Environmental decision support: rare events, monitoring and inversion, and uncertainty quantification

Philip Jonathan

Lancaster University

Responsible research, Birmingham, 2026
(slides at *www.lancs.ac.uk/∼jonathan*)

Acknowledgement and overview

Acknowledgement

- Cambridge, Durham, Lancaster, Melbourne, Shell

Overview

- This is a talk about responsible research in environmental decision support
- My take on "responsible research"
- Remote sensing (airborne, line-of-sight, satellite)
- Extremes (non-stationary marginal, multivariate)

Responsible research

What is it?

- Research which genuinely tries to leave to world and everyone and everything it touches in a better state
- Is scientifically rigorous
- Avoids harm to people, animals and environment
- Serves the public good

Responsible research

In academia?

- Tries to manage “dual-use dilemmas”
  - e.g. Understanding of atomic physics from early 1900s led to Hiroshima and Chernobyl; but it also led to unprecedented widespread and lasting public good
  - Where could AI lead? ... and countless more mundate examples
- Provides “open-source” science
  - Transparent, accessible, and shared freely
- Addresses “real-world societal needs”
  - Not transient “flavour of the month” or “hobby-horse” topics
  - Not “chasing citations” or “top journals” or “oneupmanship”
  - But what about ”research for research’s sake” and unanticipated future impact?

Responsible research

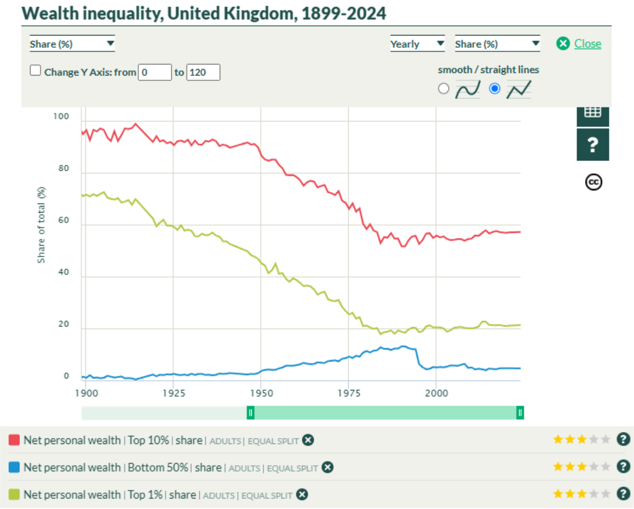
In heavy industry and fossil fuel sectors?

- Mitigates immediate harm
  - Methane detection and abatement
  - Safe decommissioning of large-scale facilities
- Mitigates medium-term harm: transforms industrial processes
  - Electricity in place of hydrocarbons as the energy source for heat-intensive reactions (e.g. needed to create plastics, chemicals, fertilizers, extract lithium from rock)
  - Depleted oil and gas reservoirs for carbon capture and storage
  - Existing pipeline infrastructure to carry green hydrogen in place of methane
  - Alternative sources of energy: renewables, small modular nuclear reactors?
- Transforms society
  - ”Addressing the economic reality of energy-dependent communities and wider society”
- Green-washing
  - Who’s the judge?

Responsible research: reality check

- Is responsible research compatible with a free-market economy?
    - Driven by “profit motive” and competitive “advantage”
        - At a corporate and individual level
    - Fundamentally misaligned with RR (driven by “public good”)
    - From a free-market perspective, RR is a cost, it reduces profit and competitive advantage $\Rightarrow$ it’s suboptimal (to say the least) to do it
- Public pressure is ultimately the handbrake
    - Public demand for regulation and other “guardrails”
    - Public demand (and resulting sources of investment) for ethical processes and products
    - Education, transparency, open science
- But does the handbrake really work?
    - Think about climate change
        - Societal inequality is increasing (e.g. in UK since late 1970s) `https://wir2026.wid.world/`
        - The wealthiest in society behave most irresponsibly regarding climate change `https://wid.world/www-site/uploads/2025/10/Climate_Inequality_Report_2025_Final.pdf`
    - Think about ill-informed or even malicious public pressure

