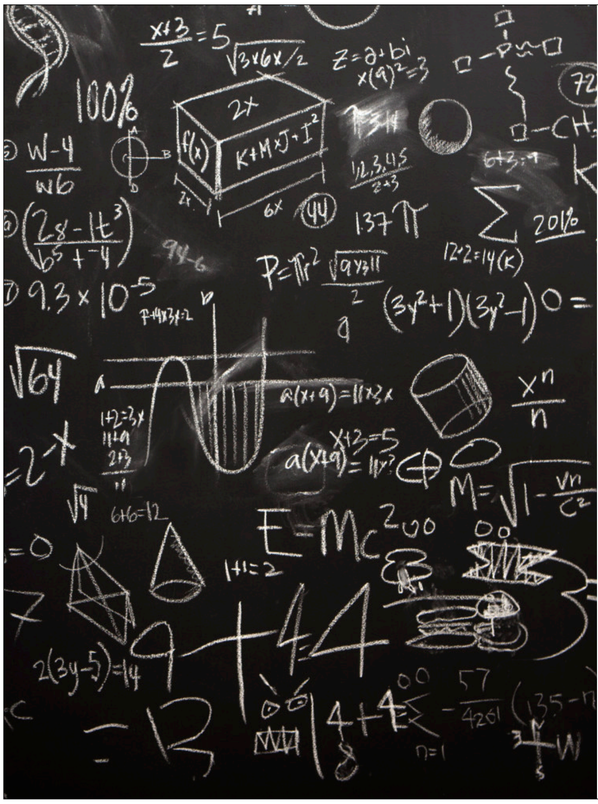


COMP4030

DATA MODELLING AND ANALYSIS

Lecture 2: Introduction to Modelling

LECTURE OUTLINE

1. Learning Outcomes
2. Introduction: a complex world
3. Systems
4. Models
5. Simulations
6. Mathematics as a natural modelling language
7. Classification of models
8. Summary (COMP4030 at a glance)

LECTURE OUTCOMES

- At the end of this lecture, you should be able to answer these questions:
  - What is a model?
  - Why are models useful?
  - What is a mathematical model?
  - How is a mathematical model defined?
  - How do you classify a mathematical model?
  - What is the purpose of mathematical models?

COMPLEX WORLD

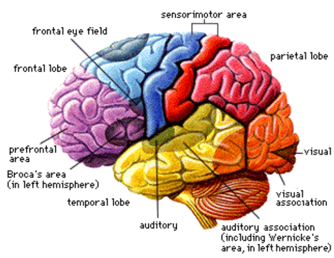
- Problems in science and engineering tend to be difficult:
  - System complexity
- What is a system?
  - A system is as an entity or collection of entities whose properties we want to study
- Modelling allows us tackle this complexity:
  - Models contribute to making these problems tractable
  - Models simplify problems and allow us to understand complex situations

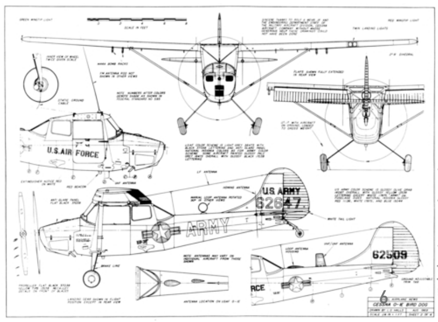
COMPLEX WORLD

- What do scientists/engineers do most of the time?
  - They try to:
    - Understand systems
    - Develop systems
    - Optimize systems
- But, everywhere you look there are complex systems

COMPLEX WORLD

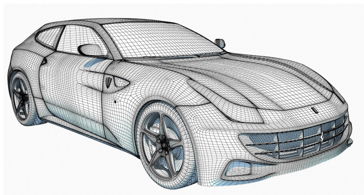
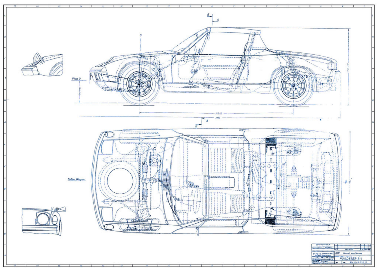
- Examples of <u>natural</u> complex systems:
  - The human brain
  - The nervous system
  - An ant colony
- Examples of an <u>artificial</u> complex system:
  - A plane
  - A computer
  - The Internet
  - Elections
- So how do we tackle complexity?





