

The background is a dark blue field filled with numerous short, diagonal streaks in shades of cyan, yellow, and orange. On the left and right sides, there are faint, white dashed lines that form irregular, wavy patterns, resembling topographic contour lines or abstract data paths.

# Fit for purpose development under climate change

Wendy Parker

Virginia Tech

[wendyparker@vt.edu](mailto:wendyparker@vt.edu)



# Overview

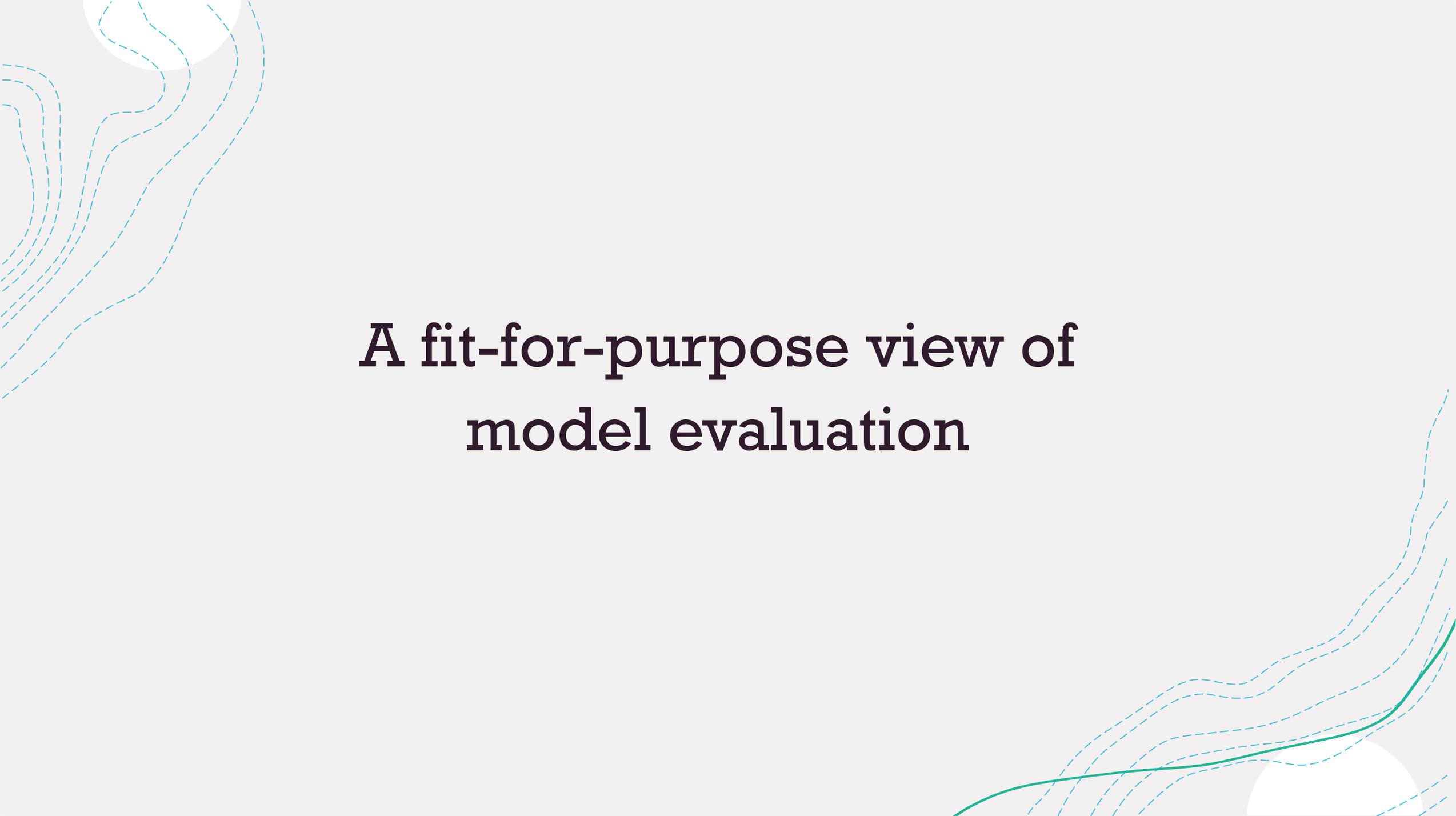
---

Part I. A fit-for-purpose view of model evaluation

---

Part II. Climate change, FEWS and fitness-for-purpose





# A fit-for-purpose view of model evaluation



## A common view of modeling: The mirror view

- + Models are representations
- + Model quality = representational fidelity
- + Improve model by adding detail and/or increasing fidelity
- + Long-run aim: perfect model