Responsible research: UK wealth inequality



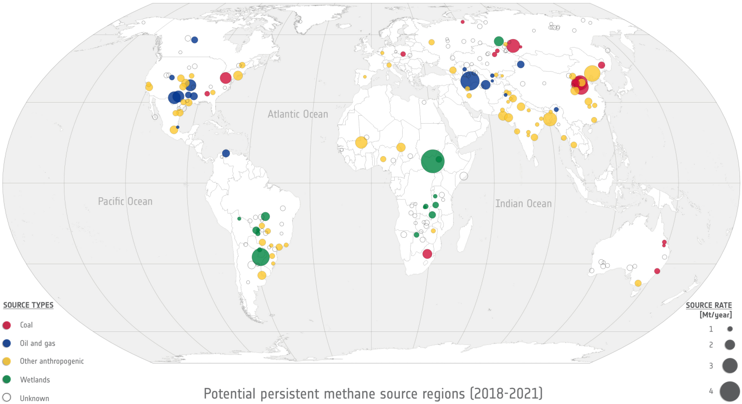
This is the context of current "UK responsible research" `https://wir2026.wid.world/`

Remote sensing of gaseous emissions

Remote sensing: why is this important?

- Greenhouse gases cause global warming with potentially catastrophic effects; CO_2 and CH_4 are the two most important greenhouse gases
- "Natural" CO_2 sources (aerobic respiration, organic decomposition, ocean-atmosphere exchange, volcanoes, wildfires), as well as "anthropogenic" sources (e.g. combustion, which we can try to reduce)
- "Natural" CH_4 sources (anaerobes in wetlands, ruminants, permafrost; ocean hydrates; wildfires), as well as "anthropogenic" sources
- CO_2 stays in the atmosphere for centuries; CH_4 for ≈ 12 years only
- CO_2 has a global warming potential of unity (by definition); GWP(CH_4 over 20 years) ≈ 80 (much worse!)
- CH_4 emissions responsible for half the extreme heat absorbed by the atmosphere since the industrial revolution
- If we can stop CH_4 getting into the atmosphere, we can provide "immediate mitigation"
- We can look for CH_4 sources locally (using point sensors, line-of-sight sensors or drones) or on a global scale using satellites

Remote sensing: super-emitters



`https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P/Top_10_persistent_methane_sources` using data from the Copernicus Sentinel-5P TROPOMI satellite

Remote sensing: fracking in the Permian basin (Texas)

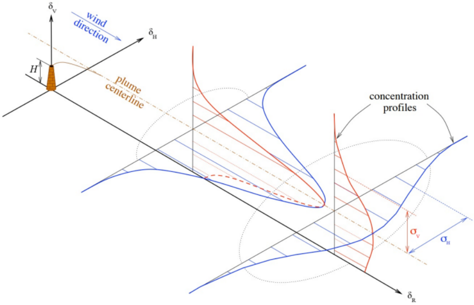
https://fossilfuel.com/the-west-texas-oil-boom-that-changed-the-world/

Remote sensing: super-emitter plumes (Lybia)

https://phys.org/news/2023-09-trio-sentinel-satellites-methane-super-emitters.html

Remote sensing: key ideas

- An unknown source of gas
- Gas transported on the wind
- We observed the gas remotely using a sensor
- Given the remote measurements and some knowledge of the wind field, can we characterise the source (size and location)?



Stockie [2011]

Remote sensing: formulation

- Statistically, we have a problem of probabilistic inversion

$$c_t = \sum_{j=1}^{n_S} a(\boldsymbol{y}_t, \boldsymbol{x}_j, \boldsymbol{w}_t, \boldsymbol{\sigma})\, s_j + b(\boldsymbol{y}_t, t), \quad t = 1, 2, 3, \ldots$$

- n_S sources of unknown strengths (s_j) at locations $(\boldsymbol{x}_j)$; n_S unknown
- Spatially-homogeneous wind field $\boldsymbol{w}_t$ at time t
- Measurement c_t at remote location $\boldsymbol{y}_t$ at time t
- Background $b(\boldsymbol{y}_t, t)$ at measurement location $\boldsymbol{y}_t$ at time t
- Coupling $a(\boldsymbol{y}_t, \boldsymbol{x}_j, \boldsymbol{w}_t, \boldsymbol{\sigma})$ (e.g. Gaussian plume) between source at $\boldsymbol{x}_j$ and measurement at $\boldsymbol{y}_t$ in wind field $\boldsymbol{w}_t$, with (potentially unknown) fixed coupling parameters $\boldsymbol{\sigma}$