COMPLEX WORLD

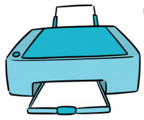
- So how do we tackle complexity?
- Simplification
  - We build a simplified description of the system, which contains information relevant to our problem
  - This “simplified description” is a model by definition
  - Depending on the level of detail required, we can build different models with different levels of detail.

COMPLEX WORLD: AN EXAMPLE

- We try, but our car doesn’t start:
  - This can be seen as a technical problem (not starting) involving a complex artificial system (car).
- Most people (~~I~~) use a simplified model of the car:
  - Fuel tank + battery
- Someone with that model, checks the fuel tank and/or battery
  - Hopefully this solves the problem.
- If not, then we send the car to a mechanic, who tries to fix the car by resorting to a more sophisticated model of the car’s engine, which contains elements such as valves, pistons, injectors, etc.

COMPLEX WORLD: ANOTHER EXAMPLE



- The same person can use models at different granularities to solve problems, too.
- An example:
  - Your parents call you and tell you that the printer is not working.
  - What do you do?

What is a model?

MODEL DEFINITION

- The following definition of a *model* is by Marvin Minsky (1965):

> ***Definition 2.1 (Model)*** *To an observer B, an object A* is a model of an object A to the extent that B can use A* to answer questions of interest about A.*

MODEL DEFINITION

- With respect to the example in the previous slide:
  - B = car driver
  - A* = tank/battery model
  - A = car

  OR
  - B = mechanic
  - A* = complex car model
  - A = car

MODEL DEFINITION

- What is the purpose of a model?
  - To answer questions
  - To solve problems
- Model quality: relative to purpose
- How detailed should a model be?
  - As detailed as possible?
  - What do you think?

MODEL DEFINITION

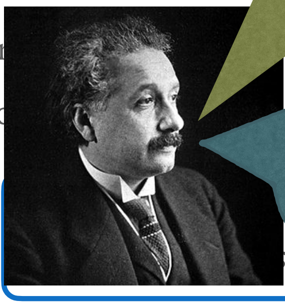
- How detailed should a model be?
- One of the main aims:
  - To understand complex systems
- If models are too complex:
  - no insight
- If models are too simple:
  - might not be accurate

The best model is the simplest one that still allows us to understand the system and solve problems related to it.

MODEL DEFINITION

- How deta
- One o
  - To u
- If models ar
  - no insight
- If models ar
  - might no

It can scarcely be denied that the supreme goal of all theory is to make <u>the irreducible basic elements as simple and as few as possible</u> without having to surrender the adequate representation of a single datum of experience.



Everything should be made as simple as possible, but no simpler.

ys

MATHEMATICAL MODELS

- A mathematical model is a model created using mathematical concepts and equations.
  - Mathematical modellers deal with a variety of real problems and translate each problem into a mathematical form
- The system needs to be observable/measurable
  - Most systems also allow for some kind of input (observable variables)
  - The output of the system needs to be collected (variable(s) we want to predict)
- Mathematical models map relationships between system inputs and outputs using mathematics and statistics

MATHEMATICAL MODELS AND MODELLING

- Example 1: Photosynthesis system:
  - **System:** plant cell
  - **Input**: light, CO_2 and water levels / **Output**: carbohydrate production
- Example 2: Optimising traffic flow near a roundabout
  - Input: junction configuration, number of cars, their speed, traffic density, etc.
  - Output: good layout / bad layout
  - **Note:** The alternative to mathematical modelling would be to build several designs, and to test them with vehicles.
    - Expensive! Time consuming! Not practical!

MATHEMATICAL MODELS: MODELS AND MODELLING

- Example 3: Considering the weather and my behaviour this past two weeks, will I play tennis today?

Input *Output*

	Outlook	Temp.	Humid.	Windy	Play
1	sunny	hot	high	false	no
2	sunny	hot	high	true	no
3	overcast	hot	high	false	yes
4	rainy	mild	high	false	yes
5	rainy	cool	normal	false	yes
6	rainy	cool	normal	true	no
7	overcast	cool	normal	true	yes
8	sunny	mild	high	false	no
9	sunny	cool	normal	false	yes
10	rainy	mild	normal	false	yes
11	sunny	mild	normal	true	yes
12	overcast	mild	high	true	yes
13	overcast	hot	normal	false	yes
14	rainy	mild	high	true	no

And what is the system?

The human context within which the decision is made (i.e. nature of the game, cultural and psychological factors, etc.)

