



# Neuromatch Academy: Modeling Practice - Summary Sheet<sup>1</sup>

## Modeling Practice

### Framing the Question

We will try to clarify the process of computational modeling, by thinking through the logic of modeling a project. We assume that you have a general idea of a project in mind, i.e. a preliminary question, and/or phenomenon you would like to understand. You should have started developing a project idea yesterday with this brainstorming demo. Maybe you have a goal in mind. We will now work through the 4 first steps of modeling (Blohm et al., 2019):

#### Framing the question:

1. finding a phenomenon and a question to ask about it
2. understanding the state of the art
3. determining the basic ingredients
4. formulating specific, mathematically defined hypotheses

The remaining steps 5-10 are covered later.

**Importantly**, we will guide you through Steps 1-4 today. After you do more work on projects, you likely have to revise some or all of these steps 'before' you move on to the remaining steps of modeling. 'Models' here can be data analysis pipelines, not just computational models...

## Modeling Practice

### Step 1 - Asking your own question

You should already have a project idea from your brainstorming yesterday. **Write down the phenomenon, question and goal(s) if you have them.**

As a reminder, here is what you should discuss and write down:

What exact aspect of data needs modeling?

- Answer this question clearly and precisely! Otherwise you will get lost (almost guaranteed)
- Write everything down!
- Also identify aspects of data that you do not want to address (yet)

Define an evaluation method!

- How will you know your modeling is good?
- E.g. comparison to specific data (quantitative method of comparison?)

For computational models: think of an experiment that could test your model

- You essentially want your model to interface with this experiment, i.e. you want to simulate this experiment

You can find interesting questions by looking for phenomena that differ from your expectations. In **what** way does it differ? **How** could that be explained (starting to think about mechanistic questions and structural hypotheses)? **Why** could it be the way it is? What experiment could you design to investigate this phenomenon? What kind of data would you need?

### Step 1 - Make sure to avoid the pitfalls!

Question is too general

- Remember: science advances one small step at the time. Get the small step right...

Precise aspect of phenomenon you want to model is unclear

- You will fail to ask a meaningful question

You have already chosen a toolkit

- This will prevent you from thinking deeply about the best way to answer your scientific question

You don't have a clear goal

- What do you want to get out of modeling?

You don't have a potential experiment in mind

- This will help concretize your objectives and think through the logic behind your goal

## Modeling Practice

### Step 2 - Understanding the state of the art & background

Here is what you should get out of it:

Survey the literature

- What's known?
- What has already been done?
- Previous models as a starting point?
- What hypotheses have been emitted in the field?
- Are there any alternative / complementary modeling approaches?

What skill sets are required?

- Do I need to learn something before I can start?
- Ensure that no important aspect is missed

Potentially provides specific data sets / alternative modeling approaches for comparison

<sup>1</sup>t Hart et al., (2022). Neuromatch Academy: a 3-week, online summer school in computational neuroscience. Journal of Open Source Education, 5(49), 118. <https://doi.org/10.21105/jose.00118>

## Modeling Practice

### Step 3 - Determining the basic ingredients

This will allow you to think deeper about what your modeling project will need. It's a crucial step before you can formulate hypotheses because you first need to understand what your modeling approach will need. There are 2 aspects you want to think about:

1. What parameters / variables are needed?
  - Constants?
  - Do they change over space, time, conditions...?
  - What details can be omitted?
  - Constraints, initial conditions?
  - Model inputs / outputs?
2. Variables needed to describe the process to be modeled?
  - Brainstorming!
  - What can be observed / measured? latent variables?
  - Where do these variables come from?
  - Do any abstract concepts need to be instantiated as variables?
  - E.g. value, utility, uncertainty, cost, salience, goals, strategy, plant, dynamics
  - Instantiate them so that they relate to potential measurements!

This is a step where your prior knowledge and intuition is tested. You want to end up with an inventory of *specific* concepts and/or interactions that need to be instantiated.

### Step 3 - Pitfalls

I'm experienced, I don't need to think about ingredients anymore

- Or so you think...

