CATS: Climate-Aware Task Scheduler for environmentally conscious developers

Sadie L. Bartholomew[1], Lincoln Colling[2], Abhishek Dasgupta[3], Anthony J. Greenberg[4], Loïc Lannelongue[5], Thibault Lestang[6], Carlos Martinez-Ortiz[7], Colin Sauzé[8], Adam Ward[8] , Andrew M. Walker[3]

[1] National Centre for Atmospheric Science, United Kingdom. [2] University of Sussex, United Kingdom. [3] University of Oxford, United Kingdom. [4] Bayesic Research, USA. [5] University of Cambridge, United Kingdom. [6] Météo-France, France. [7] Netherlands eScience Center, Netherlands. [8] National Oceanography Centre, United Kingdom

What is the problem?

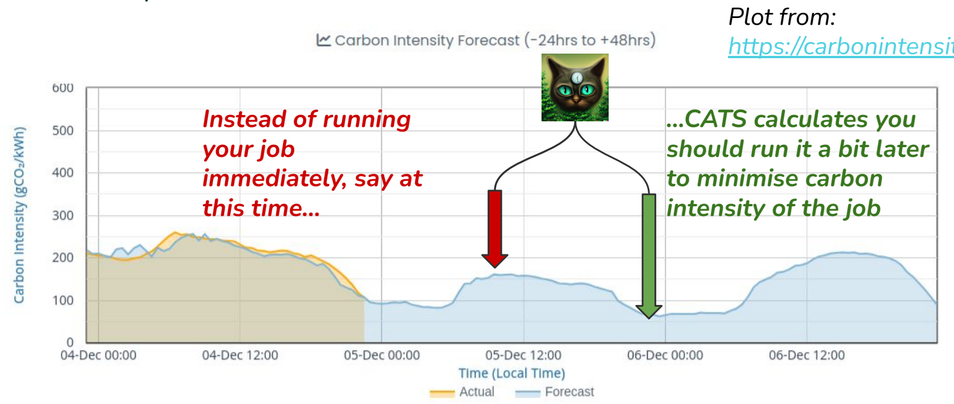
(High performance) computing *always requires energy* (electricity etc.) - how can we do it in a sustainable way to not exacerbate the climate crisis?

Image credits: https://i.imgflip.com/208mpa.jpg, from IT Crowd (Channel 4)

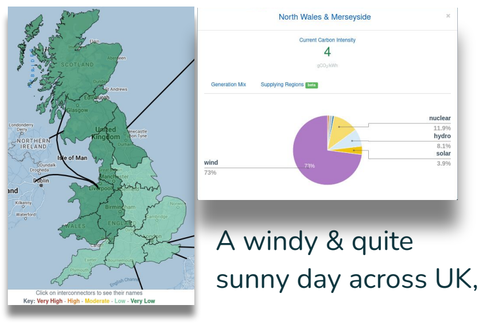
How does CATS work?

The **C**limate **A**ware **T**ask **S**cheduler calculates the optimal time to run a job to minimise its carbon intensity. It uses data from the National Energy System Operator (NESO) via it's carbonintensity.org.uk API. This provides 48 hour carbon intensity forecasts.

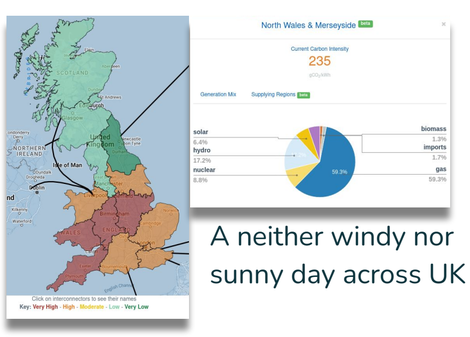
Plot from: *https://carbonintensity.org.uk/*



Grid carbon intensity varies considerably both geographically and in time. Typical values 0-400g CO_2e/kWh. Windy and/or sunny days have low carbon intensity (<50g) and windless cloudy days have a high carbon intensity (>200g).

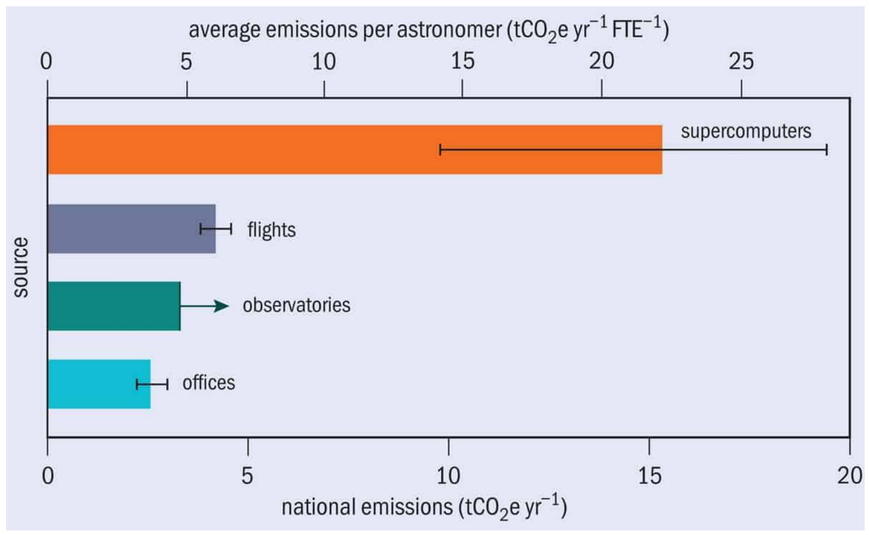


A windy & quite sunny day across UK,



A neither windy nor sunny day across UK

How much impact does (High Performance) Computing have?



Breakdown of the four sources of **Australian astronomers' emissions** considered in one study from 2019.

*Source: B. Li et al., 2023, Toward Sustainable HPC: Carbon Footprint Estimation and Environmental Implications of HPC Systems
† Source: A. Stevens et al, 2019, The imperative to reduce carbon emissions in astronomy

Basic CATS usage

```
cats -d <job duration in mins> --loc <postcode>
```

```
date
Mon  2 Dec 19:00:38 GMT 2024
cats -d 30 --loc RG1

WARNING:root:config file not found
WARNING:root:Unspecified carbon intensity forecast service, using carboninten
sity.org.uk

Best job start time                    = 2024-12-03 00:00:42.252413+00:00
Carbon intensity if job started now    = 270.48 gCO2eq/kWh
Carbon intensity at optimal time       = 192.91 gCO2eq/kWh
```

Download CATS

You can download CATS from:
https://github.com/GreenScheduler/cats

Or run: `pip install climate-aware-task-scheduler`

Directly schedule jobs using CATS

Use the `--scheduler` argument. We currently support the `at` and `sbatch`. To run a Python script `work.py` expected to take an hour on a computer in Warwick use:

```
cats -d 60 --loc CV4 --scheduler sbatch --command 'python work.py'
```

Limitations

CATS only works when a system isn't running at 100% load. It is best suited for smaller clusters that aren't always busy.

Currently CATS must be run by the user submitting the job(s). We are working on a SLURM plugin to create a "green" queue that prioritises carbon intensity.

Further Reading

CATS Documentation - https://cats.readthedocs.io/

JOSS Paper - https://joss.theoj.org/papers/10.21105/joss.08251

SOFTWARE SUSTAINABILITY INSTITUTE

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

UNIVERSITY OF OXFORD

National Oceanography Centre

Border image credits: 'Climate Stripes' infographic designed by Prof. Ed Hawkins (University of Reading), see showyourstripes.info