated by a study on fireplace more than 200 years ago.<sup>8</sup>

A formal model, on the other hand, is an abstract but unambiguous description of a physical object or phenomenon, usually in logical or mathematical constructs. Commonly used mathematical constructs for building formal models are rate equations, such as differential equations,<sup>9</sup> or state transitions, such as petri nets.<sup>10</sup> Other modeling formalisms had been reviewed by Machado et al.<sup>11</sup> Besides differentiating models by their underlying formalism, models can also be classified by their informational complexity. At the lowest level, the presence or absence of associations between 2 entities (such as cells or molecules) can be represented as a Boolean term. This results in a binary state petri net or Boolean network, which can be used to study cycles and reachability. For example, a metabolic pathway or gene regulation network is essentially a Boolean network.<sup>12</sup> At the next level, the stoichiometries between each state can be added,<sup>13</sup>

resulting in a steady state model, which can be simulated and used to study the state or concentrations of various molecules at homeostasis using methods such as flux balance analysis.<sup>14-16</sup> At the highest level, the transitions between each state can be represented using a set of differential equations and this result in a kinetic model.<sup>17</sup> The main advantage of kinetic models over steady state model is the ability to simulate each state (or molecules in the context of metabolic modeling) across time, in addition to the ability to model homeostasis.<sup>12</sup> This will allow for identification of bottlenecks in the metabolism.<sup>18</sup>

With advancements in computing power, formal models, being constructed in mathematical or logical format, render it possible for computation. Such computation or execution of formal models is commonly known as simulation. Hence, by definition, one can “build a model”, then “run or execute a model” or “run or execute a simulation” but in no means, one can “build a simulation”. The ability to render execution of a model; thus, an executable or simulatable model; defines the fundamental difference between a material model and a formal model. In addition, there are two kinds of modeling strategy in the current literature<sup>19</sup> forward and reverse modeling. Reverse modeling starts from experimental data and seeks potential causalities suggested by the correlations in the data, captured in the structure of a mathematical model. Forward modeling starts from known, or suspected, causalities, expressed in the form of a model, from which predictions are made about what to expect.

Thus, simulation can be seen as the utilization of a formal model. Very commonly, the purpose of building a formal model is to simulate. As such, in this article, I will use the term “model building” to specifically refer to the process of building a formal model, “simulation” to specifically refer to the process of simulating a formal model, and “modeling” to refer to the both the process of model building and simulation.

The most important purpose of modeling is to gain insights into behaviors that are not obvious from the model itself<sup>20</sup> by reading the mathematical equations. In spite of this, there is an ongoing debate

on the role of models and simulations in biology and science at large, even to this day. The major arguments against modeling in biology can be summarized in three questions:

- i. Is a model/simulation useful?<sup>3</sup>
- ii. Is a model/simulation representative?<sup>21</sup>
- iii. Is it worth spending the time and effort engaging in modeling?<sup>22</sup>

For the rest of this article, I will attempt to answer these questions and take the view that modeling is both useful and can be representative; hence, it is worth the time and effort to engage in this endeavor.

## All models are wrong but some are useful

Are models useful? In order to answer this question, we have to ask, what are the uses of models? Thus, if models exhibit its required utility, models are; therefore, useful. To start off, I argue that models are an integral part of learning. Learning from models and learning by models are collectively known as model-facilitated or model-based learning, which is the formation of mental models of knowledge by the learner.<sup>23,24</sup> In a classical monograph titled “What the Best College Teachers Do”,<sup>25</sup> Ken Bain pointed out that knowledge from teacher to student is not transferred but built in the student and the process of doing so is likened to assisting the student to build a mental model of the knowledge, followed by reasoning using the built mental model. This is then followed by critically examination of the model and the role of the teacher is to bring the student’s mental model to a point of “expectation error”, where the mental model of knowledge breaks down, which is analogous to finding the limitations of models. As Bain<sup>25</sup> puts it:

“The best college and university teachers create what we might call a natural critical learning environment in which they embed the skills and information they wish to teach in assignments (questions and tasks) students will find fascinating – authentic tasks that will arouse curiosity, challenge students to rethink their assumptions and examine their mental models of reality”.