I can't think of any ingredients

- Think about the potential experiment. What are your stimuli? What parameters? What would you control? What do you measure?

I have all inputs and outputs

- Good! But what will link them? Thinking about that will start shaping your model and hypotheses

I can't think of any links (= mechanisms)

- You will acquire a library of potential mechanisms as you keep modeling and learning
- But the literature will often give you hints through hypotheses
- If you still can't think of links, then maybe you're missing ingredients?

## Modeling Practice

### Step 4 - Formulating your hypothesis

Once you have your question and goal lined up, you have done a literature review (let's assume for now) and you have thought about ingredients needed for your model, you're now ready to start thinking about *specific* hypotheses.

Formulating hypotheses really consists of two consecutive steps:

1. You think about the hypotheses in words by relating ingredients identified in Step 3
  - What is the model mechanism expected to do?
  - How are different parameters expected to influence model results?
2. You then express these hypotheses in mathematical language by giving the ingredients identified in Step 3 specific variable names.
  - Be explicit, e.g.  $y(t) = f(x(t), k)$  but  $z(t)$  doesn't influence  $y$

There are also 'structural hypotheses' that make assumptions on what model components you hypothesize will be crucial to capture the phenomenon at hand.

**Important:** Formulating the hypotheses is the last step before starting to model. This step determines the model approach and ingredients. It provides a more detailed description of the question / goal from Step 1. The more precise the hypotheses, the easier the model will be to justify.

### Step 4 - Pitfalls

I don't need hypotheses, I will just play around with

- Hypotheses help determine and specify goals. You can (and should) still play...

My hypotheses don't match my question (or vice versa)

- This is a normal part of the process!
- You need to loop back to Step 1 and revisit your question / phenomenon / goals

I can't write down a math hypothesis

- Often that means you lack ingredients and/or clarity on the hypothesis
- OR: you have a "structural" hypothesis, i.e. you expect a certain model component to be crucial in explaining the phenomenon / answering the question

## Modeling Practice Example - The Train illusion

### Step 1 - Asking your own question

The train illusion occurs when sitting on a train and viewing another train outside the window. Suddenly, the other train 'seems' to move, i.e. you experience visual motion of the other train relative to your train. But which train is actually moving?

Often people have the wrong percept. In particular, they think their own train might be moving when it's the other train that moves; or vice versa. The illusion is usually resolved once you gain vision of the surroundings that lets you disambiguate the relative motion; or if you experience strong vibrations indicating that it is indeed your own train that is in motion.

We asked the following (arbitrary) question for our demo project: "How do noisy vestibular estimates of motion lead to illusory percepts of self motion?"

### Step 2 - Understanding the state of the art & background

This is where you would do a literature search to learn more about what's known about self-motion perception and vestibular signals. You would also want to examine any attempts to model self-motion, perceptual decision making and vestibular processing.

### Step 3 - Determining the basic ingredients

We determined that we probably needed the following ingredients for our model:

- Vestibular input:  $v(t)$
- Binary decision output:  $d$  - time dependent?
- Decision threshold:  $\theta$
- A filter (maybe running average?):  $f$
- An integration mechanism to get from vestibular acceleration to sensed velocity:  $\int$

### Step 4 - Formulating your hypothesis

Our main hypothesis is that the strength of the illusion has a linear relationship to the amplitude of vestibular noise. Mathematically, this would write as  $S = k \cdot N$  where  $S$  is the illusion strength and  $N$  is the noise level, and  $k$  is a free parameter. we could simply use the frequency of occurrence across repetitions as the 'strength of the illusion'. We would get the noise as the standard deviation of  $v(t)$ , i.e.

$$N = E[v(t)^2],$$

where  $E$  stands for the expected value. Do we need to take the average across time points? doesn't really matter because we have the generative process, so we can just use the  $\sigma$  that we define

## Modeling Practice

### Modelling the Question

How-to-model guide.

