# The potential for models to fill data gaps and enhance decision making

David Byrne (NOC)
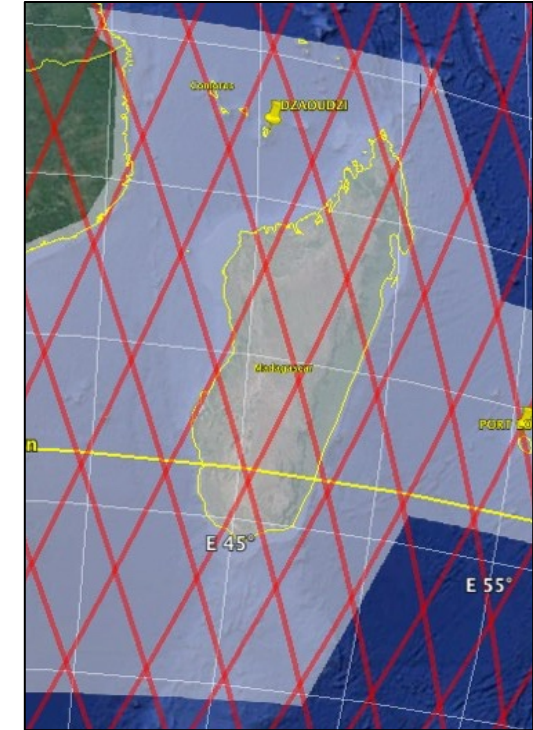
11th and 12th February 2020, Antananarivo, Madagascar

Observations Are Great

- Observations <u>attempt</u> to tell us what's really happened.

- In-situ and remote observations both have problems

- In-situ data can be sparse in space

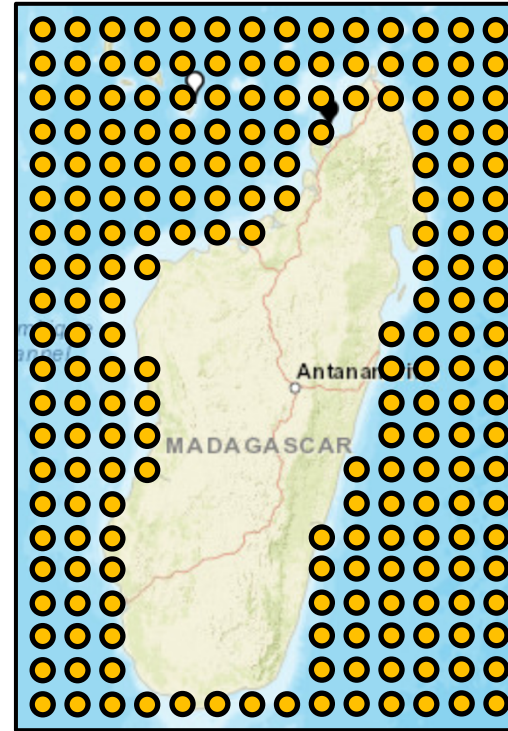- Satellite data can be sparse in time. Good for creating climatological data.
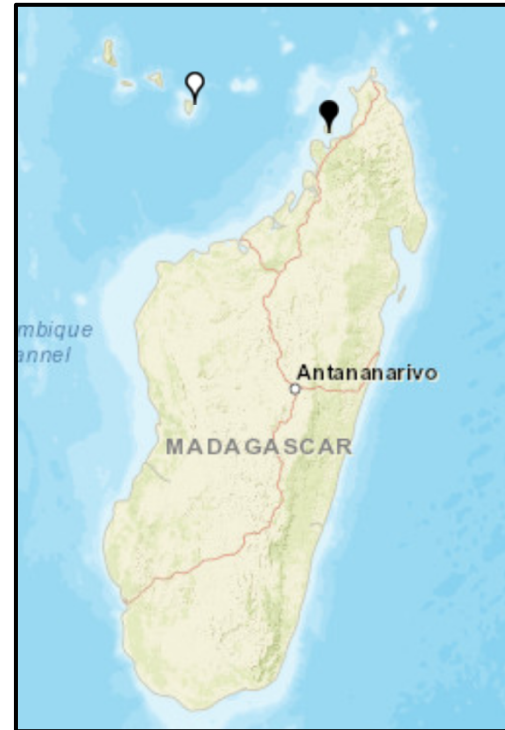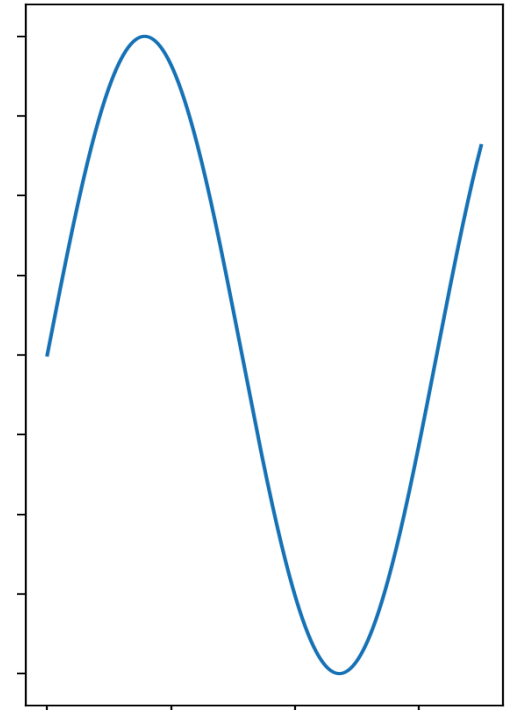