Here is a personal experience when I was trying to learn about DNA replication. As DNA is double stranded and replication can only occur from a 5’ to 3’ direction, it means that one strand of the DNA (known as the lagging strand) has to be “copied backwards” from a seemingly 3’ to 5’ direction, which is a contradiction. Work by Sakabe & Okasaki<sup>26</sup> showed that the lagging strand is actually copied in short fragments in a 5” to 3” direction, which eventually came to be known as Okazaki fragments. The primers of Okazaki fragments are subsequently removed by a combination of enzymes. However, I could not understand why the enzymes did not zip pass and remove the entire downstream Okasaki fragment instead. My instructor realized this contradiction and suggested that the chemical structure of ribose (in the primer) and deoxyribose (in the DNA section of Okasaki fragment) may play a role.<sup>27</sup> Hence, my mental model of DNA replication was correct but incomplete as it did not take into account of the chemical structure between ribose and deoxyribose, which lead to an expectation error (why the entire downstream Okasaki fragment was not removed?). This led me to modify my mental model.

The advancement of science depends on constant challenging and re-examination of perceived notions and assumptions, and to incorporate new information from latest research work into existing models. In the process of doing so, mental models of knowledge are refined and used as basis for extension, making model a powerful

tool for learning scientific knowledge and reasoning.<sup>28</sup> Despite the advantages of using models in learning,<sup>25,28</sup> there are perceived notions that modeling may not be useful as models may be too specific and difficult to generalize.<sup>20,29</sup> These notions may be indirectly refuted by recent efforts promoting the usefulness of modeling in biology<sup>30,31</sup> and by a large volume of studies in other fields.<sup>32</sup> Underpinning the usefulness of modeling is the balance between specificity and generalization.

I will further argue that model building is perennial in all fields of science on the basis that the simplest model, a single mathematical equation, is a codified model between the inputs and outputs. For example, Newton’s Second Law of Motion states that force is proportional to the product of mass and acceleration, is a formal model between force, mass, and acceleration. Similarly, in microbiology, Monod’s equation is a century old equation which formalizes the relationship biomass to the concentration of limiting substrate, and is still in use in many studies today.<sup>33,34</sup>

Hence, models are meant to be specific - the intent of a model is for one and only one purpose. For instance, a scale model of an airplane for aerodynamic testing in a wind tunnel will not be applicable for propulsion testing. If the same scale model can be used for another purpose not in its original specification or intent, it is a bonus and not a requirement. Conversely, if the same model cannot be used for another purpose other than its original intent, it is expected rather than a criticism. Thus, all models are wrong except when used for its intended purpose. This is also related to the choice of modeling techniques. For example, if the requirement is to identify bottlenecks in redox reactions, then it is necessary to build a kinetic model.

To push this argument further, it can also be said that a model need not be correct in all aspects. In fact, the areas in which the model is wrong may be generalized as the limitations of the model. For example, Mendelian genetics is a model of inheritance which had been used for more than a century as the basis for studying inherited disorders.<sup>35</sup> Despite so, instances of inheritance that fail to follow Mendelian genetics are widely known.<sup>36</sup> Yet, this does not mean that either model is wrong. In fact, both models are correct in a limited way. Thus, all models are wrong except when used for its intended purpose and with limitations. Indeed, even the seemingly ubiquitous Monod’s equation<sup>37,38</sup> in microbiology has its own limitations. The same can be said for Newtonian mechanics, which is known to be irrelevant when the speed of the object is near light speed.

In fact, a model need not be highly accurate to be useful. For example, Karr et al.,<sup>39</sup> constructed a whole cell metabolic model of *Mycoplasma genitalium*. Although after modeling all 525 genes and its metabolism, Karr et al.,<sup>39</sup> managed to achieve a correlation of 0.82 (reported as R<sup>2</sup> of 0.68) between observed experiments and simulated results. In the strictest sense, Karr et al.,<sup>39</sup> model can be deemed to be “having room for improvements” as the correlation is not near perfect (that is, R<sup>2</sup>>0.95) but this did not prevent the model to be used to gain insights into various metabolic processes of *M. genitalium*. This suggests that a model only needs to be sufficiently accurate or correct in some aspects but not all. This work<sup>39</sup> had even been commented in the same issue of the journal as the “dawn of virtual cell biology”.<sup>40</sup> Such large scale models had been attempted in other organisms and had provided insights<sup>41</sup> even for industrial applications.<sup>42</sup> Larger scales of models of various single cell models had been attempted for studying the emergence of microbial ecology<sup>43</sup> and it is not expected for these models to be entirely highly to be useful. Hence, there is no practical reason to be hung up on high accuracy of models, especially when

it is generally known that laboratory experiments may have varied degrees of reproducibility due to inaccuracies of instruments (such as, imprecision of micropipetters) and skills of the researcher (such as, not adhering to protocols down to the seconds). In many cases, the purpose of the models is to provide insights, which have to be experimentally validated, but nevertheless drive research directions.

