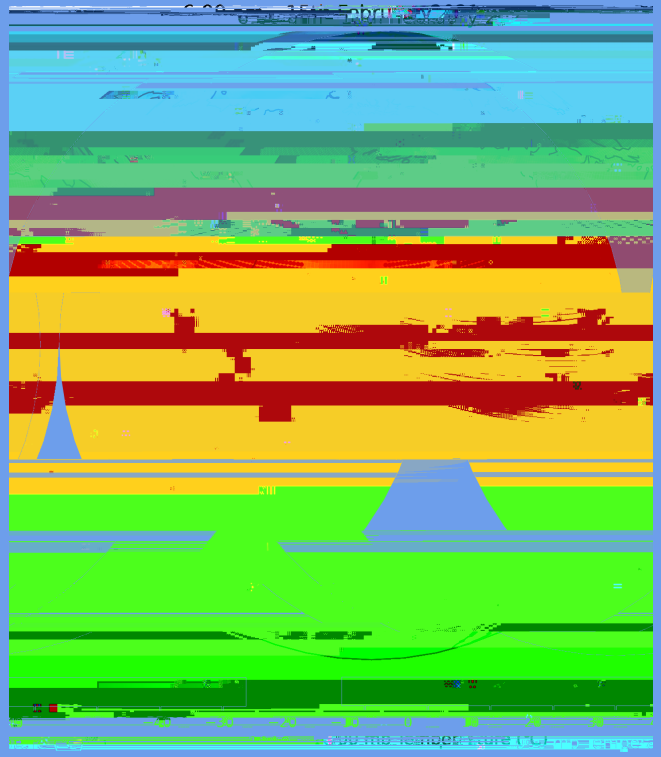
Approaching meaningful insight into the climate system. My experiences so far.



Tom Keel

About me

Sept 2015 to June 2018 – Geography & Geocomputation (KCL GEOG)

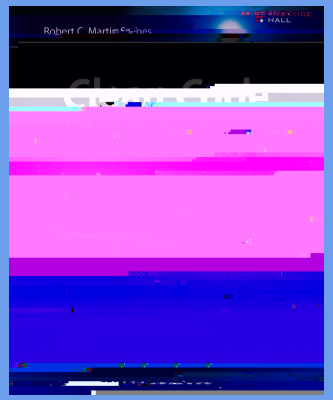
Sept 2018 to Sept 2019 – Spatial Data Science and Visualisation (UCL CASA)

Sept 2019 to Sept 2020 – *Bits and Bobs*

Sept 2020 to right now – London NERC DTP PhD (UCL GEOG)

Experience:

Data scientist, GIS software developer.

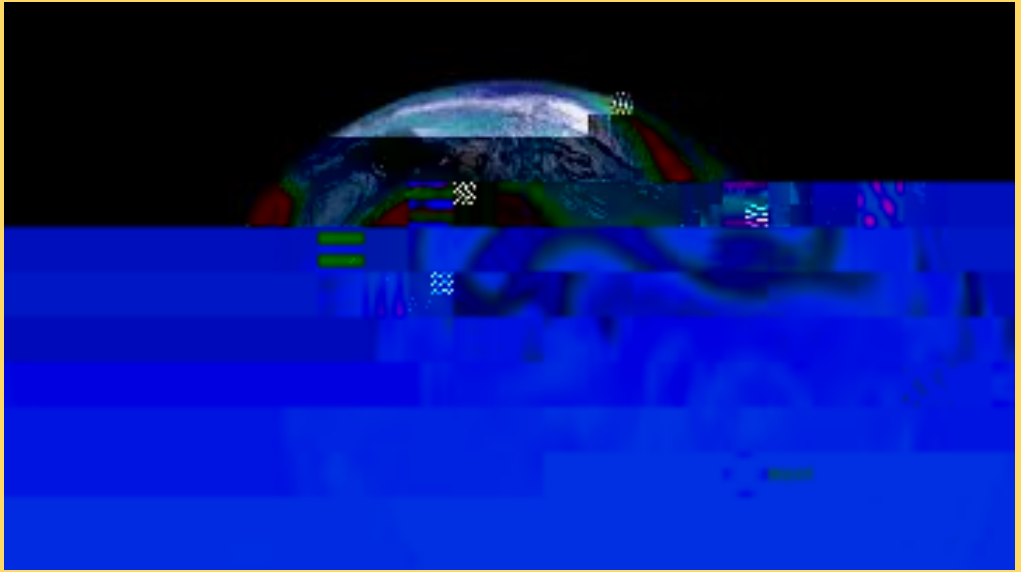


About this presentation

About my research
About climate
About meaningful insight
My approach to my project
Role of open-source
Summary of key points