In-situ
e.g. Tide Gauges

Remotely sensed
e.g. Satellite

Gaps in Data

- Ideally: observations everywhere, continuously.

- Having both (or either) is very difficult.
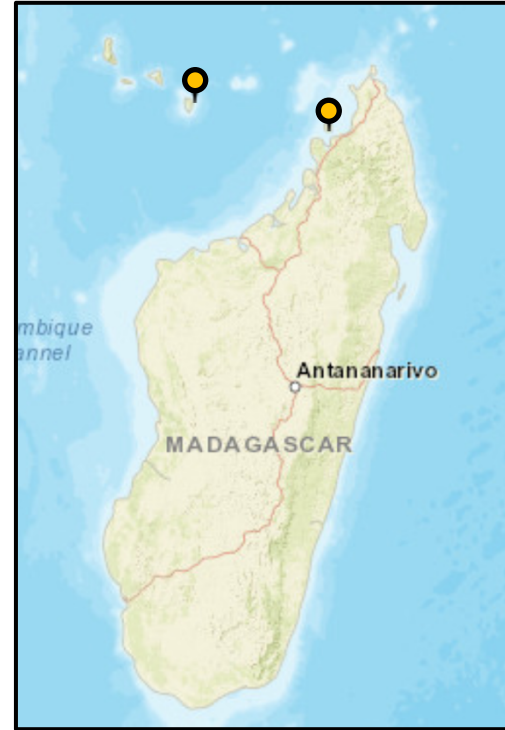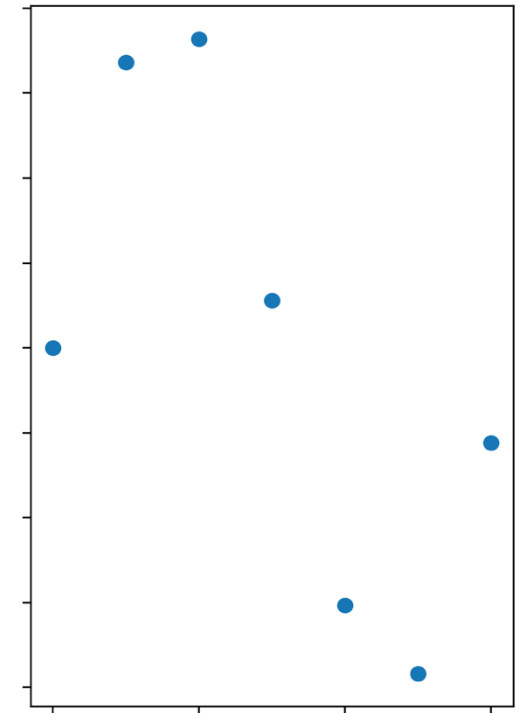

Space

Time

Gaps in Data

- Ideally: observations everywhere, continuously.

- Having both (or either) is very difficult.

- Prohibitively expensive and impractical (e.g. data storage)

Models offer a _partial_ solution for filling in gaps

Space

Time

What is a model?

A model is a <u>representation</u> of nature used to understand, forecast or demonstrate

<u>Examples of model types</u>

1. Physical

2. Statistical

3. Parametric

4. Numerical

None are perfect!!

What is a model?

A model is a <u>representation</u> of nature used to understand, forecast or demonstrate

<u>Examples of model types</u>

1. Physical

2. Statistical

3. Parametric

4. Numerical

None are perfect!!

A Toy Ocean: Inputs and Outputs

Input Examples

- Winds

- Atmospheric Pressure

- Temperature Flux (sun, air temperature)

- Salinity Flux (rainfall, river outflow)

- Nutrients

Output Examples

- Sea surface height

- Currents (speed/direction)