Albert Einstein is reported to have said to the effect that *“no amount of experimentation can ever prove me right; a single experiment can prove me wrong”*.<sup>44</sup> Science should be worried about models without limitations, and by extension, worry about models that are always right. This will only imply that such models have not been carefully examined. Thus, all models have to be wrong in some ways to render its greatest use.

At the end of the day, science had always been in the business of proving models to be wrong and that is the basis of hypothesis testing. Personally, I had thrown away more models than I am willing to count. Like a tree today is sitting on a mass grave of past plants and animals, the models that I have today are standing on a graveyard of models. Yet, all these dead or failed models, which are common place and un-publishable, taught me something about my own understanding of the field and more importantly, highlighting my mis-understanding of the field.

### Modeling, the only experimental tool

In the most basic form, a model can be seen as the box or system between the input or stimuli, and the output or response.<sup>45</sup> The main goal of science is to understand and explain how the response(s) is/are a result of the input(s).<sup>46</sup> Taking Monod's equation as an example, the concentration of limiting substrate is the input to calculate the biomass, and the equation itself explains the logical relationship between the concentration of limiting substrate and the biomass.

This implies that there are 3 uses of models.<sup>45</sup> Firstly, by knowing the input and the model, the output or response can be predicted. Secondly, by knowing the model and the required response, the input can be controlled. Lastly, by knowing the input and response, the system can be understood.

However, these 3 uses are not independent. Very often, a model has to be relatively understood before it can be used for prediction or control. If a model is not substantially understood, it often gives wrong predictions and the model has to be refined and improved upon. Scientific experiments can then be seen as a procedure to generate a set of corresponding input-output pairs on which the underlying phenomenon can be understood; thus, modeled.

Although experimentation in biology is common (as evident by the large volume of experimental publications), laboratory experimentation is both difficult and unethical in many areas of biology,<sup>47</sup> such as epidemiology and evolution. Real-world epidemiological experiments will require the willful dissemination of infectious agents into the general public as the input to track the speed and routes of transmission as the output. Clearly, this is not ethical and should never be allowed. Hence, most of epidemiology is the collection of data on existing infections from the general public<sup>48</sup> for the purpose of modeling the mode of infection<sup>49</sup> using methods such as curve fitting techniques.<sup>50</sup> Only then, can outcomes of disease transmission be studied and tested in a virtual context before using it for forecasting. Similarly, studying evolution experimentally is difficult and expensive.<sup>51</sup> Although some experimental evolution studies have been done,<sup>52-55</sup> they are restricted to evolution of fast-

growing bacterial cells. It is impossible to study human evolution in an experimental setting, both in terms of ethics and in duration.

In these ethically and/or practically impossible situations, modeling appears to be the only viable experimental tool. Computer simulations of virtual organisms (commonly known as “digital organisms” or “artificial life”) had been used instead (reviewed in)<sup>56</sup> and may provide some insights into human evolution.<sup>57</sup> Moreover, there are areas that intersect between evolution and epidemiology, such as the evolution of antibiotics resistance. There have been contradictory studies showing that antibiotics resistance can possibly revert once the specific antibiotic is dis-used.<sup>58</sup> Although antibiotics resistance had been studied experimentally,<sup>59</sup> it is only performed in controlled laboratory settings as it is unacceptable to willfully induce one or more antibiotics resistance in the human or animal population at large. Hence, modeling of virtual organisms has been used as an experimental tool in attempt to break the above contradiction and find that contradictory results are caused by statistical variations in the de-evolution of antibiotics resistance as the 95% confidence interval can vary from reversion to non-reversion of antibiotics resistance.<sup>[58,60]</sup>

Hence, the question of whether models are useful appears to be rather muted in situations whereby models are the only feasible instruments of study.

### Modeling, the economic tool

It is without a doubt that biology is complex and there can be a lack of thorough understanding in underlying biology or the lack of representative data but Gunawardena<sup>19</sup> argues that these limitations do not warrant the call for not modeling. I will further argue that such limitations, especially the lack of representative data, may be what render modeling useful – to build an initial model using available data, and using the model to uncover what data is needed to improve and refine the model. Using the analogy of learning a new language - nobody attempt to learn a new language by memorizing all the grammatical rules and dictionary of words prior to constructing a simple greeting. Recently, Chowdhury et al.,<sup>61</sup> had demonstrated the use of models as a means to consolidate on-going research results.