- **Key insight**: sources are “spiky” whereas background is “smooth” allowing identification of both

Remote sensing: methodology and contributions

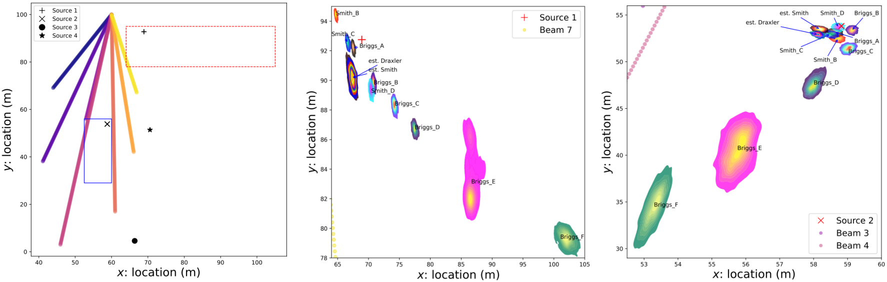
Methodology

- MCMC, full conditionals where possible
- MH with gradient-based (mMALA) proposals otherwise
- More recently, MLP

Contributions

- MCMC quantification of land-fill emissions: Hirst et al. [2013]
- Line-of-sight inversion: Hirst et al. [2017]
- Permian basin (satellite): Roberts et al. [2022]
- MCMC for estimation of coupling parameters: Newman et al. [2025a]
- Deep learning for inversion: Newman et al. [2025b]

Remote sensing: illustrative results



Newman et al. [2025a]. "est" are source estimates with opening angles estimated as part of the MCMC. Other estimates make explicit assumptions (often guessed) about opening angles.

Remote sensing: what next?

- Remote sensing (and associated inference) rapidly growing field
- SRON: `https://www.sron.nl/en/pillars/science/earth/methane/`
- GHGSat: `https://www.ghgsat.com/`

- Evolving regulatory frameworks
- Satellites, line of sight and point sensors
- Independent verification of industrial reporting of greenhouse gas emissions

- "Earth observation"
- `https://www.gov.uk/government/collections/earth-observation-eo`
- `https://www.earthdata.nasa.gov/learn/earth-observation-data-basics`
- `https://earth.esa.int/eogateway`

Characterising extreme ocean environments

Extreme environments: why is this important?

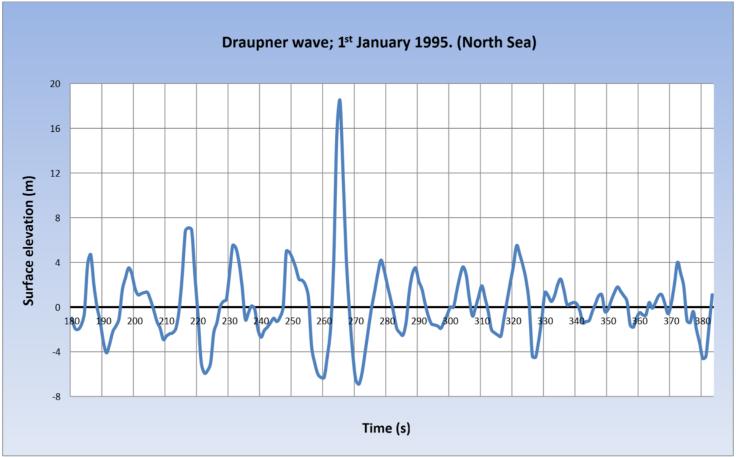
- Extreme environments are **dangerous to life** and human infrastructure
  - Extreme ocean storms, earthquakes, floods, wildfires, volcanos, ...
  - Want to build processes and structures that are "safe" given an inherently "risky" environment
  - e.g. "This structure must withstand the 1-in-1000-year event"
- Influence and improve **existing regulatory environment**
  - Historically, lack of data and methodology for good risk analysis, so existing industry standards use "sub-optimal" approaches
  - Modern methods are **more** complete and coherent (in terms of physical and statistical understanding)
  - But people must be **convinced** to switch to better approaches
- Combine best new ideas from academia with best practice from practical experience