- Ocean Temperature

- Ocean Salinity

- Chlorophyll/phytoplankton

- Waves

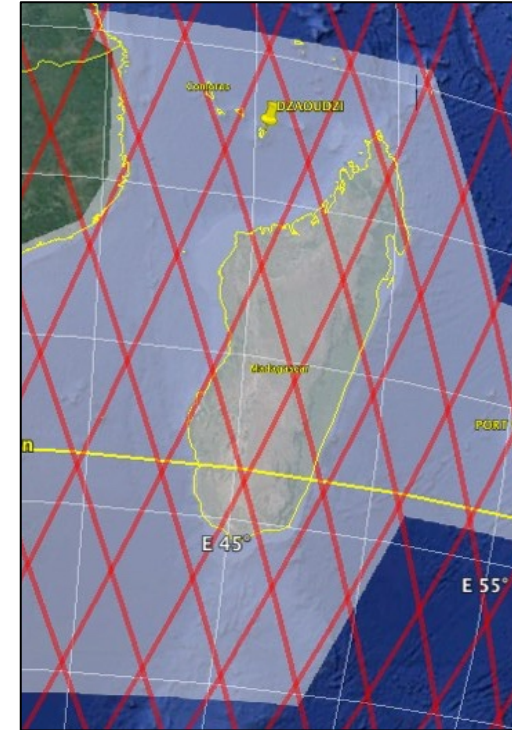**What you get out depends on what you put in**

Numerical Model

- "Computer Simulation."

- Built using physics, mathematics and programming.
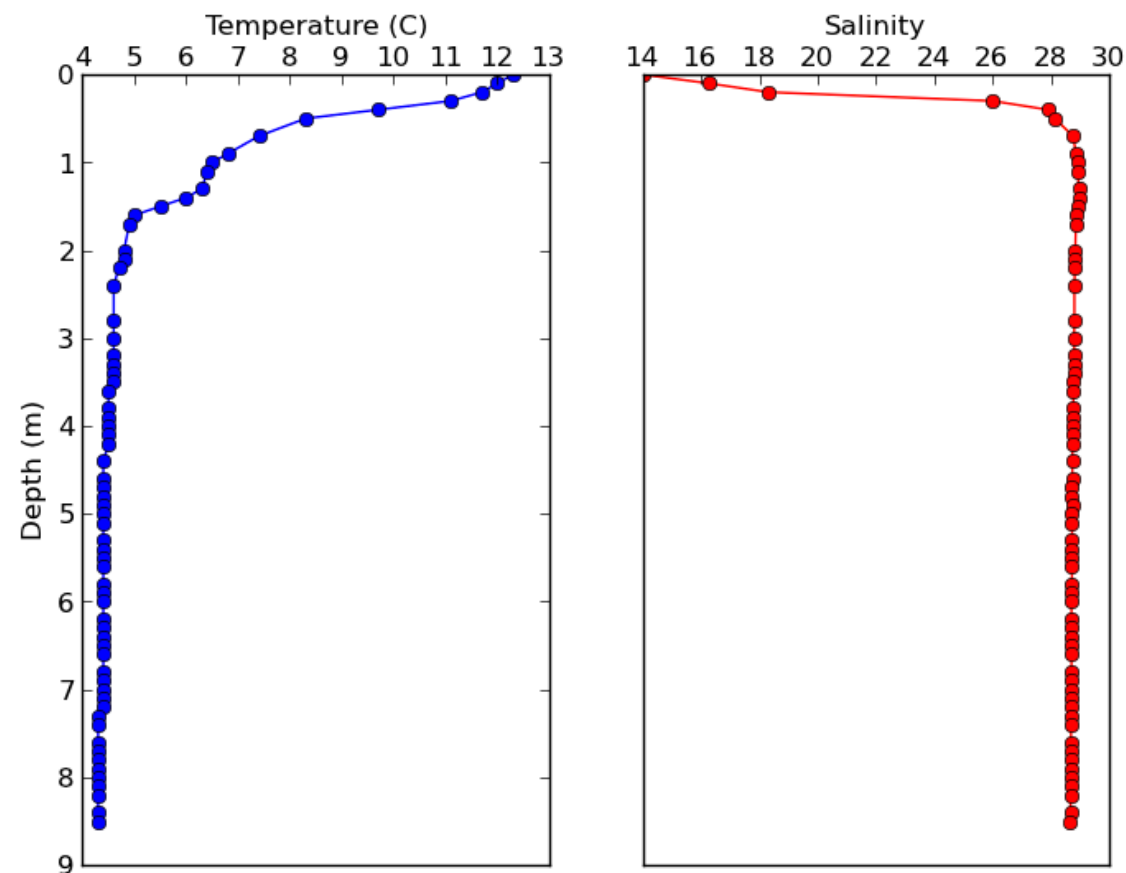
(NASA)