Modern science is not just complex but expensive to study experimentally. Multi-billion dollar scientific equipment, such as the Hubble telescope and Large Hadron Collider, had been built. Biology is of no exception as it can often take millions of dollars to build a modern biology laboratory. This has been suggested to be widening a disparity between wealth of different research groups or even countries.<sup>47</sup> Modeling may present itself as an economic leveling tool in this aspect, as well as a tool towards efficient use of research dollars.

There have been a number of recent metabolic engineering studies demonstrating the use of modeling to reduce the number of experiments needed<sup>62</sup> to optimize specific metabolite production.<sup>63,64</sup> Thus, resulting in more efficient use of research dollars as a result of higher research and development throughput and is likely to result in shorter time between research and industrial production.

Machado et al.,<sup>63</sup> modeled a 4-enzymatic step for curcumin production from phenylalanine and/or tyrosine, and carried out *in silico* modeling experiments to isolate mutants with potential for high curcumin production. Of the possible enzyme over-expressions, Machado et al.,<sup>63</sup> suggested that over-expression of acetyl-coA carboxylase, which catalyzes the acetyl-coA to malonyl-coA, might have the biggest impact in terms of increased curcumin production.

This result is un-expected as acetyl-coA carboxylase is not the first enzyme in the sequence that utilizes phenylalanine or tyrosine, nor the enzyme directly responsible to curcumin synthesis; hence, suggesting a non-intuitive result.

Weaver et al.,<sup>64</sup> modeled a 6-enzyme pathway from mevalonate to amorphadiene with the main of modifying enzyme kinetics, such as turnover number and Michaelis-Menten constant of the enzymes as well as concentrations of enzymes, in order to maximize amorphadiene production. Hence, Weaver et al.,<sup>64</sup> went one step further compared to Machado et al.,<sup>63</sup> as Weaver et al.,<sup>64</sup> is interested to which property of the enzyme (turnover number and/or Michaelis-Menten constant) is more important. Through sensitivity analysis, Weaver et al.,<sup>64</sup> found that the turnover number and expression level of amorphadiene synthase, which catalyzes the last step to product amorphadiene, are important in the final production rate of amorphadiene. However, the Michaelis-Menten constant of amorphadiene synthase has negligible impact on amorphadiene production rate. Weaver et al.,<sup>64</sup> validated this modeling findings experimentally.

These recent studies<sup>63,64</sup> demonstrate the use of modeling to reduce the number of optimization trials to run in order to get a potentially economically bacterial strain for industrial fermentation. They also<sup>63,64</sup> further underline a usefulness of modeling as a window to access the system on hand.<sup>65</sup> This view is supported by Lung,<sup>66</sup> whom suggests that models are useful as “*models can provide insight into the behavior of complex systems and sometimes yield results which are different from the intuitive predictions.*” Taken together, modeling may be a potential tool for improved stewardship of limited research funding.

## Simulacrum

Recently, there are also reported works using modeling as a tool for clinical settings,<sup>67,68</sup> suggesting that efforts spent to learn about modeling is worthwhile.<sup>67</sup> This brings us to the last issue – what makes a model representative and by extension, how interpretable are the simulation results? These appear to be new philosophical questions arising from modeling but Frigg & Reiss<sup>69</sup> argue that these seemingly new philosophical questions have analogies in experimentation and thought experiments; hence, they are not new questions but a rehashing of existing questions.

I am inclined towards Frigg & Reiss.<sup>69</sup> In the field of biology, the term “model” has been used in many different contexts to refer to an experimentation platform (such as model organism) or an abstraction (such as fluid-mosaic model of cell membrane) and the question of representativeness had been addressed according; for example, “is zebrafish a good model for developmental biology?”<sup>70</sup> Unless we want to use humans as model organism for human development, we are not left with much choice but to use appropriate experimental models; which balance practicality, scientific value, and ethics; as well as acknowledging the limitations of such models. However, this may also be a fatalistic point of view.

Another way to view this is using the relationship between population and sample as an analogy. Here, population refers to the object of interest but is practically impossible to study in its entirety. On the other hand, a sample is a reduced population or a representative population, which is analogous to a model and is feasible for study. Then, what makes a sample a representative population? The answer will be “probability theory”-every reasonable feature in the population has to be proportionally represented in the sample. It is important to note that it may not be possible to represent every single feature from the population onto the sample. Very often, features

of low probabilities get left out and this has to do with sample size. For example, if the side-effect of a drug in question happens 1 in 100thousand, it is unlikely to be represented in a sample of 10thousand patients administered the drug. This forms the inherent limitation of the sample, and by extension, the model. After analyzing the sample, the analysis results have to be interpreted in the context of the population. This is the work of “statistical theory”. Within this interpretation, the sampling limitations must be reflected.