# Some limitations of the mirror view

1. We build models not just to represent, but to achieve other scientific & practical goals.
  - + predicting X, understanding Y
  - + protecting people from Z, improving decisions about D...
2. Whether a goal is achieved almost always depends on more than how the model represents the world. It also depends on:
  - + the methodology employed, including how we get initial conditions, computing power, bias correction, communication of results, etc.
  - + features of the model user, such as their background knowledge
3. Models don't need to represent the world with perfect fidelity to be useful; a perfect model might even be less useful for some purposes!

# A better view of modelling

## Mirror view

- + Models are representations
- + Model quality = representational fidelity, i.e., fit with world
- + Improving a model means adding detail and/or increasing fidelity
- + Long-run aim: perfect model

## Fit-for-purpose view

- + Models are representational *tools*
- + Model quality = *fitness-for-purpose*, which depends on fit with world, user, method, *jointly*.
- + Improving a model means adjusting it to better serve purposes of interest
- + Long-run aim: a set of models that is fit for a wide range of purposes of interest



# It makes a difference to practice

## Mirror view

- + General / overall performance metrics
- + Model uncertainty is uncertainty about 'true' equations and parameter values

## Fit-for-purpose view

- + Purpose-tailored performance metrics
- + Model uncertainty is uncertainty about which equations and parameter values would give results that meet our needs
  - + best parameter value might not be 'true' value, given other model imperfections
  - + best model version might be different for different purposes

# The fit-for-purpose view: not just for a model

- + A single model
- + An ensemble of models
- + A forecast system
- + A forecast
- + A dataset
- + ...

*If this tool/resource is used by this user, following this methodology, will purpose P be achieved?*