Extreme waves: spectacular scale

Offshore Portugal, 24m wave height, November 2017 (The Guardian)

- Nazaré is a great source of huge coastal waves

Extreme waves: spectacular scale



Laser readings, 1 January 1995. Wave 25.6m, crest 18.5m (Statoil / Equinor)

- Maximum recorded wave height > 30m (multiple events, various sources)
- Maximum recorded significant wave height : 19.0m (buoy, North Atlantic, 4 Feb 2013, WMO)

Extreme waves: impact damage

Norwegian Dream, Atlantic, 2007
(gcaptain.com)

Ike, Gulf of Mexico, 2008
(Joe Richard)

Extreme environments: non-stationary marginal

- Environmental extremes of (univariate) Y vary continuously with multidimensional covariates Ω
- Asymptotic theory gives form of distribution of exceedances of high threshold ψ

$$Y|(\Omega, Y > \psi) \sim GP(\xi, \sigma, \psi), \text{ generalised Pareto with parameters } \xi, \sigma, \psi$$

- Inferences should reflect sources of uncertainty fairly
- Need statistical and computational efficiency

- **Predict extreme quantiles** of Y
- **Assess risk** (or expected loss $\mathbb{E}(L)$) for system $S = s_0$ due to Y and structural response R

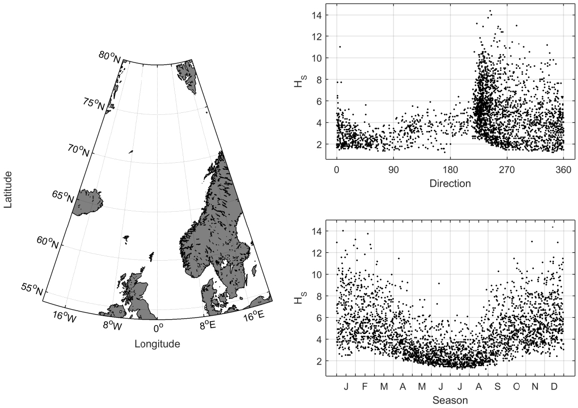
$$\mathbb{E}(L|S = s_0) = \int_r \int_y \int_\omega L(r|S = s_0) f_{R|Y}(r|y) f_{Y|\Omega}(y|\omega) f_\Omega(\omega) d\omega dy dr$$

- Use cases: Offshore and coastal design, weather windows and alerts

- Jones et al. [2018], Hansen et al. [2020] ,Towe et al. [2021]

Directional-seasonal: application

Storm peak significant wave height at northern North Sea location; clear directional and seasonal variability in storm severity; directional variability more dramatic at around 225°; seasonal variability more gradual.



Directional-seasonal: the model

Model for size of occurrences of large storms

- Hybrid density $f(y|\xi, \sigma, \alpha, \gamma, \psi, \tau) =$
$$\begin{cases} \tau \times f_{TW}(y|\alpha, \gamma) & \text{for } y \leq \psi, \text{ a truncated Weibull (or similar)} \\ (1-\tau) \times f_{GP}(y|\xi, \sigma) & \text{for } y > \psi \end{cases}$$
- Threshold non-exceedance probability τ to be inferred
- Physics suggests parameters $\alpha, \beta, \rho, \xi, \sigma, \psi$ and τ vary smoothly with covariates θ, ϕ
- Values of $\eta \in \{\alpha, \beta, \rho, \xi, \sigma, \psi, \tau\}$ on some index set of covariates take the form $\boldsymbol{\eta} = \boldsymbol{B}\boldsymbol{\beta}_\eta$ with basis $\boldsymbol{B}$ (for covariate domain) and basis coefficients $\boldsymbol{\beta}_\eta$
- Randell et al. [2016], Zanini et al. [2020]