Here, I argue that the inter-relationship between population, samples, probability theory, and statistics, forms the basis of simulacrum. At the fundamental level, the simile to probability theory and statistics, when applied to modeling, are verification and validation, respectively.

In layman’s terms, verification and validation can be asked in the following questions:

- i. For verification, “are we building the model right?” That is, is the model free from implementation errors and/or logical errors?<sup>71</sup>
- ii. For validation, “are we building the right model?” That is, is the model addressing the needs and requirements?<sup>71</sup>

In essence, verification is concerned with internal or within-model “correctness” while validation is concerned with external “correctness”.<sup>72</sup>

Using the work of Massoud et al.,<sup>73</sup> verification can be performed at 2 stages: model building stage, and internal verification of the model. In building stage, the real-world system (analogous to population) is formulated into a model (analogous to sample) via inductive generalization, also known as semantic link. This calls for the identification of a set of minimum parameters used to describe the real world system and to form associations based on observations in the real world. This results in a specific model, which has to be verified internally. In this stage (model verification), the main purpose is to ensure that the semantic associations between each parameter is logically meaningful based on the current knowledge in the field. By Occam’s Razor, simple associations are preferred over complex associations, and a set of small number of associations is preferred over a larger set.

Once all associations are verified, testing of the model is performed for accuracy by matching the model’s behaviors with the observations in which the model is built from. This step is mandatory as it determines the degree in which the model is an accurate representation of the system under study.<sup>3,20</sup> It is common in mathematical models that stochastic variations arise from imprecisions in floating-point calculations, which are commonly known as rounding-off errors. Hence, it may be important to consider whether such errors are within acceptable range and whether a change of scales may reduce such errors. For example, 1% error in molarity represents 10milli-moles whereas 1% error in milli-molar represents 10 micro-moles. In addition, the importance of each parameter can be further estimated using sensitivity analysis<sup>74</sup> and it may be possible to eliminate parameters of low importance for the purpose of simplifying the model. Several iterations of model refinements and testing are usually carried out to produce an ‘adequate’ model that is verified with acceptable accuracy. The model can then be said to be representative of the system under study.

Only then, can the model be validated. Validation can be done in primarily 2 ways, as described by Hicks et al.,<sup>75</sup> namely, (1) comparing

simulation results with additional observations, and (2) comparing simulation results with that of other validated models.

When comparing simulation results with additional observations, it is important to note that these observations cannot be used for model building; implying that a set of new observations must be used for validity testing. A commonly used technique is known as cross validation,<sup>76</sup> also known as rotational estimation. In cross validation, the entire set of observations is randomly divided into 2 or more equal sub-sets where 1 sub-set is reserved for validation testing and the remaining sub-sets are used for model building. The error or accuracy of the model is then reported as average (mean of means) and standard error.

In event where other validated models are available, it is also common to compare results between 2 or more models. However, there are 2 common issues with this approach. The parameters used in other models may be different from the set of parameters used in the current model, which may be a result of different assumptions and considerations during model building. The purpose of the models may also be different. This is likely to result in differences in simulation outputs as comparison should be made only when 2 models uses the same set of parameters. The differences in the simulation output will be confounded with the difference in parameters used, giving another obstacle in interpretation.

Besides the above 2 methods, Sargent<sup>71</sup> encourages the elucidation of operational validity to determine whether the model's output has the accuracy required for its intended purpose. A major aspect of operational validity is accuracy, which has been dealt with in the above verification and validation. The glaring aspect that had not been addressed is the limitations of the model, which can be defined as the boundaries of the applicability of the model. In this case, extreme parameter conditions have to be tested to estimate the conditions in which the model fails; thus, setting operational boundaries for the model. Within these operational boundaries, the model is able to give insights into the system under study (analogous to statistics providing insights into the population).

## Conclusion

Modeling has been an important aspect of science<sup>71,77</sup> but there is an ongoing debate on the role of modeling in biology and science at large, even to this day. In this article, I argue that modeling can be useful tool<sup>72,78</sup> to aid in many areas of biological studies. The use of modeling extends beyond being a research tool but may be an economical tool to reduce the number of experimental trials needed, which may lead to a better use of limited research funding. Thus, it is worth spending effort to learn and incorporate modeling into a biologist toolkit.

## Acknowledgements

None.

## Conflict of interest

The author declares no conflict of interest.