# **Climate Change, FEWS and Fitness-for-purpose**

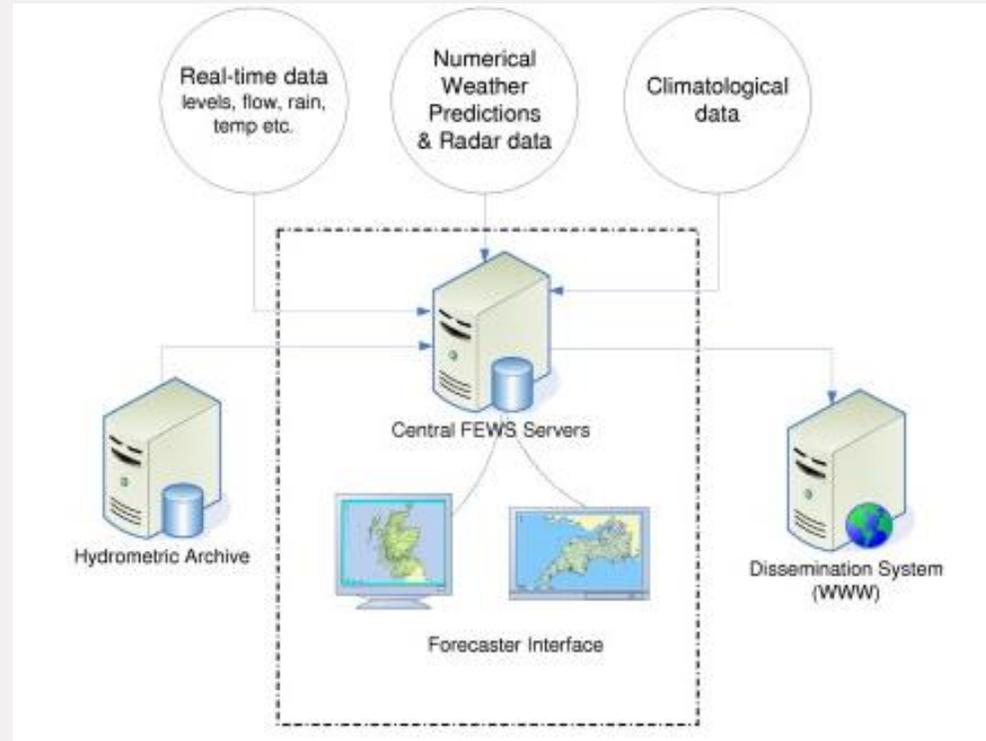
# Evaluating fitness-for-purpose

## + **Ingredients/Design** – do the components of our forecast system have the right features for the job?

- + use right kinds of data, given purpose?
- + models include key physical processes, given purpose?
- + sufficient resolution, given purpose?
- + ensembles to sample uncertainty?
- + manageable user interface?
- + ...

## + **Performance**

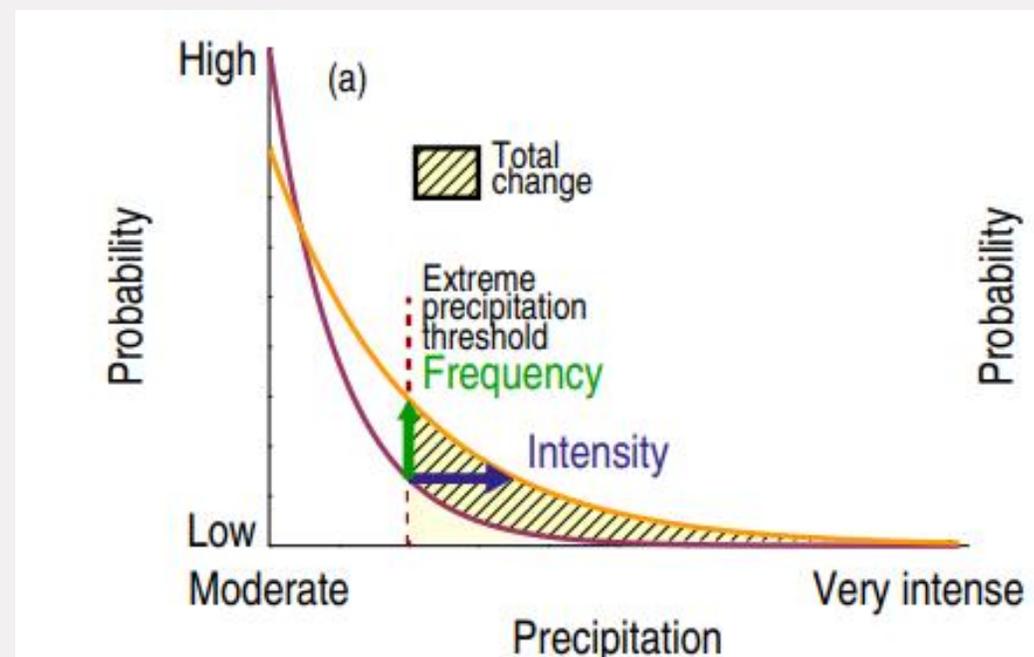
- + scores on relevant performance metrics (verification statistics)
- + hindcast tests
- + situational biases seen in experience
- + salient successes and busts



Werner et al. 2013

# Adding climate change to the mix

- + **Forecast system purpose:** Provide timely warnings of flooding events in our area to help protect lives and property
- + The range of events that our forecast system needs to handle may expand under climate change, e.g. with more extreme events, new weather patterns, ...
- + Existing forecast system may have reduced fitness for purpose, if it was not designed to handle the new types of events.