#### Implementing the model

5. Toolkit selection
6. Planning the model
7. Implementing the model

#### Model testing

8. Completing the model
9. Testing and evaluating the model

#### Publishing

10. Publishing models

### References

1. Blohm G, Kording KP, Schrater PR (2020). A *How-to-Model Guide for Neuroscience* eNeuro, 7(1). <https://doi.org/10.1523/ENEURO.0352-19.2019>
2. Mensh B, Kording K (2017). *Ten simple rules for structuring papers*. PLOS Comput Biol 13(9): e1005619. <https://doi.org/10.1371/journal.pcbi.1005619>

## Modeling Practice

### Step 5 - Selecting the toolkit

In selecting the right toolkit, i.e. the right mathematics, computer science, engineering, or physics, etc approaches, you should consider the following important rules:

1. Determine the level of abstraction
2. Select the toolkit that best represents the ingredients and hypotheses
3. Toolkit should express all needed relationships, mechanisms and concepts
4. Keep it as simple as possible!

#### Guiding questions:

What is the most appropriate approach to answer your question?

- What level of abstraction is needed?
- Determine granularity / scale based on hypotheses & goals
- Stay as high-level as possible, but be as detailed as needed!!!

#### Select the toolkit

- Requires prior knowledge about flexibility / limitations of toolkit
- Often more than one option possible
- Some toolkits are more flexible, span a wider range of behaviour and/or are lumpable
- Also determines how the model will be solved, i.e. simulated

Viewing modeling as a decision process might help providing clarity regarding different model types, and how framing the problem and stating your goals influences the toolkit selection. Don't be afraid to pursue goals that no one else pursues; diversity of models should be encouraged because it results in complementary model considerations.

### Step 5 - Pitfalls

Choosing a toolkit for the toolkit's sake (e.g. DL because it's cool to do deep learning)

- this will prevent you to really answer your research question and/or speak to your hypotheses

Being at the wrong level of abstraction (see W1D1 intro)

- too complex toolkits will have too many parameters you can't meaningfully constrain, and/or that add needless complexity
- too simple toolkits will lack means to implement the details you care about

Not knowing any toolkits

- this highlights a lack of literature review and/or background work to learn about the tools used by the field

## Modeling Practice

### Step 6 - Planning / drafting the model

Planning the model involves thinking about the general outline of the model, its components and how they might fit together. You want to draw a model diagram, make some sketches and formalize necessary equations. This step will thus outline a plan of implementation. Once you have that plan, this will hugely facilitate the actual implementation of the model in computer code.

Your model will have:

- **inputs:** the values the system has available - this can be broken down into data, and parameters
- **outputs:** these are the predictions our model will make that you could potentially measure (e.g. in your idealized experiment)
- **model functions:** A set of functions that perform the hypothesized computations.

You will thus need to define a set of functions that take your data and some parameters as input, can run your model, and output a prediction for a hypothetical measurement. **Guiding principles:**

Keep it as simple as possible!

Don't get lost in details

Draft on paper: start with a flow diagram

- Draw out model components (boxes)
- What influences what? (arrows)

Then consider each model box separately

- Draft internal workings in terms of equations
- This might require a lot of work...
- Relate box inputs to box outputs!
- Keep in mind that the model should include a way to relate model variables to measurements
- Use the question, ingredients, and hypotheses to ensure you have all required components

**Goal:** Put in place all the components of the hypothesized relationships and explanations.

### Step 6 - Pitfalls

I don't need to draft the model, I have it clearly in my mind

- you might think you do, but experience shows you're likely missing many important aspects

I can just make a rough draft

- the more detailed the draft, the easier it will be to implement the model in computer code
- rough drafts tend to forget important details that you need to think about, e.g. signals needed (where do they come from?), parameters to specify (how to constrain them?), etc.

Draft is too detailed or not detailed enough

- too detailed: you're writing our recursions, etc
- not detailed enough: you have no idea what's inside "boxes"

## Modeling Practice

### Step 7 - Implementing the model

This is the step where you finally start writing code! Separately implement each box, icon, or flow relationship identified in Step 6. **Test** each of those model components separately! (This is called a *unit test*). Unit testing ensures that each model components works as expected/planned.

#### Guiding principles:

Start with the easiest possible implementation

- Test functionality of model after each step before adding new model components (unit tests)
- Simple models can sometimes accomplish surprisingly much...