About my research

Title: A shifting jet-stream in a changing climate: Exploring the response of the polar jet-stream in the Northern Hemisphere to various climate futures.



What are jet-streams?

- **Fast and fluid:** Streams of fast wind which occur in regions known as jet streaks
- **High and broad:** 8-12 km in between the troposphere and stratosphere
- **Complex and unknowable?**

Jet-streams – Why are they important?

Dynamical properties:

- Form at the boundaries between “air-masses”
- Create cyclones and anticyclones at surface

Jet-streams as patterns:

-

Jet-stream – Link to weather

Storm tracks!

Jet-streams – Impacts

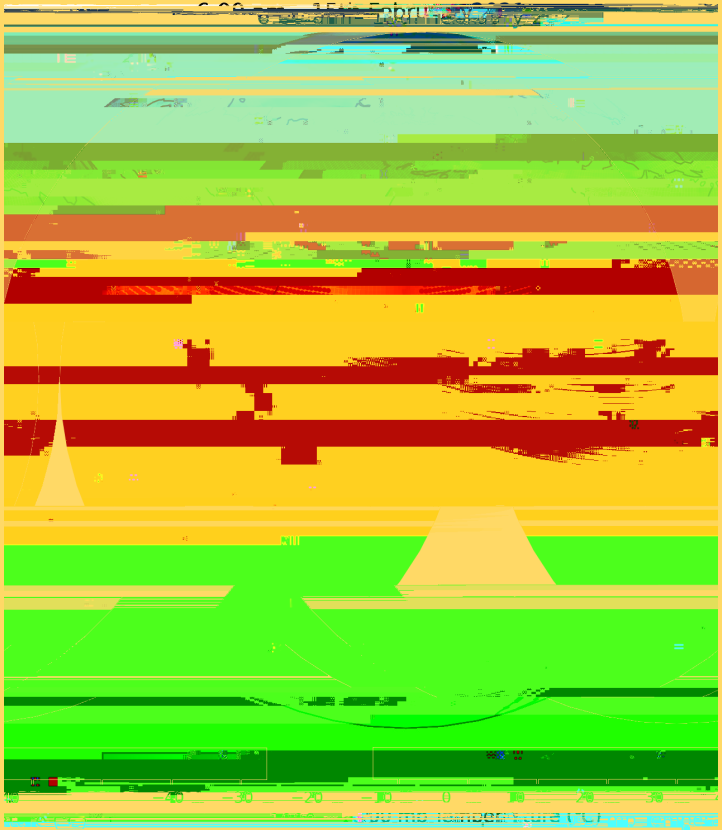
Cold-air transport:

- Texas snowstorm February 2021
- Beast from the East(s)

Persistence:

- European heatwaves (2003, 2010, 2021)
- Droughts

Note: Northern Hemisphere!

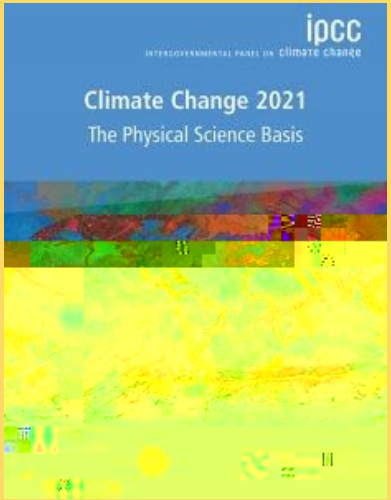


Jet-streams – Changes

Jet-streams – IPCC AR6

“The extratropical jets and cyclone tracks have likely been shifting poleward in both hemispheres since the 1980s with marked seasonality in trends (medium confidence)”. (IPCC AR6 2.3.1.4.3)

“There is low confidence in projected poleward shifts of the Northern Hemisphere mid-latitude jet and storm tracks due to large internal variability and structural uncertainty in model simulations”. (IPCC AR6 TS-38)



What is change in climate?