## References

- George EPB, Norman RD. Empirical Model Building and Response Surfaces. *Technometrics*. 1988;30(2):229–231.
- Thaddeus T. All Models are Right...Most are Useless. *JSM Proceedings*. 2009.
- Zoltan D. Mathematical Models in Philosophy of Science. 2nd ed. *International Encyclopedia of the Social & Behavioral Sciences*. USA; 2015. p. 791–799.
- Gustav B. On Physicalistic Models of Non-Physical Terms. *Philosophy of Science*. 1940;7(2):151–158.
- Jouhten P. Metabolic Modelling in the Development of Cell Factories by Synthetic Biology. *Comput Struct Biotechnol J*. 2012;3:e201210009.
- Ray D, Ye P. Characterization of the Metabolic Requirements in Yeast Meiosis. *PLoS One*. 2013;8(5):e63707.
- Rosenblueth Ar. The Role of Models in Science. *Philosophy of Science*. 1945;12(4):316–321.
- Peale C, Peale R. *Description of Some Improvements in the Common Fire-Place, Accompanied with Models, Offered to the Consideration of the American Philosophical Society*. USA: Transactions of the American Philosophical Society; 1802. p. 320–324.
- Van der Vaart HR. A Comparative Investigation of Certain Difference Equations and Related Differential Equations: Implications for Model-Building. *Bulletin of Mathematical Biology*. 1973;35:195–211.
- Lee DY, Zimmer R, Lee SY, et al. Colored Petri Net Modeling and Simulation of Signal Transduction Pathways. *Metab Eng*. 2006;8(2):112–122.
- Machado D, Costa RS, Rocha M, et al. Modeling Formalisms in Systems Biology. *AMB Express*. 2011;1:45–45.
- Samal A, Jain S. The Regulatory Network of E. coli Metabolism as a Boolean Dynamical System Exhibits Both Homeostasis and Flexibility of Response. *BMC Systems Biol*. 2008;2:21.
- Baldan P, Cocco N, Marin A, et al. Petri Nets for modeling metabolic pathways: a survey. *Nat Comput*. 2010;9:955–989.
- Kauffman KJ, Prakash P, Edwards JS. Advances in Flux Balance Analysis. *Curr Opin Biotechnol*. 2003;14:491–496.
- Kim J, Fabris M, Kim MK, et al. Flux Balance Analysis of Primary Metabolism in the Diatom *Phaeodactylum tricornutum*. *Plant J*. 2016;85(1):161–176.
- Matsuoka Y, Shimizu K. Metabolic Flux Analysis for *Escherichia coli* by Flux Balance Analysis. *Methods Mol Biol*. 2014;1191:237–260.
- Smallbone K, Simeonidis E, Swainston N, et al. Towards a Genome-Scale Kinetic Model of Cellular Metabolism. *BMC Systems Biology*. 2010;4:6.
- Mannan AA, Toya Y, Shimizu K, et al. Integrating Kinetic Model of E. coli with Genome Scale Metabolic Fluxes Overcomes Its Open System Problem and Reveals Bistability in Central Metabolism. *PLoS One*. 2015;10(10):e0139507.
- Gunawardena J. Models in Biology: 'Accurate Descriptions of Our Pathetic Thinking'. *BMC Biol*. 2014;12:29–29.
- Peck SL. Simulation as Experiment: A Philosophical Reassessment for Biological Modeling. *Trends Ecol Evol*. 2004;19(10):530–534.
- Rice CC. Factive Scientific Understanding without Accurate Representation. *Biology & Philosophy*. 2015;31(1):81–102.
- Petrinco MI, Pagano E, Desideri A, et al. Information on Center Characteristics as Costs' Determinants in Multicenter Clinical Trials: Is Modeling Center Effect Worth the Effort? *Value Health*. 2009;12(2):325–330.
- Spector MJ. Model-Facilitated Learning. *Encyclopedia of the Science of Learning*. 2012:2316–2317.
- Buckley BC. Model-Based Learning. *Encyclopedia of the Science of Learning*. 2012:2300–2303.

