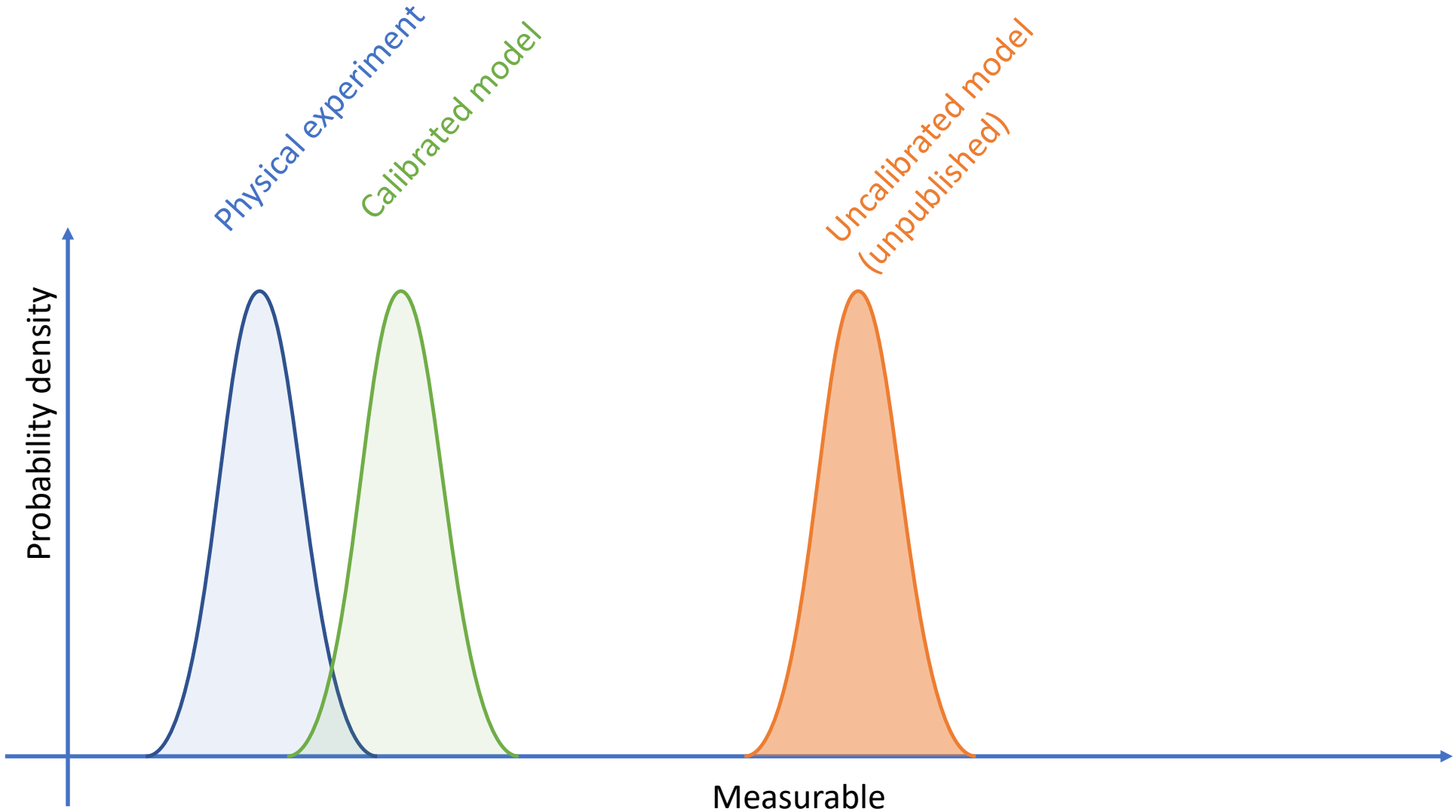
Concept and vision: FluidFlower benchmark

Agenda today

- Jan Martin Nordbotten: Vision, concept and process
- Martin Fernø: Overview of data and processes
- Bernd Flemisch: Reporting and communication
- Questions, remarks, suggestions, and clarifications

A visual to get everyone on the same page…

Caricature:
A physical model and a simulation



Published modeling studies do not accurately reflect the forecasting skill for CO_2 storage

What can we learn from weather/climate forecasting

Perspective | Open Access | Published: 16 July 2021

Standard assessments of climate forecast skill can be misleading

James S. Risbey, Dougal T. Sq
Didier P. Monselesan, Thomas S.
Tozer

Nature Communications **12**, Art

Abstract

Assessments of climate fore
perspective, we use forecast
bias-correction on skill. Mar