Add / remove different model elements

- Gain insight into working principles
- What's crucial, what isn't?
- Every component of the model must be crucial!

Make use of tools to evaluate model behavior

- E.g. graphical analysis, changing parameter sets, stability / equilibrium analyses, derive general solutions, asymptotes, periodic behaviour, etc.

**Goal:** Understand how each model component works in isolation and what the resulting overall model behavior is.

### Step 7 - Pitfalls

Building the whole model at once without testing components

- you will make mistakes. Debug model components as you go!
- debugging a complex model is close to impossible. Is it not working because individual components are not working? Or do components not "play nice" together?

Not testing if individual components are important

- It's easy to add useless components to a model. They will be distracting for you and for readers

Not testing model functionality step by step as you build up the model

'd be surprised by what basic components often can already achieve... E.g. our intuition is really bad when it comes to dynamical systems

Not using standard model testing tools

- each field has developed specific mathematical tools to test model behaviors. You'll be expected to show such evaluations. Make use of them early on!

## Modeling Practice

### Step 8 - Completing the model

Determining what you're done modeling is a hard question. Referring back to your original goals will be crucial. This is also where a precise question and specific hypotheses expressed in mathematical relationships come in handy.

**Note:** you can always keep improving our model, but at some point you need to decide that it is finished. Once you have a model that displays the properties of a system you are interested in, it should be possible to say something about your hypothesis and question. Keeping the model simple makes it easier to understand the phenomenon and answer the research question.

#### Guiding principles:

Determine a criterion

Refer to steps 1 (goals) and 4 (hypotheses)

- Does the model answer the original question sufficiently?
- Does the model satisfy your own evaluation criteria?
- Does it speak to the hypotheses?
- Can the model produce the parametric relationships hypothesized in step 4?

Make sure the model can speak to the hypothesis. Eliminate all the parameters that do not speak to the hypothesis.

**Goal:** Determine if you can answer your original research question and related hypotheses to your satisfaction. If the original goal has not been met you need to go back to the drawing board!

### Step 8 - Pitfalls

Forgetting to specify or use a criterion for model completion (in Step 1!)

- This is crucial for you not to get lost in an endless loop of model improvements

Thinking the model can answer your question / hypotheses without checking

- always check if all questions and hypotheses can be answered / tested
- you will fail on your own benchmarks if you neglect this

You think you should further improve the model

- This is only warranted if your model cannot answer your hypotheses / questions and/or meet your goals
- remember: you can always improve a model, but you want to focus on the question / hypotheses / goals at hand!

## Modeling Practice

### Step 9 - Testing and evaluating the model

Every models needs to be evaluated quantitatively. There are many ways to achieve that and not every model should be evaluated in the same way. Ultimately, model testing depends on what your goals are and what you want to get out of the model, e.g. qualitative vs quantitative fit to data.

#### Guiding principles:

By definition a model is always wrong!

- Determine upfront what is "right enough" for you

Ensure the explicit interfacing with current or future data

- model answers the questions/hypotheses/goals with a sufficient amount of detail

Quantitative evaluation methods (see also W1D3)

- Statistics: how well does the model fit data?
- Predictability: does the model make testable predictions?
- Breadth: how general is the model?

Comparison against other models (BIC, AIC, etc.)

- This is often not easy to do in a fair way... Be nice and respectful to other models. Does the model explain previous data? (this is called the subsumption principle in physics!)

A good model should provide insight that could not have been gained or would have been hard to uncover without the model

Remember, a model is a working hypotheses; a good model should be falsifiable!

**Goal:** You want to demonstrate that your model works well. You also want to make sure you can interpret the model's meaning and findings. I.e. what did the model allow you to learn that was not apparent from the data alone?

### Step 9 - Pitfalls

Thinking your model is bad

- does it answer the question / hypotheses and meet your goals? Does it provide the level of explanation and insights you set out to gain? **Then it's probably good enough!**

Not providing any quantitative evaluation

- Just do it, it's something that's expected

Not thinking deeply about your model and what you can learn from it

- this is likely the most important pitfall. You want to learn as much as you can from a model, especially about aspects that you cannot gain from the data alone
- model interpretation can open new avenues for research and experimentation. That's part of why we want to model in the first place! **A model is a hypothesis!**

## Modeling Practice

### Step 10 - Publishing the model

#### Guiding principles:

Know your target audience!