25. Bain K. *What the Best College Teachers Do*. England: Harvard University Press; 2004.
26. Sakabe K, Okazaki R. A Unique Property of the Replicating Region of Chromosomal DNA. *Biochim Biophys Acta*. 1966;129(3):651–654.
27. Zheng L, Shen B. Okazaki Fragment Maturation: Nucleases Take Centre Stage. *Journal of Molecular Cell Biology*. 2011;3:23–30.
28. Quillin K, Thomas S. Drawing-to-Learn: A Framework for Using Drawings to Promote Model-Based Reasoning in Biology. *CBE Life Sci Educ*. 2015;14(1):es2.
29. Gavrilits S. Models of Speciation: Where Are We Now. *J Hered*. 2014;105(Suppl 1):743–755.
30. Maeng HJ, Chow ECY, Fan J, et al. Physiologically Based Pharmacokinetic (PbPK) Modeling: Usefulness and Applications. *Encyclopedia of drug metabolism and interactions*. 2012.
31. Wang P, Zhu BT. Usefulness of Molecular Modeling in Characterizing the Ligand-Binding Sites of Proteins: Experience with Human Pdi, Pdp1 and Cox. *Curr Med Chem*. 2013;20(31):3840–3854.
32. Bishop MA, Trout J. 50years of successful predictive modeling should be enough: Lessons for philosophy of science. *Philosophy of Science*. 2002;69:S197–S208.
33. Koutinas M, Kiparissides A, Silva-Rocha R, et al. Linking genes to microbial growth kinetics: an integrated biochemical systems engineering approach. *Metab Eng*. 2011;13(54):401–413.
34. Garnier A, Gaillet B. Analytical solution of Luedeking-Piret equation for a batch fermentation obeying Monod growth kinetics. *Biotechnol Bioeng*. 2015;112(12):2468–2474.
35. Paciolla M, Pescatore A, Conte MI, et al. Rare mendelian primary immunodeficiency diseases associated with impaired NF- $\kappa$ B signaling. *Genes Immun*. 2015;16(4):239–246.
36. Cuzin F, Rassoulzadegan M. Non-Mendelian epigenetic heredity: gametic RNAs as epigenetic regulators and transgenerational signals. *Essays Biochem*. 2010;48(1):101–106.
37. Kovárová-Kovar K, Egli T. Growth Kinetics of Suspended Microbial Cells: From Single-Substrate-Controlled Growth to Mixed-Substrate Kinetics. *Microbiol Mol Biol Rev*. 1998;62(3):646–666.
38. Regaudie-de-Gioux A, Sal S, López-Urrutia A. Poor correlation between phytoplankton community growth rates and nutrient concentration in the sea. *Biogeosciences*. 2015;12:1915–1923.
39. Karr JR, Sanghvi JC, Macklin DN, et al. A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*. 2012;150(2):389–401.
40. Freddolino PL, Tavazoie S. The Dawn of Virtual Cell Biology. *Cell*. 2012;150(2):248–250.
41. Chew YH, Wenden B, Flis A, et al. Multiscale Digital Arabidopsis Predicts Individual Organ and Whole-Organism Growth. *Proc Natl Acad Sci U S A*. 2014;111(39):4127–4136.
42. Simeonidis E, Price ND. Genome-Scale Modeling for Metabolic Engineering. *J Ind Microbiol Biotechnol*. 2015;42(3):327–338.
43. Zomorodi AR, Segrè D. Synthetic Ecology of Microbes: Mathematical Models and Applications. *J Mol Biol*. 2016;428(5):837–861.
44. Wynn CM, Wiggins AW, Harris S, et al. *The Five Biggest Ideas in Science*. USA: John Wiley and Sons; 1997.
45. Haefner JW. *Modeling Biological Systems: Principles and Applications*. Springer Science & Business Media; 2005.
46. Friedman M. Explanation and Scientific Understanding. *The Journal of Philosophy*. 1974;71(1):5–19.
47. Ling MHT. The bioinformaticist's/computational biologist's laboratory. *MOJ Proteomics & Bioinform*. 2016;3(1):00075.
48. Buring JE. Primary Data Collection: What Should Well-Trained Epidemiology Doctoral Students be Able to Do? *Epidemiology*. 2008;19(2):347–349.
49. Bellan SE, Pulliam JR, Scott JC, et al. How to Make Epidemiological Training Infectious. *PLoS Biology*. 2012;10(4):e1001295.
50. Chernov N. *Circular and Linear Regression: Fitting Circles and Lines by Least Squares*. Chapman & Hall/CRC Monographs on Statistics and Applied Probability, 2010:117.
51. Batut B, Parsons DP, Fischer S, et al. In Silico Experimental Evolution: A Tool to Test Evolutionary Scenarios. *BMC Bioinformatics*. 2013;14(Suppl 15):S11.
52. Lee CH, Oon JSH, Lee KC, et al. Escherichia coli ATCC 8739 Adapts to the Presence of Sodium Chloride, Monosodium Glutamate, and Benzoic Acid After Extended Culture. *ISRN Microbiology*. 2012;2012:965356.
53. Goh DJW, How JA, Lim JZR, et al. Gradual and Step-Wise Halophilization Enables Escherichia coli ATCC 8739 to Adapt to 11% NaCl. *Electronic Physician*. 2012;4(3):527–535.
54. How JA, Lim JZR, Goh DJW, et al. Adaptation of Escherichia coli ATCC 8739 to 11% NaCl. *Dataset Papers in Biology*. 2013;2013:219095.
55. Loo BZL, Low SXZ, Aw ZQ, et al. Escherichia coli ATCC 8739 Adapts Specifically to Sodium Chloride, Monosodium Glutamate, and Benzoic Acid after Prolonged Stress. *Asia Pacific Journal of Life Sciences*. 2014;7(3):243–258.
56. Ling MHT. Applications of Artificial Life and Digital Organisms in the Study of Genetic Evolution. *ACSIJ*. 2014;3(4):107–112.
57. Castillo CFG, Ling MHT. Digital Organism Simulation Environment (DOSE): A Library for Ecologically-Based In Silico Experimental Evolution. *ACSIJ*. 2014;5(3):44–50.
58. Castillo CFG, Ling MHT. Resistant Traits in Digital Organisms Do Not Revert Preselection Status despite Extended Deselection: Implications to Microbial Antibiotics Resistance. *BioMed Research International*. 2014;2014:648389.
59. Gullberg E, Cao S, Berg OG, et al. Selection of Resistant Bacteria at Very Low Antibiotic Concentrations. *PLoS pathogens*. 2011;7:e1002158.
60. Castillo CFG, Chay ZE, Ling MHT. Resistance Maintained in Digital Organisms Despite Guanine/Cytosine-Based Fitness Cost and Extended De-Selection: Implications to Microbial Antibiotics Resistance. *MOJ Proteomics & Bioinforma*. 2015;2(2):00039.
61. Chowdhury A, Khodayari A, Maranas CD. Improving Prediction Fidelity of Cellular Metabolism with Kinetic Descriptions. *Current Opinion in Biotechnology*. 2015;36:57–64.
62. Motta S, Pappalardo F. Mathematical Modeling of Biological Systems. *Briefings in Bioinformatics*. 2013;14:411–422.
63. Machado D, Rodrigues LR, Rocha I. A Kinetic Model for Curcumin Production in Escherichia coli. *BioSystems*. 2014;125:16–21.
64. Weaver LJ, Sousa MM, Wang G, et al. A Kinetic-Based Approach to Understanding Heterologous Mevalonate Pathway Function in E. coli. *Biotechnol Bioeng*. 2015;112(1):111–119.
65. Weisberg M. Biology and Philosophy symposium on Simulation and Similarity: Using Models to Understand the World. *Biology & Philosophy*. 2015;30:299–310.
66. Lung W. *Modeling of Phosphorus Sediment Water Interactions in White Lake, Michigan*. USA: University of Michigan; 1975.
67. Michor M, Beal K. Improving Cancer Treatment Via Mathematical Modeling: Surmounting the Challenges Is Worth the Effort. *Cell*. 2015;163(5):1059–1063.
68. Dolgin E. The mathematician versus the malignancy. *Nat Med*. 2014;20(5):460–463.