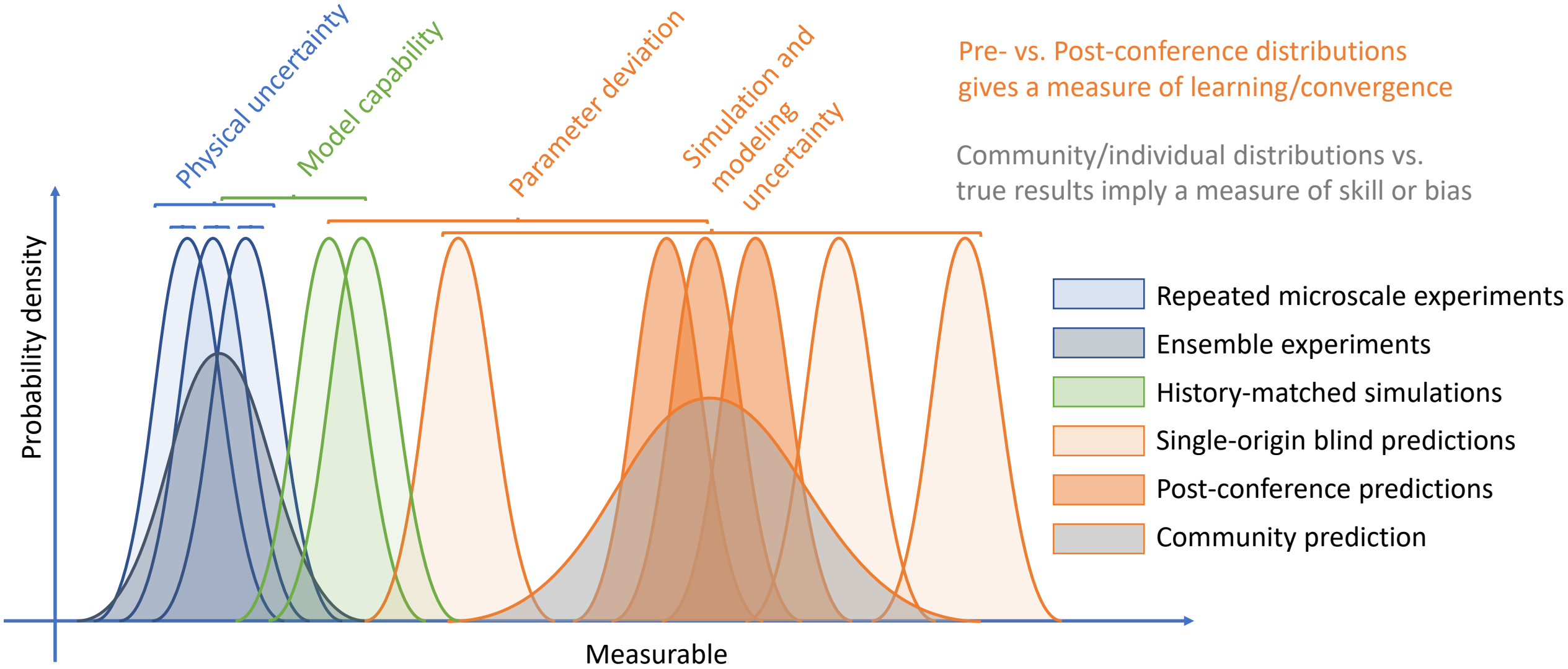
Lessons for forecast skill assessment

Progress towards the goal of improving climate forecasts needs to be quantified by fair skill assessments. At present, seasonal forecasts are typically assessed by the institutions that produced them and are not easily comparable across institutions. In many papers and hindcast/forecast archives the bias correction method is not clearly stated or not consistent across models[10], rendering any subsequent skill assessment or comparison of dubious use. To remedy this, skill assessments should be performed by open communities on open platforms. Model groups need to provide the raw hindcast outputs so that all models can be subject to identical bias-corrections and so be meaningfully compared.

Due to our continuous learning (and review), a fair assessment of predictive capabilities is not possible within a single research group

A community effort is the only viable way

Vision:
Quantify all uncertainties and bias from prediction to flow



Pre- vs. Post-conference distributions gives a measure of learning/convergence

Community/individual distributions vs. true results imply a measure of skill or bias

Sample questions we can address

- What is the underlying uncertainty of coupled multi-phase flow relative to the macro-scopic parameters we measure?
- Can numerical simulation tools adequately simulate these processes?
- What is the uncertainty/bias of lab measured parameters relative to in situ parameters?
- Can our community reliably predict the coupled processes associated with CO_2 storage?
- What is the added value of multi-group assessments for prediction quality?
- Is our perceived uncertainty of the modeling and simulation process an adequate approximation of true uncertainty?