MATHEMATICAL MODELS: REPRESENTATION

- Inputs and outputs often represented by numerical values
- Input (x) and output (y) relationship as a function:
  - $y = f(x)$
- If we can find such a function:
  - Simplified description
  - Can answer our questions
  - This is consistent with our definition of a model
- Mapping internal mechanisms onto mathematical structures allows us to:
  - Understand,
  - Develop, and
  - Optimize systems

MATHEMATICAL MODELS: DEFINITION

- Many definitions are possible
  - Again we opt for a very general one

> ***Definition 2.2 (Mathematical Model)*** *A mathematical model is a triplet (S, Q, M) where S is a system, Q is a set of questions* $Q = \{Q_1, Q_2, \ldots, Q_m\}$ *relating to S, and M is a set of mathematical statements* $M = \{\Sigma_1, \Sigma_2, \ldots, \Sigma_n\}$ *which can be used to answer Q.*

- The order of (S,Q,M) follows chronology
- The model its linked to its ***purpose***
- Which element of (S,Q,M) guarantees a purpose?

MATHEMATICAL MODELS: REDUCED SYSTEMS

- State variables (or dependent variables)
  - Computed from system parameters and mathematical statements in M. For example:
    - s_1 = car "startability" (0 = no "startability"; 1 = good "startability")
    - $M = \{\Sigma_1\}$ where Σ_1 = "if ($p_1 = 0$) OR ($p_2 = 0$) then $s_1 = 0$ else $s_1 = 1$"
    - Q = "Can the car start?"

MATHEMATICAL MODELS: REDUCED SYSTEMS

- Reduced set of parameters (e.g. $S_r = \{p_1, p_2\}$), e.g.:
  - p_1 = fuel level / p_2 =battery level
- State variables (or dependent variables)
  - Describe the state of S in terms of M, which are required to answer Q.
- System parameters can be:
  - constants
  - variables (i.e. independent variables)
  - functions

MATHEMATICAL MODELS: BOXES

- Depending on the system, we will be dealing with:
- Black-box Models:
  - We know inputs and their outputs
  - But we don't know the process
- White-box Models:
  - We know the input-output mapping **and** the relationships (parameters) governing their mapping
- Grey-box models: the most common case
  - Start with an unknown process
  - As we carry out more experiments, we learn (discover) more input/output relationships and parameters
- Gradually move from black boxes towards white boxes

MATHEMATICAL MODELS: HOW TO BUILD MODELS

- The steps you will usually follow are:
  1. **Problem Identification:** what do we want to know? What is its purpose and objective? How will be measure outcome? What is Q?
     - This is not easy!
  2. **System Analysis:** What are the sources of facts and data? Are they reliable? What data is needed? What is out there in the literature? Who are the experts in this? Is there available data, or will new data need to be collected? If so, how? What is S and Sr?
     - Simplification is key! We can't/shouldn't express all factors (variables) of a real system.
     - What influences system's behaviour in a way that is relevant to our question? Variables.
     - The variables we want to understand (predict): *dependent variables*.
     - All other variables: *independent variables*.

MATHEMATICAL MODELS: HOW TO BUILD MODELS

3. **Formulate mathematical model:** Look for the simplest model (using diagrams if needed). Draw relationships and equations connecting input and outputs. What is (are) M?
   - Collect more data if needed to explain behaviour, assigning proper units.
4. **Obtain Mathematical Solution:** Use algebra, numerical methods, calculus, programming or simulations to solve the model. Does it have any limitations? Does M reflect the relationship between Q and S?
   - Careful not to conclude that passing a few "validation tests" implies that the model is correct.
   - The model is consistent with (and validated by) the current data, but may be inconsistent with future data.

MATHEMATICAL MODELS: HOW TO BUILD MODELS

5. **Interpret the model:** examine the results obtained. Is the model robust? Are there any limitations on the data? Is the model usable? Can it be shared with others? Is it understandable? Does it answer Q?
6. **Compare with reality:** can your results be tested against real data? Do your solutions make sense? Do your predictions match the real data? Has it fulfilled its purpose? Can it be improved, or simplified?
   - Where does it work? Consistency vs. correctness vs. theories vs. laws
   - Will it need to be adjusted/maintained?
7. **Report back!**

MATHEMATICAL MODELS: STEPS IN MODEL CONSTRUCTION (EXPECTATIONS)

1. *Identify the Question or Problem*
2. *System Analysis*
3. *Formulate Mathematical Model*
4. *Obtain Mathematical Solution*
5. *Interpret the Solution*
6. *Compare with reality*

Although different authors might present the modelling steps slightly differently, it is usually quite easy to map the different descriptions

MATHEMATICAL MODELS: STEPS IN MODEL CONSTRUCTION (REALITY)