What is weather?:

- Short-term variations in atmospheric variables.

What is climate?:

- An average of weather conditions over a particular region.

Obstacles for characterising change:

- **Space-time continuity:** no separation between scales.
- **Interactions:** between different parts of climate system.

What are climate phenomena?

Phenomenon: An occurrence, circumstance, or fact that is perceptible by the senses.

Climate phenomenon: An observable event perceptible in data.

Examples of climate phenomena:

- Monsoons
- Beast from the East
- El Nino, La Nina

What makes them real?

What are climate metrics?

Metrics: A set of numbers that give you information about a particular process or activity.

Climate metrics: indices, statistics and algorithms used to isolate and characterise a given climate phenomena over a given time period and location.

Examples: days with temperatures above 35°C; frost-days; ENSO index

Problems:

- You can show anything with metrics *“all metrics are wrong, but some are useful”*

Gaining meaningful insight

It is a research *community* working together that can get through to ‘meaningful’ insight (*hint: open-source*).

Problems for my research area (jet-streams):

- Jet-streams show signal in various measured variables.
- BUT: We cannot reliably collect observations about them
- AND: We have multiple sources of information about them (past & future).

Climate Reanalysis:

- **Data assimilation strategy:** combination of observations, climate models and interpolation
-

Reading between the lines

We never use one model, one scenario, one metric *INSTEAD* we use an ensemble and read between the lines.

Finding trends:

- **Model agreement:** Probability of X occurring given the ensemble.
- Visualisation is our story-telling device

My process

To approach solving problem we first need to define the bounds of the problem:

- What is the jet? What is a change in the jet?
- Which space-time context is most useful for understanding a jet?

To approach the software required, we first need to define what the community might find useful:

Philosophy A: to create a solution using the least amount of components.

Philosophy B: to create a solution that is decoupled.

Philosophy C: take a reductive view of the problem we are trying solve (it is just data at the end of the day)

First steps: Look at literature

Context from literature

30 metrics found!

Storylines from literature:

1. *Mean* latitude.
2. Waviness.
3. Preferred positions.

Contexts of understanding:

1. Jet as continuous.
2. Jet as segmented.
3. Jet as emergent.

A metric for diagnosing jet-stream latitude over the North Atlantic.

Reductive view:

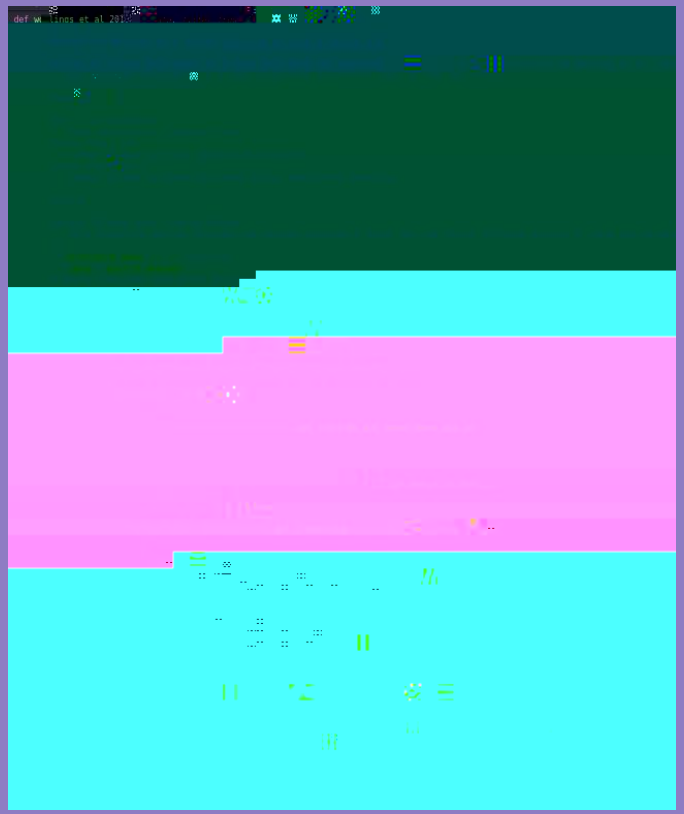
-

Step 2. Write de-coupled functions for metric

- Stick to Single Responsibility Principle (SRP)
- Each function takes same argument: *data*
- Refactor, Refactor, Refactor!

```
# Step 1: Calculate long and/or plev mean
zonal_mean = jetstream_metrics_utils.get_zonal_mean(data)

# Step 2: Apply 8-day lancoz filter
lancoz_filtered = data = jetstream_metrics_utils.apply_lanczos
```