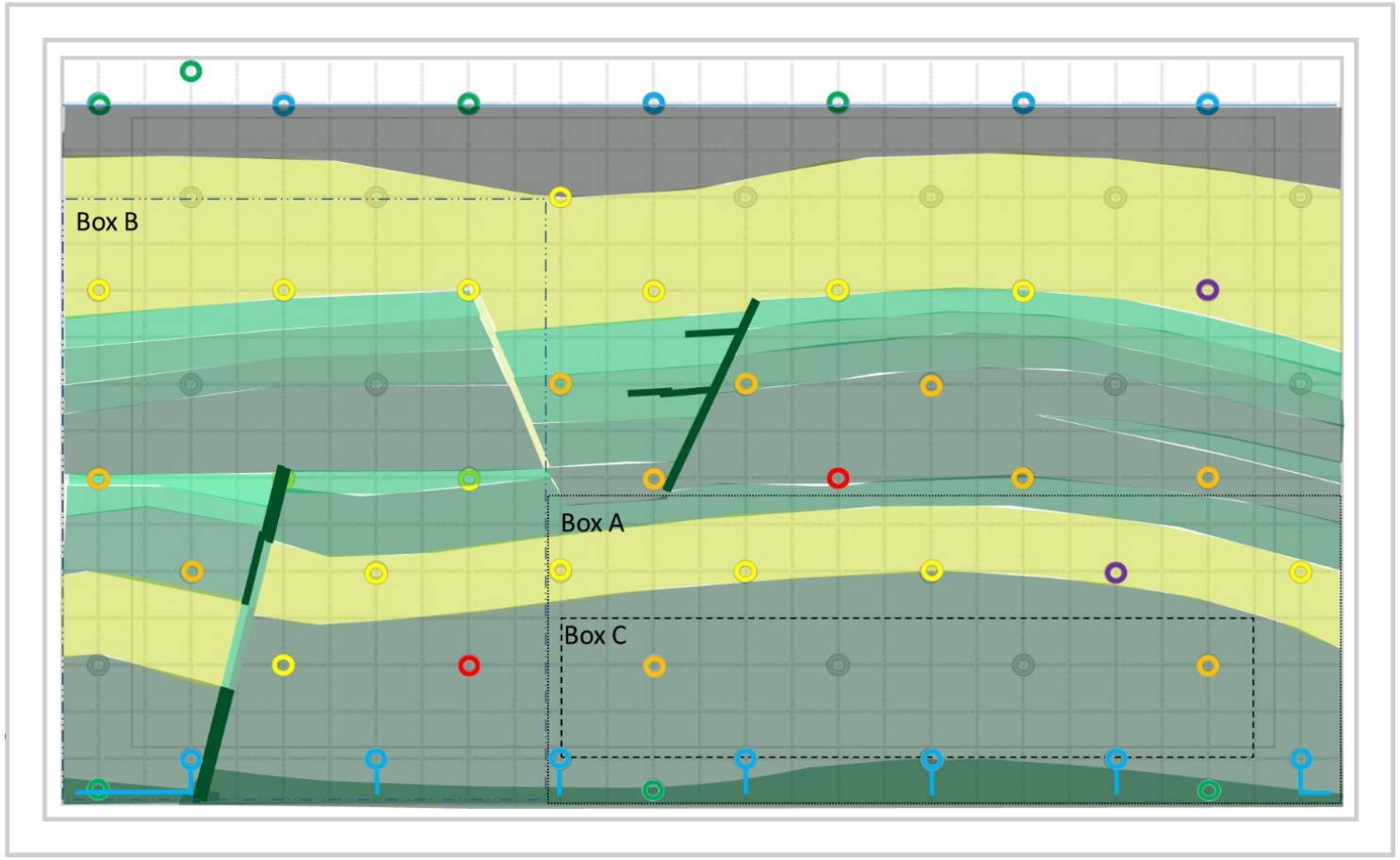
The results will necessarily depend on both context and measureable

Two sets of measurables

- Pointwise results from numerical simulation
  - Allows for quantitative comparison and analysis of inter-group variability
- Aggregate results related to key CO2 storage questions
  - Allows for comparison with experiment
  - Allows for «human judgement»
- For aggregate results we ask for
  - Most likely mean value (P50 of mean)
  - 10% and 90% confidence interval for mean value (P10 and P90 for mean)
  - Most likely variability in repeated experimental results (P50 of standard deviation)
  - 10% and 90% confidence interval for variability (P10 and P90 for standard deviation)

Key CO_2 storage questions

- Mechanical stability of overburden
  - Maximum pressure at sensors
- Time of maximum leakage risk
  - Time of maximum free-phase CO_2 in Box A
- Ability to accurately model physics
  - Phase distribution in Box A
- Ability to handle uncertain geological features
  - Phase distribution in Box B
- Importance of convective mixing
  - Integral measure of mixing in Box C
- Integrity of overburden
  - Total mass of CO_2 in the seal within box A.



Setup

- Dept. Physics (UiB) conducts all experiments, and are blind to everything else.
- Dept. of Civil and Environmental Engineering (MIT) conducts an «enhanced blind» study, with access to additional analog data, but blind to the actual experiment and everyone else.
- Dept. of Mathematics (UiB) conducts the history matching (with access to experiments) and are blind to external simulations.
- IWS (Stuttgart) coordinates the international benchmark study and manages communications.
- External participants conduct their own simulations:
  - Initially blind to both the experiments, history matching, and each other.
  - Later only blind to experiments and history matching (see below).
- We will organize a virtual conference for external participants where simulation results are shared and discussed
  - During this conference, Magne Jørgensen from Simula Research Laboratory will also gather «soft» information in order to better distinguish various sources of (initial) uncertainty.
- We complete the study with a physical/hybrid workshop near Bergen 1-2 February 2022, where we present and discuss all results, and make a detailed publication plan.