Filling Data Gaps
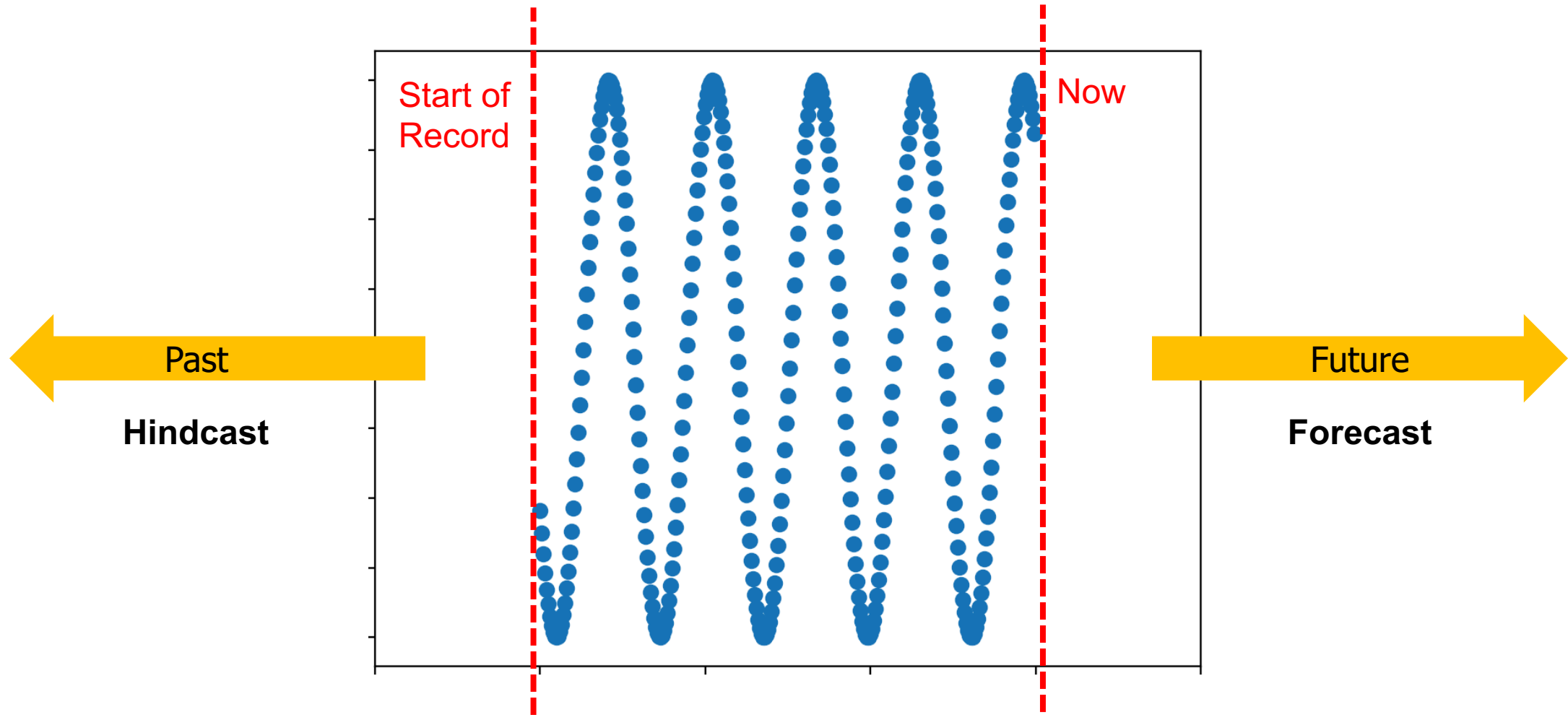

In-situ
e.g. Tide Gauges

Remotely sensed
e.g. Satellite

The Third Dimension

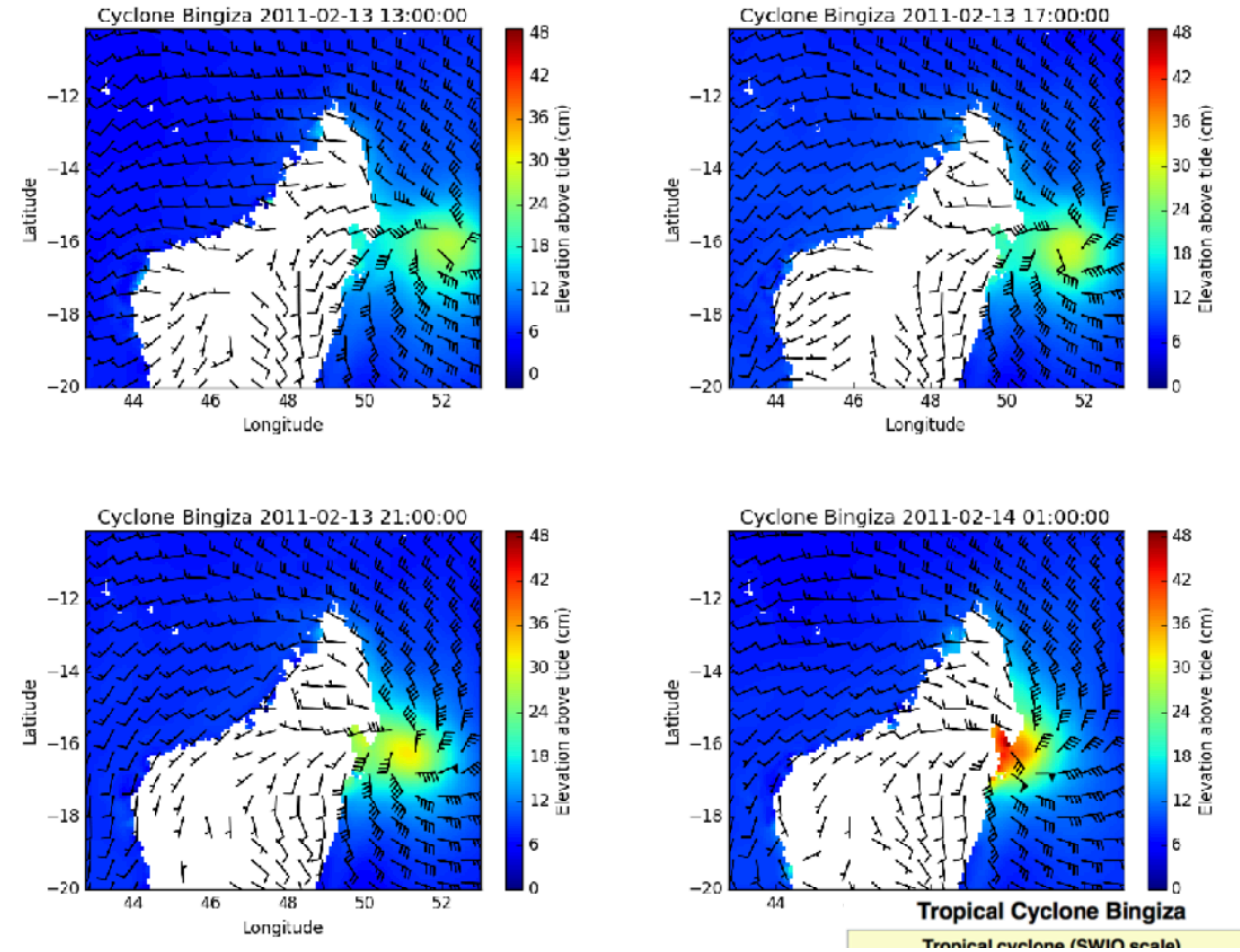Models can supply 3D data. Satellites only see surface data.
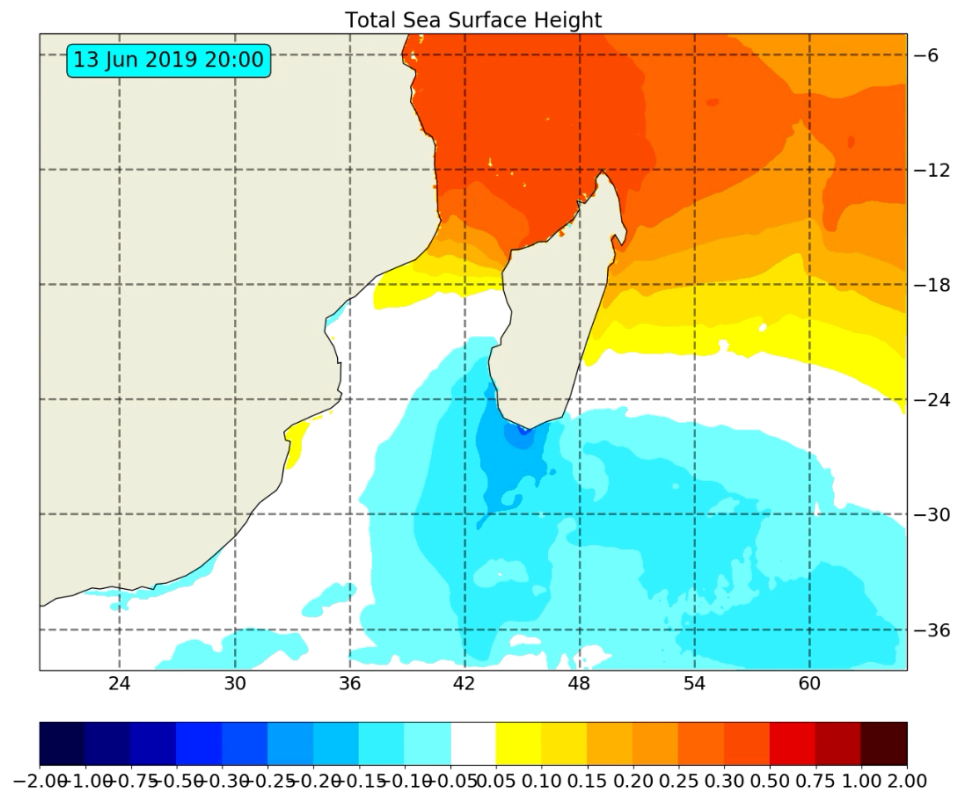
The Biggest Gaps

Forecasting: CRISC

Estimates of variables for some future period

Hindcasting

Simulations of the past

Can be used to:

- Understand what happened during an event (e.g. a storm)

- Understand and quantify risks using multiple simulations

- Create climatological data.