Myhre et al 2019

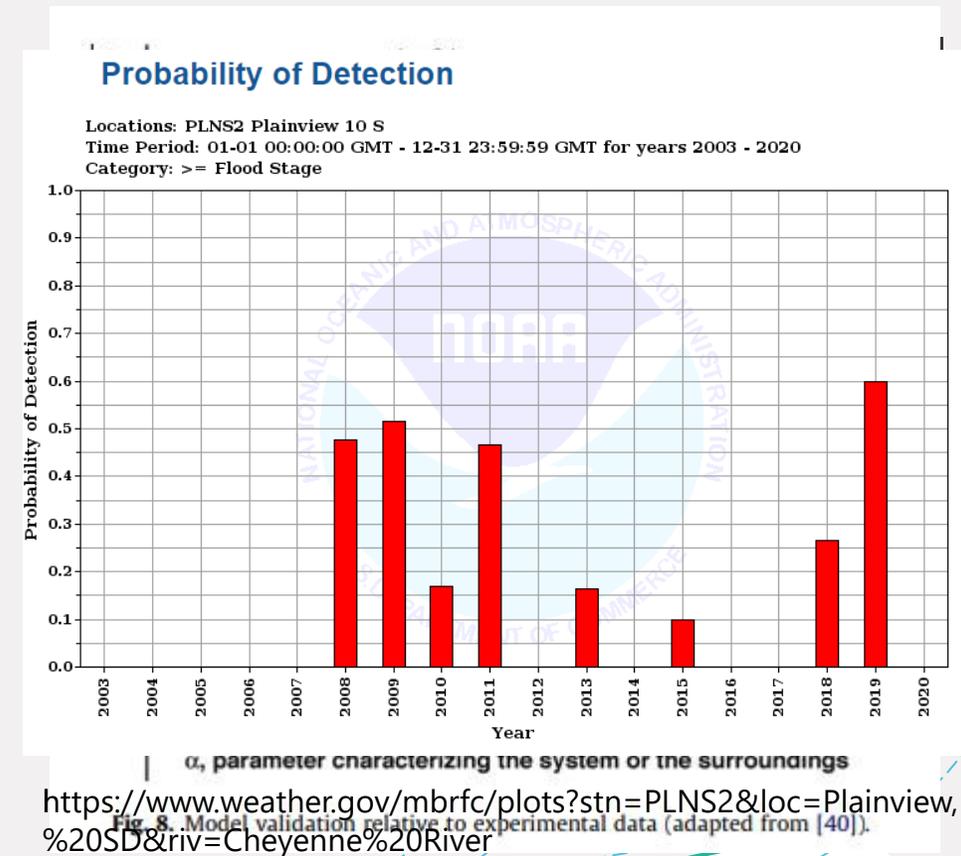
# Fit-for-purpose evaluation under climate change

## + Ingredients/Design

- + Can we *expect* the fitness of our forecast system to be reduced, given an expanded application domain? Which components? How can we remedy this?

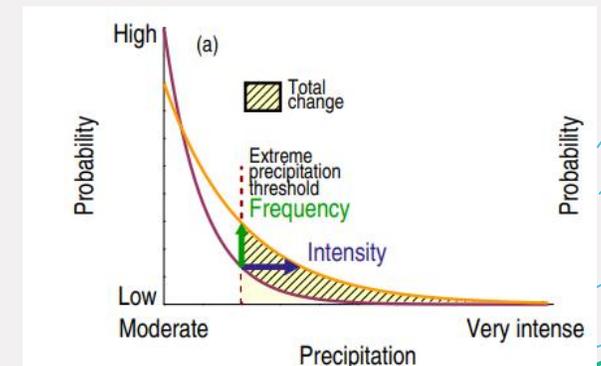
## + Performance

- + Are verification statistics absent for types of events that will materialize; is performance changing?
- + Adjusting of model guidance in light of our own experience/judgement may also need updating.



# Questions for reflection...

- + How do we evaluate the fitness for purpose of our forecast system now?
- + How might our purposes (e.g. types/range of events) change in future?
- + How might components of our forecast system need adjusting to maintain/increase fitness?
- + Can we start working on those adjustments now? What are priorities?
- + How will we evaluate the fitness of our updated system?



## Some references

**Myhre**, G. et al. 2019. "Frequency of extreme precipitation increases extensively with event rareness under global warming", *Scientific Reports*.

**Parker**, W.S. 2020 "Model Evaluation: An Adequacy-for-Purpose View", *Philosophy of Science*.

**Roy**, C.J. and W.L. **Oberkampf** 2011. "A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing", *Computer Methods in Applied Mechanics and Engineering*.

**Werner**, M. et al. 2013. "The Delft-FEWS flow forecasting system", *Environmental Modelling and Software*.