1. *Identify the Question or Problem*
2. *System Analysis*
3. *Formulate Mathematical Model*
4. *Obtain Mathematical Solution*
5. *Interpret the Solution*
6. *Compare with reality*

MODEL CONSTRUCTION: THE SIMPLIFICATION CONUNDRUM

- As mentioned earlier, the modelling process hardly ever proceeds sequentially through all the steps
- The modelling process is most often an iterative process with successive simplifications and refinements

Model Simplification	Model Refinement
Restrict the problem scope	Expand the problem scope
Neglect variables	Add new variables
Combine subsets of variables	Consider each variable in detail
Set some parameters to be constant	Set other parameters to be variable
Assume linear relationships	Consider non-linear relationships
Incorporate more assumptions	Reduce the number of assumptions

MATHEMATICAL MODELS: THE IMPORTANCE OF DATA

- In simple systems, internal relationships might be evident
  - Car example: the relationship between battery and fuel
- In more complex systems:
  - Data (observations/measurements) will be needed to ascertain these relationships.
    - In other words: inputs and outputs
    - These will be obtained through experimentation within the system.
      - Certain inputs are introduced in the system, and their associated output(s) are collected
    - This collection of inputs and (associated) outputs: a dataset

MATHEMATICAL MODELS: DEFINITION

- Many definitions are possible
  - Again we opt for a very general one

Definition 2.2 (Mathematical Model) *A mathematical model is a triplet (S, Q, M) where S is a system, Q is a set of questions* $Q = \{Q_1, Q_2, \ldots, Q_m\}$ *relating to S, and M is a set of mathematical statements* $M = \{\Sigma_1, \Sigma_2, \ldots, \Sigma_n\}$ *which can be used to answer Q.*

CLASSIFICATION OF MODELS: SYSTEM INFO

- Phenomenological model

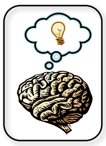
    Definition 2.3 (Phenomenological Model) *A mathematical model (S, Q, M) is called phenomenological if it was constructed based on experimental data only, using no a priori information about S.*

- Mechanistic model

    Definition 2.4 (Mechanistic Model) *A mathematical model (S, Q, M) is called mechanistic if some of the statements in M are based on a priori information about S.*

CLASSIFICATION OF MODELS: SYSTEM INFO

- Phenomenological models are also referred to as:
    - empirical models
    - statistical models
    - data-driven models
    - black box models
- Mechanistic models
    - complete knowledge
        - ??? box models
    - partial knowledge
        - ??? box models or semi-empirical models



CLASSIFICATION OF MODELS: SYSTEM INFO

- Pros & cons, in general:
  - Phenomenological models tend to:
    - [+] be more general and applicable to many different kinds of problems
    - [+] require less effort to set up
    - [-] provide less insight
    - [-] provide less predictive power
  - Mechanistic models typically allow:
    - [+] deeper insights
    - [+] more accurate predictions
    - [-] but require more effort to build
    - [-] are not universally applicable

CLASSIFICATION OF MODELS: TIME AND SPACE

- Static/dynamic models

  Definition 2.5 (Static/dynamic models)

  A mathematical model (S,Q,M) is called:

  - *dynamic,* if at least one of its system parameters or state variables depends on time
  - *static* otherwise
- Distributed/lumped models

  Definition 2.6 (Distributed/lumped models)

  A mathematical model (S,Q,M) is called:

  - *distributed,* if at least one of its system parameters or state variables depends on a space variable
  - *lumped* otherwise

CLASSIFICATION OF MODELS: ADDITIONAL CLASSIFICATIONS

- Natural/artificial
- Stochastic/deterministic
- Continuous/discrete
- Direct / inverse
- Linear / nonlinear
- Analytical/numerical
- Research / Management
- Field of application

Note: there is a continuous spectrum between each of these extremes

MODELS IN THE REAL WORLD: DEFINITIONS

- Robustness
  - A model is deemed to be robust if the conclusions drawn from it do not depend heavily on specific parameter settings.
  - If parameters can change somewhat without affecting the conclusions, then the model is robust.

MODELS IN THE REAL WORLD: DEFINITIONS

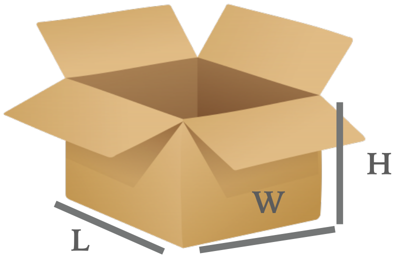
- Fragility
  - A model is deemed to be fragile if the conclusions drawn from it are very prone to change with small variations to parameter settings.