- Similar to CRISE use cases except longer time period, more data

Hindcasting

Some hindcast data exists, easy to provide:

- Global model, using NEMO model

- Available for 1960 – 2015

- Daily or monthly averaged data

- Inputs: Winds, atmospheric pressure, tidal forcing, bathymetry, temperature, salinity. No river input.

- Output variables available: Water level, currents, temperature, salinity.
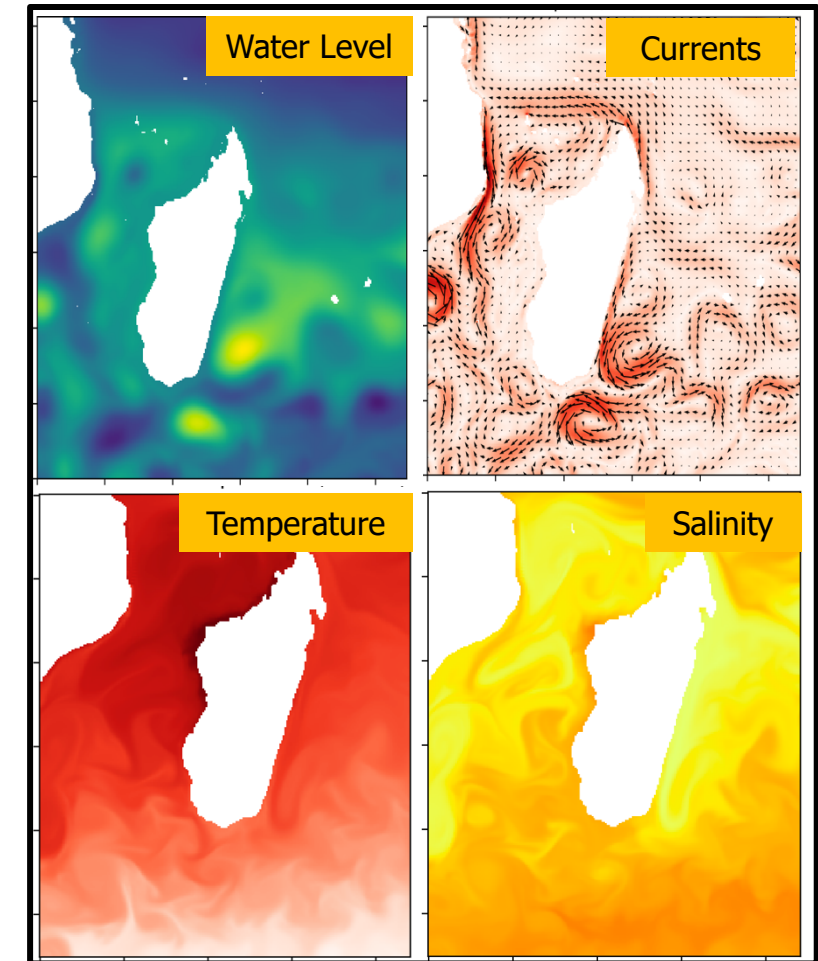
Hypothetical Modelling

Model can be modified to test how things change under different conditions/scenarios.

Example

- Storm surges/waves under different sea level rise and weather scenarios

- Run a model with as if sea level has risen some amount

- Run a model with modified version of real atmospheric data.