69. Frigg R, Reiss J. The Philosophy of Simulation: Hot New Issues or Same Old Stew? *Synthese*. 2008;169:593–613.
70. Veldman MB, Lin S. Zebrafish as a developmental model organism for pediatric research. *Pediatr Res*. 2008;64(5):470–476.
71. Sargent RG. Verification and validation of simulation models. *Journal of Simulation*. 2013;7:12–24.
72. Ruphy S. Computer Simulations: A New Mode of Scientific Inquiry? *The Role of Technology in Science: Philosophical Perspectives*. 2015:131–148.
73. Massoud TF, Hademenos GJ, Young WL, et al. Principles and Philosophy of Modeling in Biomedical Research. *FASEB J*. 1998;12(3):275–285.
74. Saltelli A, Ratto M, Andres T, et al. *Global sensitivity analysis: the primer*. USA: John Wiley & Sons; 2008.
75. J Hicks JL, Uchida TK, Seth A, et al. Is My Model Good Enough? Best Practices for Verification and Validation of Musculoskeletal Models and Simulations of Movement. *J Biomech Eng*. 2015;137(2):020905.
76. Kohavi R. *A study of cross-validation and bootstrap for accuracy estimation and model selection*. Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence. 1995. 2:1137–1143.
77. Frigg R, Hartmann S. Models in Science. *The Stanford Encyclopedia of Philosophy*. 2012.
78. Winsberg E. Computer Simulation and the Philosophy of Science. *Philosophy Compass*. 2009;4(5):835–845.