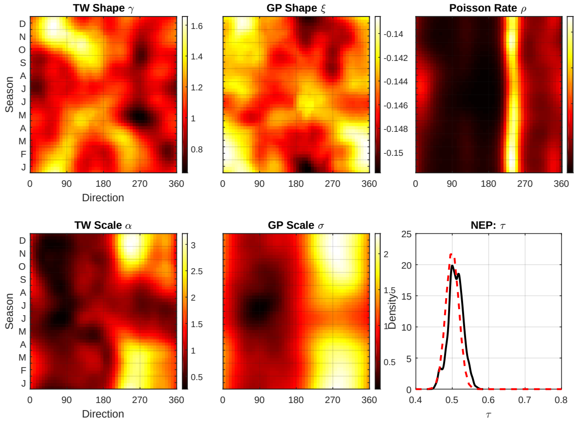
Directional-seasonal: inference

- Sampling from full conditionals
- Gibbs sampling when full conditionals available in closed form
- Metropolis-Hastings (MH) within Gibbs otherwise, using suitable proposal mechanisms (mMALA)

- Roberts and Stramer [2002], Girolami and Calderhead [2011], Xifara et al. [2014]

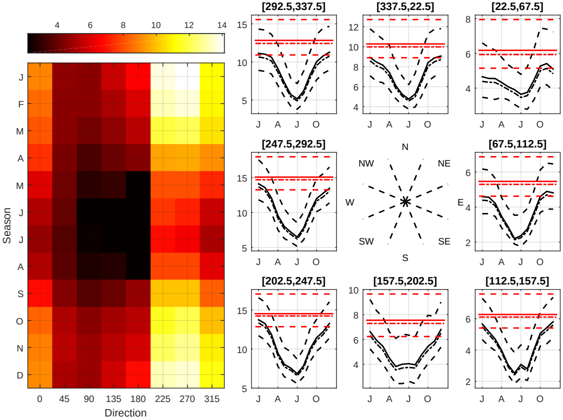
Directional-seasonal: parameter estimates

Prior for τ in red, posterior in black; all parameters except ξ and τ suggest strong directional variation; seasonal variation less pronounced but clear for α, σ (and ρ); ξ effectively constant; sample not informative about τ.



Directional-seasonal: return values

Predictive distribution of the 100-year maximum (in metres); directional and seasonal variability of the median estimate on lhs; seasonal variation of predictive distribution for directional octants (2.5%, 37%, median and 97.5% values) in black; corresponding omni-seasonal estimates in red; large difference between S and SW; smooth seasonal variation.



Extreme environments: the future

Use case

- Risk analysis for **multivariate spatio-temporal** processes

Models

- Max-stable processes, copulas, conditional extremes, SPAR, engineering models
  - Ross et al. [2017], Tendijck et al. [2019], Tendijck et al. [2023], Murphy-Barltrop et al. [2024], Sando et al. [2024]
- **Key insight**: dependence **in body** and dependence **in tail** are **different** e.g. $[X, Y] \sim N(0, [1\ \rho; \rho\ 1])$, $\rho < 1$, $\lim_{x \to \infty} \Pr(Y > x | X > x) = 0$

Risk assessment

- Expected loss $\mathbb{E}(L)$ for system $S = s_0$ due to **multivariate** Y and **multivariate** structural response R

$$\mathbb{E}(L|S = s_0) = \int_r \int_y \int_\omega L(r|S = s_0) f_{R|Y}(r|y) f_{Y|\Omega}(y|\omega) f_\Omega(\omega)\, d\omega\, dy\, dr$$

Summary

Responsible research in environmental decision support: summary

Scientific rigour

- In my experience of "conventional applied mathematical sciences", it's genuine unintended lack of insight that leads to lack of scientific rigour: "cock-up" rather than "conspiracy"
- But wilful ignorance is no defence (e.g. with respect to use of LLMs)

Avoid harm

- Usually, in "conventional applied science", it's genuine unintended lack of insight that leads to (potential) harm
- At the industrial "coal face", my impression is that the profit motive and competitive advantage can sometimes be overwhelming (or "blinding")

Public good

- Ultimately subjective
- The future is unknown
- But these are no reason to stop us **trying** to behave **rationally** and with a **clear conscience**

References

A. Brezger and S. Lang. Generalized structured additive regression based on Bayesian P-splines. *Comput. Statist. Data Anal.*, 50:967–991, 2006.

I. D. Currie, M. Durban, and P. H. C. Eilers. Generalized linear array models with applications to multidimensional smoothing. *J. Roy. Statist. Soc. B*, 68: 259–280, 2006.