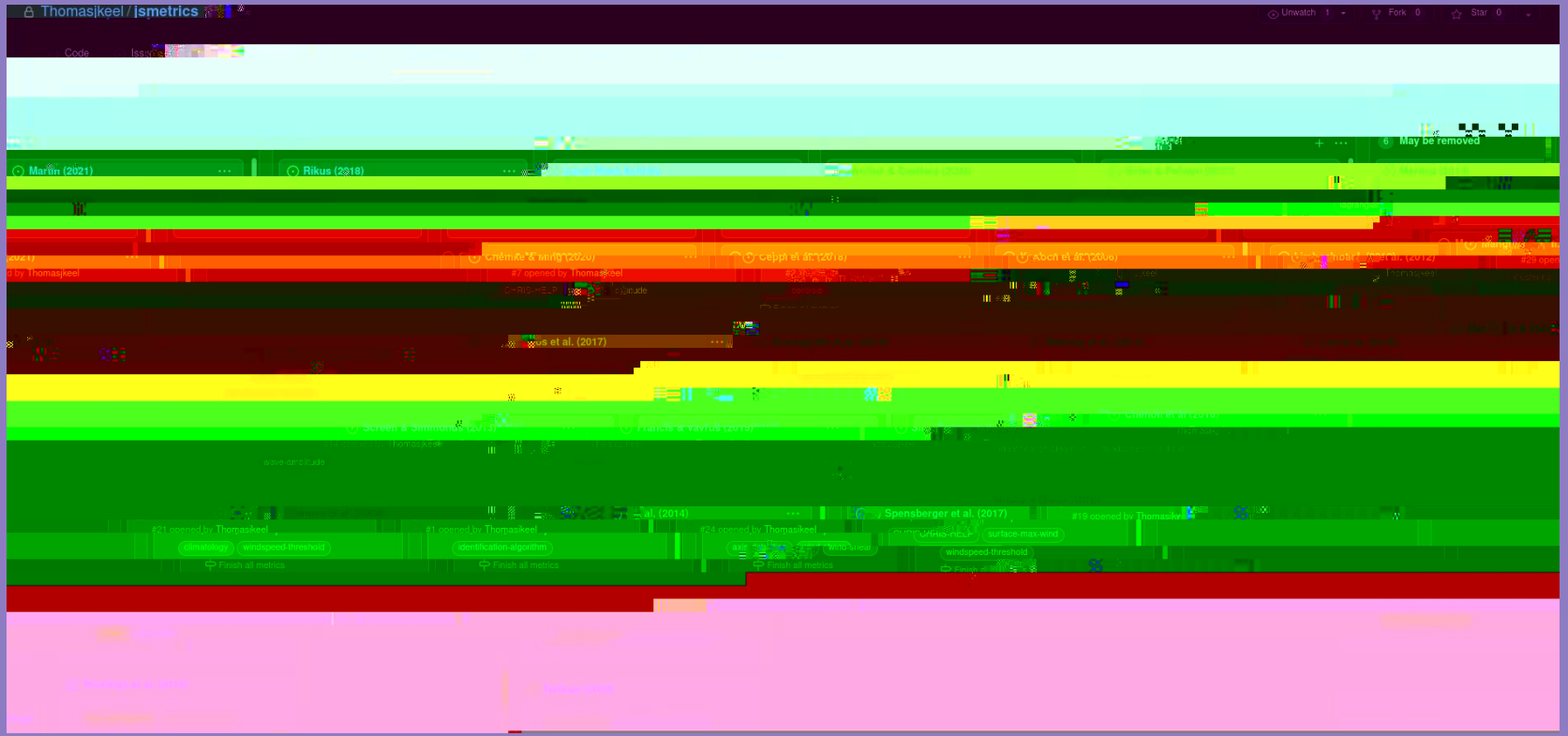


Step 3. Add to Python Module

- Keep it organised, keep it scalable
- Standard inputs, Standard outputs
- Version, manage dependencies (xarray), test

Managing the module with GitHub

Managing the module with GitHub



Initial experiment on JASMIN supercomputer

When: mid-October 2021

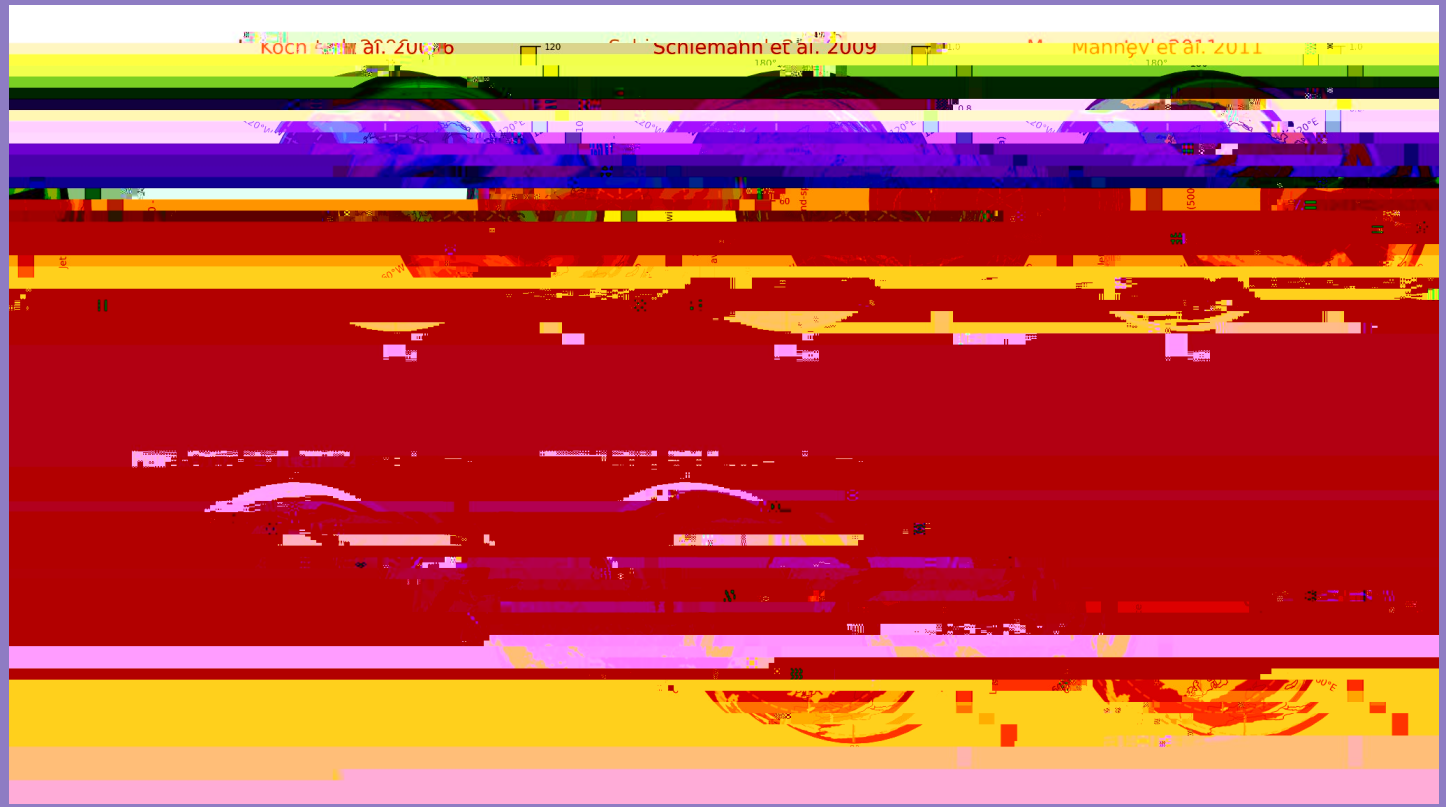
Runtime: 40 hours

Number of metrics: 10

Data:

Initial experiment results

Initial experiment results



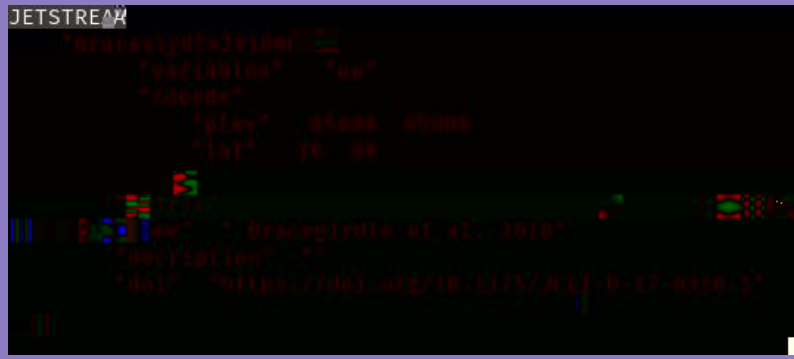
Initial experiment results

Initial research-atron experiment on JASMIN

When: *2nd* February 2022

Runtime: 1 hours

Number of metrics: 1 (Jet-latitude metric)



Data: 187 datasets from 7 modelling groups of projections between Jan 2020 and Jan 2040

Variables: u-component wind

Size: between 0.1-10 GB per dataset

Outputs: Various log, plot on JASMIN

Running with the Research-atron

Findings:

- Data is not always standard
- Slight discrepancy between modelling groups

```
Number of datasets found: 1078
Number of datasets after date range subset (20200101 to 20400101
```

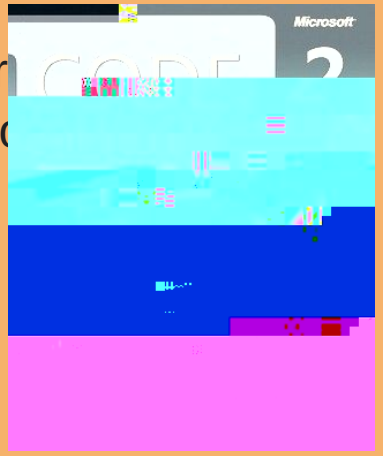
Open-source

Wait a minute, that's socialism!

For climate research:

- Is about getting enough people the right tools (think computer, then think software)
- We have an opportunity to pool problem solving in a new and exciting way.

Programming languages have always facilitated/enhanced climate r
we are not at the ceiling of possibility, or have all the tools we can p
for research.



Example: xclim

- Takes metrics that already exist (in literature), but Pythonises them and makes them run fast.
- Huge inspiration for my own-code and this presentation.
- Thanks to Raquel Alegre, Jamie Quinn and Clair Barnes for getting me involved

One issue with this form of open-source:

- Hiding too much complexity from those who like problem solving

Data-driven vs Theory-driven research

Comments from my own experience:

- Solving complex problems with data when being reductive.
- “Some solution exists!”
- When can we side step theory?

Example of need for theory:

- Woolings et al. 2010. Where knowing something about the jet-streams helps.

BUT: No theory needed when:

- Software used as a tool
- Machine learning algorithms and ‘black-box’ methodologies

Summary of key points

- All metrics are wrong, some are useful.
- Reading between the lines with climate information.
- (Climate) scientists are often self-confessed gate-keepers of (climate) science knowledge but there is a big opportunity to use open-source is an opportunity to get the tools in front of more people.

Any questions?

github: @Thomasjkeel
email: thomas.keel.18@ucl.ac.uk