“How do mathematical models simplify reality, and when do they fail?”

Mathematical models have the power to change the world, and they have been for hundreds of years. From their impacts on disease spread control in the COVID 19 pandemic and the recent flight path of the Artemis II shuttle, mathematical models have been the underlying backbone of applied science.

A model is a tangible representation of an object, system or concept. Models are seen at every point in life, whether that be recreational or professional. Professionally, they are especially valuable, both in illustration, such as in models created by an architect, and in prediction, such as in the testing of the aerodynamics of a £20 million 1995 McLaren F1 [1]. All these models share something in common: they aim to predict how it might interact with its specific environment. This is done by undertaking key assumptions, which often allows complex dynamic systems to be simplified. Subsequently, these models allow their key system aspects to be highlighted. As a result, designers can make edits to their product and improve it before committing their resources to a faulty product. This is especially true and valuable in high-speed engineering projects.

In the case of a mathematical model (MM), they should require a greater degree of abstraction. Instead of modelling physical objects, MMs aim to describe relationships between variables using specific equations under a certain system. Here is where the power lies. By using a sufficiently accurate model, an MM can take an input and produce a reasonable output about the variables’ future relationships. Ultimately, making a 100% accurate model is nearly impossible with large chaotic systems and would require immense amounts of computing power to manage many variables and equations. Instead, key assumptions can be made to simplify the model to increase its efficiency and usefulness. However, this can also be an MM’s greatest weakness, as excessive simplifications can lead to limitations and inaccuracies. This balance is the key in evaluating the validity of a prediction made by a model.

The dynamic systems that MMs aim to replicate contain huge amounts of data and many less significant variables. Overall, it means that only a few core aspects of the system are needed to create an accurate prediction, allowing for a much simpler model. For example, modelling the velocity of water in a pipe is very hard to do if you decide to model each molecule. Modelling 10ml of water for 1 microsecond would require orders of magnitude of computing power greater than what we currently have access to, even on the supercomputer Frontier [2]. For instance, a computer would have to simulate 3.35x10^23 molecules’ [3] position, speed and acceleration vectors [4]. Obviously, this is not an option. Instead, we can use the Navier-Stokes equations [5] to model the flow of the liquid through the pipe, if we undertake several key assumptions. This equation still lets us predict large-scale patterns across the fluid, such as vortices, without having to examine individual motions of the water molecules. Modelling the fluid as a continuous ‘curve’ subsequently allows data to be visually presentable and easier to understand [6].

In addition to an MM’s ability to predict future outcomes, simplify complex systems, and present data cleanly, they offer a plethora of other opportunities which are extremely useful. Primarily, one of the main implicit benefits of an MM is its ability to aid decision-making even if the model is not completely accurate. Based on sufficiently probable future predictions, manipulation of future outcomes becomes possible. One example of this is the MMs, which are used by hedge funds to predict the movement of the stock market, enabling them to buy and sell at high speed to gain profit based on the model [7]. Moreover, they let complex systems be studied, which can lead to branching research. One application of this branching research is optimisation. Optimisation is the process of changing a system in favour of a certain variable(s) to increase/decrease the output of other related variables. Wide scale optimisation could not be facilitated without MMs. This is because altering an MM does not effect the real system, so precious resources are not wasted manipulating the real-world system, on a trial-and-error based method. A small example of optimisation is analysis of queue times in coffee shops. A shorter queue means that less people leave the queue out of frustration, so the coffee shop increases its revenue by optimising time efficiency. Another equally useful application of an MM is their ability to study general systems to understand large underlying processes. This can be done by reintroducing/omitting certain variables within the model and then evaluating the effect these changes have on the output. This is especially easy to see in MMs based on ecosystems, which are highly complex and dynamic, due to the interdependence of each organism on the others. By removing a certain producer or predator, the population number of a species can dramatically increase or decrease, leading to secondary and further subsequent effects. These effects can be directly tracked using the MM over time [8]. This is precisely what allows systems to be studied and understood.

Alternatively, MMs have a high potential to output inaccurate predictions through many means. What makes MMs so strong is also what can make them fail due to oversimplification. Oversimplification can occur when a key variable(s) is omitted/missed from the model, which has significant exponential impacts. Another similar downfall can occur when incorrect assumptions are used, as is often done in simplified physics models, where air resistance is ignored. Surprisingly, this can lead to the model producing accurate results, but for the wrong system under different circumstances. On the contrary, this argument can also be reversed when MMs from 1 specific system are adapted to another similar, but different system. The outputs are still tailored to the original system, leading to the generalisation of results which could be beneficial, but will increase the uncertainty of the data. A good example of this is weather models [9], where different meteorological processes work in different places due to location and related factors. Data quality is another significant factor which can decrease the effectiveness of a model. Evidently, MMs are suited to a set of data, yet if that data itself is shallow/inaccurate, then the reliability of its predictions is limited, and not suited to the real system. It is like analysing the biodiversity of the Pacific Ocean from a bathtub of water.