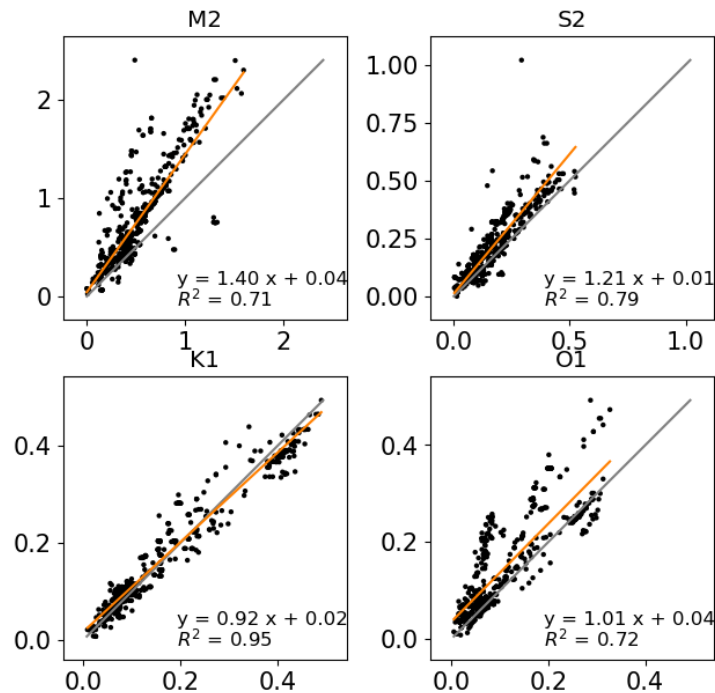- See how past storm surges/waves would have been different
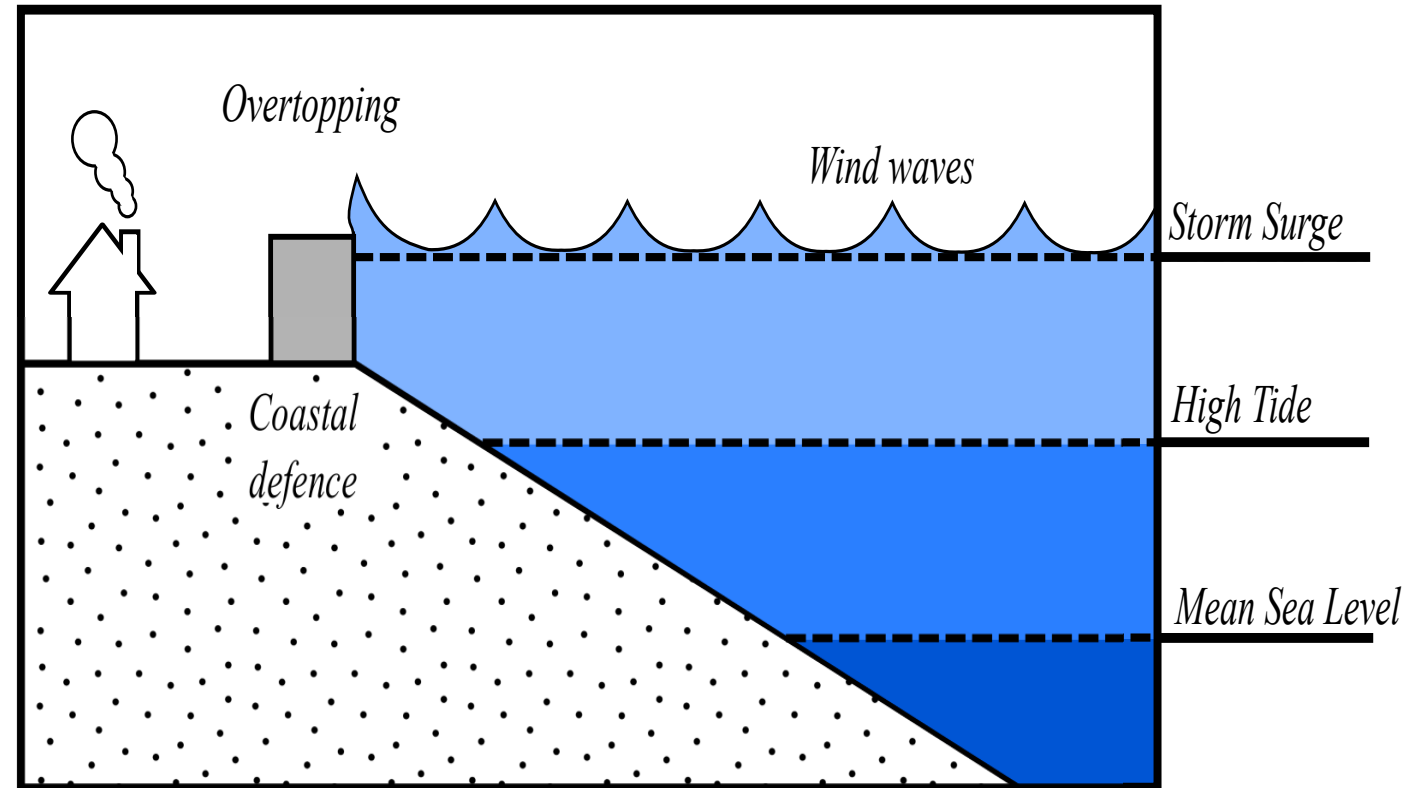
Limitations: Accuracy and Validation