MODELS IN THE REAL WORLD: DEFINITIONS

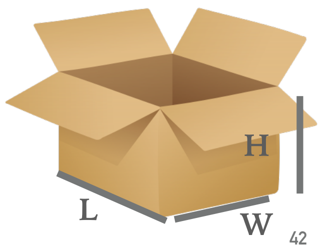
- Sensitivity
  - Models usually lie somewhere along a sensitivity spectrum between extremely robust and extremely fragile models.
  - The more robust a model is, the less sensitive it is.
  - Conversely, the more fragile a model is, the more sensitive it is.

CLASSIFICATION OF MODELS EXAMPLE: BOX OPTIMISATION

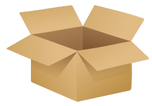
- A food company wants to maximise the amount of items they can ship on boxes they send to supermarkets.
- Create a model to calculate the space inside their current shipping boxes.
- You know the box's height (H), width (W), and length (L).
- **S? S_r? M? Q?**
- **What assumptions are we making?**



CLASSIFICATION OF MODELS EXAMPLE: BOX OPTIMISATION

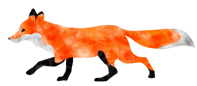


CLASSIFICATION OF MODELS EXAMPLE: BOX OPTIMISATION



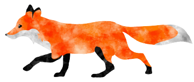
- Classification?
  - Phenomenological / mechanistic
  - Static / Dynamic
  - Lumped / Distributed
  - Natural / artificial
  - Stochastic / deterministic
  - Continuous / discrete
  - Direct/ Inverse
  - Research / Management
  - Field of application?

CLASSIFICATION OF MODELS EXAMPLE 3: POPULATION CHANGES OF RED FOXES

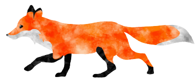


- We want to model the population of urban (red) foxes in London.
  - Model changes in a monthly basis.
- Foxes reproduce quickly and die slowly, which can be carry a potential risk of overpopulation.
- In London:
  - Foxes reproduce at a monthly rate of r = 0,25
  - Foxes die at a monthly rate of d = 0.1
- Initial number of foxes in London is estimated at 150.
- **S? S_r? M? Q?**
- **What assumptions are we making?**

CLASSIFICATION OF MODELS EXAMPLE 3: POPULATION CHANGES OF RED FOXES



CLASSIFICATION OF MODELS EXAMPLE 3: POPULATION CHANGES OF RED FOXES



- Classification?
  - Phenomenological / mechanistic
  - Linear / non-linear
  - Static / Dynamic
  - Lumped / Distributed
  - Natural / artificial
  - Stochastic / deterministic
  - Continuous / discrete
  - Direct/ Inverse
  - Research / Management
  - Field of application?

THE "DON'TS" OF MATHEMATICAL MODELLING

- According to Golomb [1]:

1. Don't confuse the model with reality.
2. Don't extrapolate too much beyond the region of fit.
3. Don't distort the facts so that "reality" fits the model.
4. Don't keep an invalidated model.
5. Don't fall in love with your model.

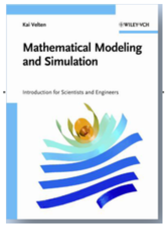
[1] Golomb, S.W. (1970) "Mathematical Models – uses and limitations". Simulation, 4(14), 197-98.

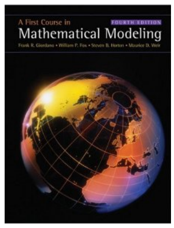
LECTURE OUTCOMES

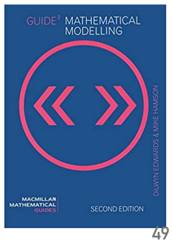
- Now, you should be able to answer these questions:
  - What is a model?
  - Why are models useful?
  - What is a mathematical model?
  - How is a mathematical model defined?
  - How do you classify a mathematical model?
  - What is the purpose of a model?

READING MATERIAL

- Books used in this lecture:
  - Velten, K. (2009), “*Mathematical Modeling and Simulation: Introduction for Scientists and Engineers*”; Wiley-VCH
  - Giordano, F.R., et al (2009)“*A First Course in Mathematical Modeling*” (4th ed.); Brooks/Cole, Cengage Learning
  - Edwards, D., Hamson M (2001) “*Guide to Mathematical Modelling - 2nd edition*”, Palgrave Mathematical Guides







THE END

Questions