- How much math details? How much explanation of math?
- What's the message?
- Should be experimentalists in most cases!!!
- Provide intuitive explanations, analogies, etc.
- Famous researcher: "a good modeller knows how to relate to experimentalists" Clearly describe what the goals, hypotheses and performance criteria were
- Prevents from false expectation of what the model should be doing

A graphical representation is worth 1000 words (or more)  
Show model simulations in parallel to data

- Much more convincing!

Publish enough implementation details

- A good model has to be reproducible!

**Goal:** Make sure your model is well received **AND USED** by the community. In order for our model to impact the field, it needs to be accepted by our peers, and order for that to happen it matters how the model is published.

### Step 10 - Pitfalls

Not thinking of your target audience

- no one will understand you if you write for yourself because only you know what you know...

Forgetting about Steps 1-4

- clearly spelling out steps 1-4 allows the reader to appreciate your goals and sets limits on model criticism
  - no one can tell you your model didn't do something you never set out to do
  - prevents unreasonable claims / rejection of paper

Thinking you don't need figures to explain your model

- your model draft is a great starting point!
- make figures that provide intuition about model behavior (just like you would create figures to provide intuition about experimental data)

my code is too messy to be published

- not an option (many journal now rightfully require it)
- code cleanly right from the start
  - model drafting should help with that
  - comment your code! Then comment more...

## Modeling Practice

### Writing the abstract of a modeling paper

Abstracts are very stereotyped pieces of writing that contain highly condensed information. To write a summary (= abstract) of your modeling, you can use these questions as a guide:

- **What is the phenomena?** Here summarize the part of the phenomena which your modeling addresses. \* **What is the key scientific question?** Clearly articulate the question which your modeling tries to answer.
- **What was our hypothesis?** Explain the key relationships which we relied on to simulate the phenomena.
- **How did your modeling work?** Give an overview of the model, it's main components, and how the modeling works. "Here we ..."
- **What did you find? Did the modeling work?** Explain the key outcomes of your modeling evaluation.
- **What can you conclude?** Conclude as much as you can *with reference to the hypothesis*, within the limits of the modeling.
- **What are the limitations and future directions?** What is left to be learned? Briefly argue the plausibility of the approach and/or what you think is essential that may have been left out.

**Instructions:** write down your answer to each of those questions (1-2 sentences each, max!). When you're done, stick the sentences together... Now you have an abstract!

## Modeling Practice

### Guidance for paper writing

There are good guidelines for structuring and writing an effective paper (e.g., [Mensh Kording, 2017]), all of which apply to papers about models. There are some extra considerations when publishing a model. In general, you should explain each of the steps in the paper:

- **Introduction:** Steps 1 2 (maybe 3)
- **Methods:** Steps 3-7, 9
- **Results:** Steps 8 9, going back to 1, 2 4

Here are some great materials to help you with paper writing:

- Ten Simple Rules for Better Figures
- How to Write Strong and Effective Figure Legends
- How to Create a Consistent Writing Schedule: 10 Tips for Writers

The audience for all of this should be experimentalists, as they are the ones who can test predictions made by your model and collect new data. In this way your models can impact future experiments, and future data can then be modeled (see modeling process schematic below). Remember your audience - it is always hard to clearly convey the main points of your work to others, especially if your audience doesn't necessarily create computational models themselves.

In addition, you should provide a visualization of the model, and upload the code implementing the model and the data it was trained and tested on to a repository (e.g. GitHub and OSF).

### Suggestion

For every modeling project, a very good exercise is to **first** write a short, 100-word abstract of the project plan and expected impact. This forces focussing on the main points: describing the relevance, question, model, answer and what it all means very succinctly. This allows you to decide to do this project or not **before you commit time writing code for no good purpose**. Notice that this is really what we've walked you through carefully in this guide!