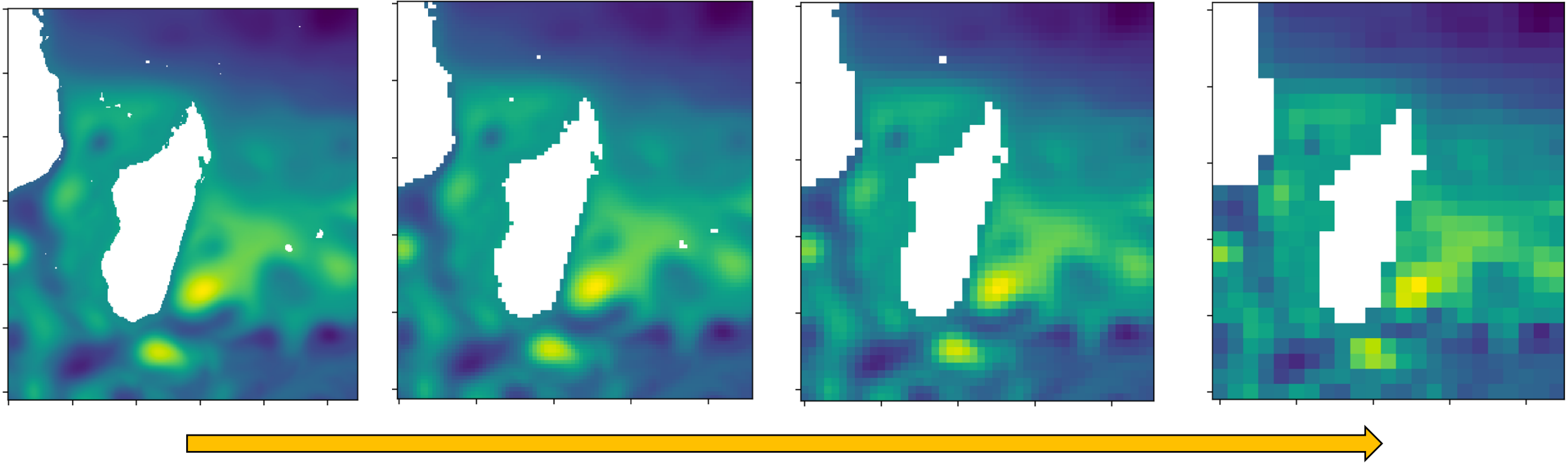**Models can have large, consistent errors which must be quantified**

A model must be tested for accuracy by comparing to as many observations as possible

Can use statistics like RMSE, MAE, correlations.

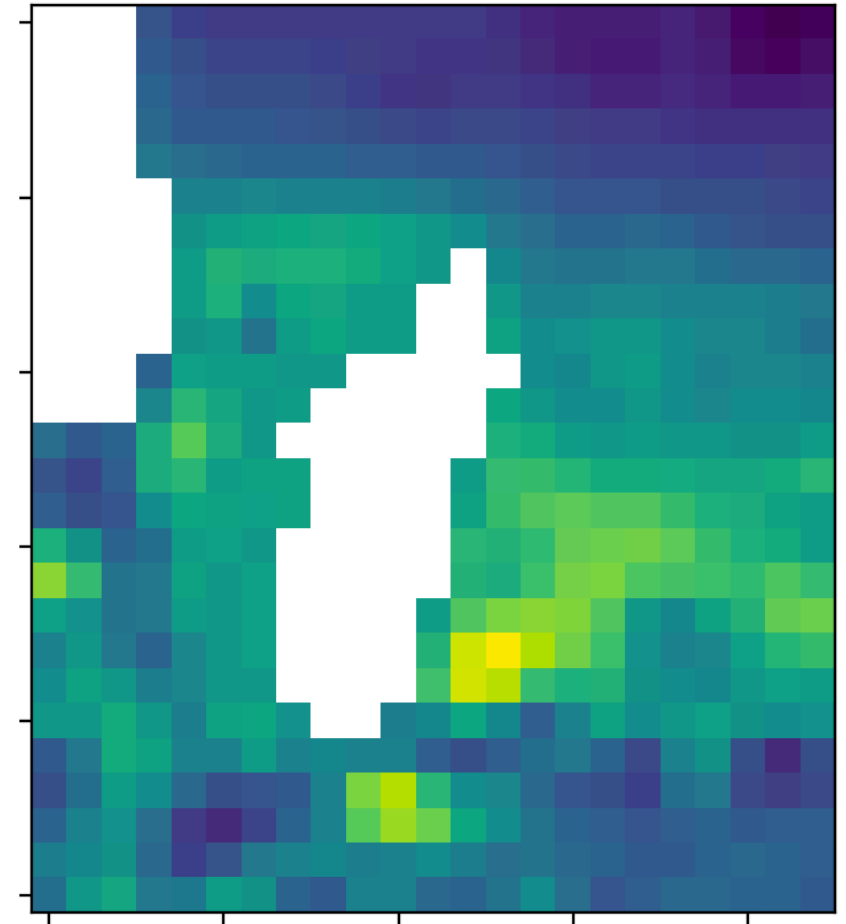# Limitations: Resolution

Less computation time
Less data storage
Less detail/reduced accuracy

Limitations: Representativity

Model data is not at 'points' like many observations. Instead it is in 'cells'.

For some variables, can be a challenge to compare directly to observations or to interpret correctly.

Thank you for listening.

Summary

1. Observations are great.

2. Model data can be used to fill in where observations aren't available.

3. Inputs to models can be modified to test different things.

4. Models can be used to look at both the future and the past.

5. Models can be used to look at hypothetical scenarios.

6. Model data has limitations that must be kept in mind: accuracy, resolution and representativity.

7. Decades of model data is available for use.