Lastly, model selection bias can play a major factor in the reliability of an MM. For instance, there are many different sets of equations which can be found that suit a set of data, yet a model which produces a more positive outcome may be favoured over the more negative one. An example is the Gaussian Copula model, which made financial risks appear less dangerous [10]. Additionally, you could argue that an Artificial Intelligence (AI) could suitably make an MM without human bias. Unfortunately, all AIs are trained on data, which will contain training data bias as a necessity, so bias must be controlled and factored into the results. This is apparent in Amazon’s hiring tool between 2015-2018, which penalised CVs containing the word ‘women’ [11]. This reduced the number of female employees which were accepted for an interview, as the training data was androcentric (containing male bias). Subsequently, MMs can amplify the effects of bias subtly, due to the speed and scale at which automated systems operate. This leads to systemic problems caused by their careless application. Therefore, another downside of MMs are their ability to deflect responsibility, which can hinder ethical development and cause widespread harm if used irresponsibly.

Sensitivity and validation are other key factors which relate to the reliability of results produced by an MM. Sensitivity refers to how much the output changes based on slight changes to the inputs. A small change to the inputs causing a large change to the outputs is called very sensitive. Sensitivity can both negatively and positively impact a model based on the scale at which it operates. If an MM is used to represent a population, high sensitivity is likely less reliable, as a difference of a few people when compared to a large starting value, such as that of London, is unlikely to have such a large impact on the population size after 10 years. This is evident as the initial population is already close to 9 million [12]. Interestingly, it can be very useful when modelling a chaotic system, such as that of a nuclear reactor. A small change to the number of slow-moving neutrons can be the difference between no fission and a large chain reaction, due to the exponential nature of nuclear decay [13].

Validation refers to the ability of a model to actually measure what it is intended to represent. Validation is critical when modelling outcomes which are far in the future. Errors grow as time passes, which leads to these errors becoming exponentiated as they are used as inputs in further equations. Also, assumptions become less clear in the future and may become invalid. This is in accordance with the Somerset floods of 2014, when the dynamic river system was shallower than expected due to sediment buildup, leading to excessive flooding [14]. The bed of a river is changing. Ultimately, this is also why a very sensitive model can be less reliable further into the future, due to the drastic disparity between results which differ by a small input as this effect is multiplied.

Lastly, MMs are constantly being improved by mathematicians on a global scale as more complex MMs are required to imitate experimental technology. There are a few trivial ways of improving an MM: making it more complex, using higher-quality data, fine-tuning assumptions, testing validity and sensitivity, continuously updating the model concerning the system it models, and increasing the processing power of the computer it operates on. Luckily, there are more interesting ways of improving MMs; for instance, stochastic processes (random walk, or a Markov process [15],[16]) can be used to attempt to model randomness using probability theory. This is very important in weather and stock market analysis, as these systems are very chaotic and hard to predict with 'limited' models. Moreover, mathematicians are working on the polynomial time (P) versus nondeterministic polynomial time (NP) problem, which is a millennium prize problem and an unsolved problem in mathematics [17]. P means that a problem can be solved 'quickly' by a computer, and NP means that a solution can be checked fast, but a solution may take a nondeterministic amount of time to solve normally. If P=NP, then many 'hard' problems become solvable, such as the travelling salesman [18]. Directly, this will have a huge capacity to impact optimisation and AI, which is driven by MMs; on the other hand, if P does not equal NP, then some problems will be fundamentally difficult to solve, limiting MMs' ability for optimisation.

MMs are evidently one of the most powerful instruments that humanity has developed. This is because they clear the fog upon chaotic systems and reduce them to simple equations. Inevitably, as processing power increases and stochastic methods are improved, MMs will soon be able to explain what we once thought was too random to predict. Yet, bias will remain. It will continue to hide in neutral assumptions, deflecting accountability. So ultimately, the strength of an MM will not lie in the mathematical elegance, but in the critical evaluation of those who build it and apply it.

References:

[1]: https://www.charlottemclaren.com/mclaren-f1-road-car-sets-record-price-at-auction/?srsltid=AfmBOoqE7VpZVcKIevheYXCHYyb-rlzjYDrediGnr-c15oNLh8QP9eRi
[2]: https://www.olcf.ornl.gov/frontier/
[3]: mol x Avogadro's constant = 10/18 x 6.023 x 10^23 = 3.461 x 10^23
[4]: https://phydemo.app/balls-simulation/, we can see the large number of variables required to simulate just a few balls (water molecules), in addition to the displacement, velocity and acceleration vectors.
[5]: https://www.grc.nasa.gov/www/k-12/airplane/nseqs.html
[6]: https://www.atlassian.com/data/charts/heatmap-complete-guide
[7]: https://personalpages.manchester.ac.uk/staff/paul.johnson-2/resources/mathFinanceWorkshop/lecture-mfw.pdf
[8]: https://education.nationalgeographic.org/resource/reintroduction-top-predator/
[9]: https://www.theweatheroutlook.com/twodata/datmdlout.aspx
[10]: https://www.wired.com/2009/02/wp-quant/
[11]: https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/
[12]: https://data.london.gov.uk/social-evidence-base/population/
[13]: https://nrl.mit.edu/reactor/fission-process/
[14]: https://www.internetgeography.net/geotopics/the-somerset-levels-flood-case-study/
[15]: https://en.wikipedia.org/wiki/Random_walk
[16]: https://en.wikipedia.org/wiki/Markov_chain
[17]: https://en.wikipedia.org/wiki/P_versus_NP_problem
[18]: https://www.lancaster.ac.uk/stor-i-student-sites/libby-daniells/2020/04/21/the-travelling-salesman-problem/