P. H. C. Eilers and B. D. Marx. Splines, knots and penalties. *Wiley Interscience Reviews: Computational Statistics*, 2:637–653, 2010.

M. Girolami and B. Calderhead. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *J. Roy. Statist. Soc. B*, 73:123–214, 2011.

H. F. Hansen, D. Randell, A. R. Zeeberg, and P. Jonathan. Directional-seasonal extreme value analysis of North Sea storm conditions. *Ocean Eng.*, 195:106665, 2020.

B. Hirst, P. Jonathan, F. González del Cueto, D. Randell, and O. Kosut. Locating and quantifying gas emission sources using remotely obtained concentration data. *Atmospheric Environ.*, 74:141–158, 2013.

B. Hirst, D. Randell, M. Jones, P. Jonathan, B. King, and M. Dean. A new technique for monitoring the atmosphere above onshore carbon storage projects that can estimate the locations and mass emission rates of detected sources. *Energy Procedia*, 114:3716–3728, 2017.

M. J. Jones, H. F. Hansen, A. R. Zeeberg, D. Randell, and P. Jonathan. Uncertainty quantification in estimation of ocean environmental return values. *Coastal Eng.*, 141:36–51, 2018.

C. J. R. Murphy-Barltrop, E. Mackay, and P. Jonathan. Inference for bivariate extremes via a semi-parametric angular-radial model. *Extremes*, 2024. URL `https://doi.org/10.1007/s10687-024-00492-2`.

T Newman, C Nemeth, M. J. Jones, and P. Jonathan. Probabilistic inversion modeling of gas emissions: a gradient-based MCMC estimation of Gaussian plume parameters. *Ann. Appl. Stat.*, 19:2937–2956, 2025a.

T Newman, C Nemeth, M. J. Jones, and P. Jonathan. Deep learning surrogates for real-time gas emission inversion. *arXiv:2408.01298.*, 2025b.

D. Randell, K. Turnbull, K. Ewans, and P. Jonathan. Bayesian inference for non-stationary marginal extremes. *Environmetrics*, 27:439–450, 2016.

C. N. Roberts, O. Shorttle, K. Mandel, M. J. Jones, R. Ijzermans, B. Hirst, and P. Jonathan. Enhanced monitoring of atmospheric methane from space over the Permian basin with hierarchical Bayesian inference. *Accepted by Environ. Res. Lett., arXiv:2111.12486*, 2022.

G. O. Roberts and O. Stramer. Langevin diffusions and Metropolis-Hastings algorithms. *Methodol. Comput. Appl. Probab.*, 4:337–358, 2002.

E Ross, M Kereszturi, M van Nee, D Randell, and P Jonathan. On the spatial dependence of extreme ocean storm seas. *Ocean Eng.*, 145:359–372, 2017.

K. Sando, R. Wada, J. Rohmer, and P. Jonathan. Multivariate spatial and spatio-temporal models for extreme tropical cyclone seas. *Ocean Eng.*, 309, 2024.

J. M. Stockie. The mathematics of atmospheric dispersion modeling. *SIAM Rev.*, 53:349–372, 2011.

S. Tendijck, E. Ross, D. Randell, and P. Jonathan. A non-stationary statistical model for the evolution of extreme storm events. *Environmetrics*, 30:e2541, 2019.

S. Tendijck, E. Eastoe, J. Tawn, D. Randell, and P. Jonathan. Modeling the extremes of bivariate mixture distributions with application to oceanographic data. *J. Am. Statist. Soc.*, 118:1373–1384, 2023.

R. Towe, E. Zanini, D. Randell, G. Feld, and P. Jonathan. Efficient estimation of distributional properties of extreme seas from a hierarchical description applied to calculation of un-manning and other weather-related operational windows. *Ocean Eng.*, 238:109642, 2021.

T. Xifara, C. Sherlock, S. Livingstone, S. Byrne, and M Girolami. Langevin diffusions and the Metropolis-adjusted Langevin algorithm. *Stat. Probabil. Lett.*, 91 (2002):14–19, 2014.

E. Zanini, E. Eastoe, M. J. Jones, D. Randell, and P. Jonathan. Flexible covariate representations for extremes. *Environmetrics*, 31:e2624, 2020.