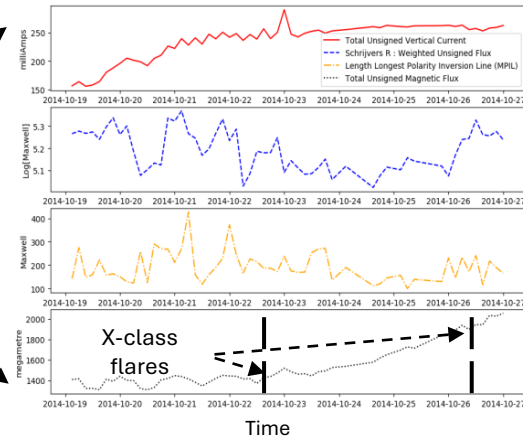
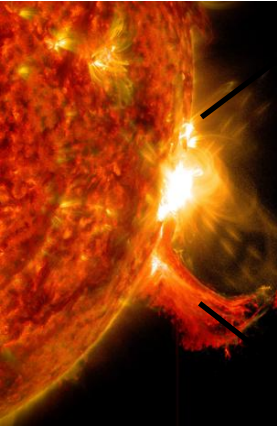


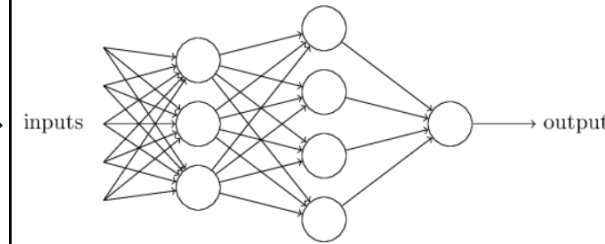
# Cause-mic Universe : Causal Approaches Probing Solar and Astrophysical Variability

## Solar Outbursts



•  
•  
•  
•  
 $y^{n:0}_{i-1}$   
 $y^{n:0}_i$   
 $y^{n:0}_{i+1}$   
•  
•  
•

## Statistical Learning or Inference



Machine (Deep) Learning  
Bayesian Networks

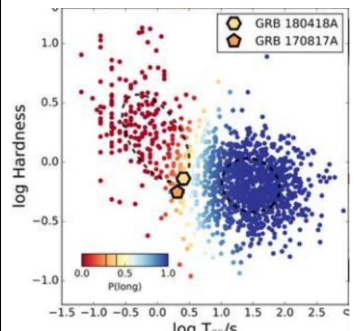
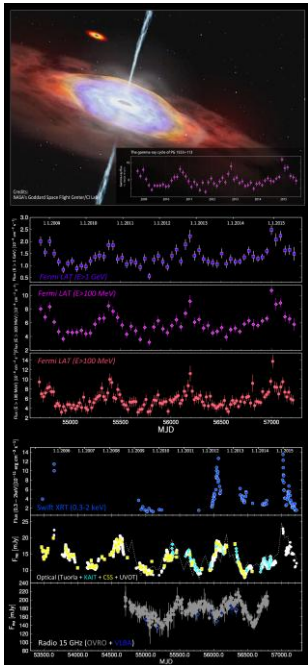
## Physics Drivers

- Competing mechanisms (variables) driving variability
- Their sequence via lags
- Transient classes (progenitors)

## Forecasts

- Solar outbursts impact space weather
- Feature-set for timely and precise ML-based forecasts (Met-Office)

## Extragalactic Sources (transients)



GRBs : Escorial et al., 2021

Nachiketa Chakraborty (University of Reading)  
[EuCAIF, 2025] [n.chakraborty@reading.ac.uk](mailto:n.chakraborty@reading.ac.uk)

with

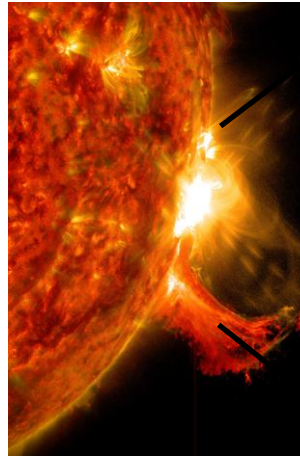
Matthew Owens, Harriet Turner, Matthew Lang, Peter Jan van Leeuwen, Amy Lien



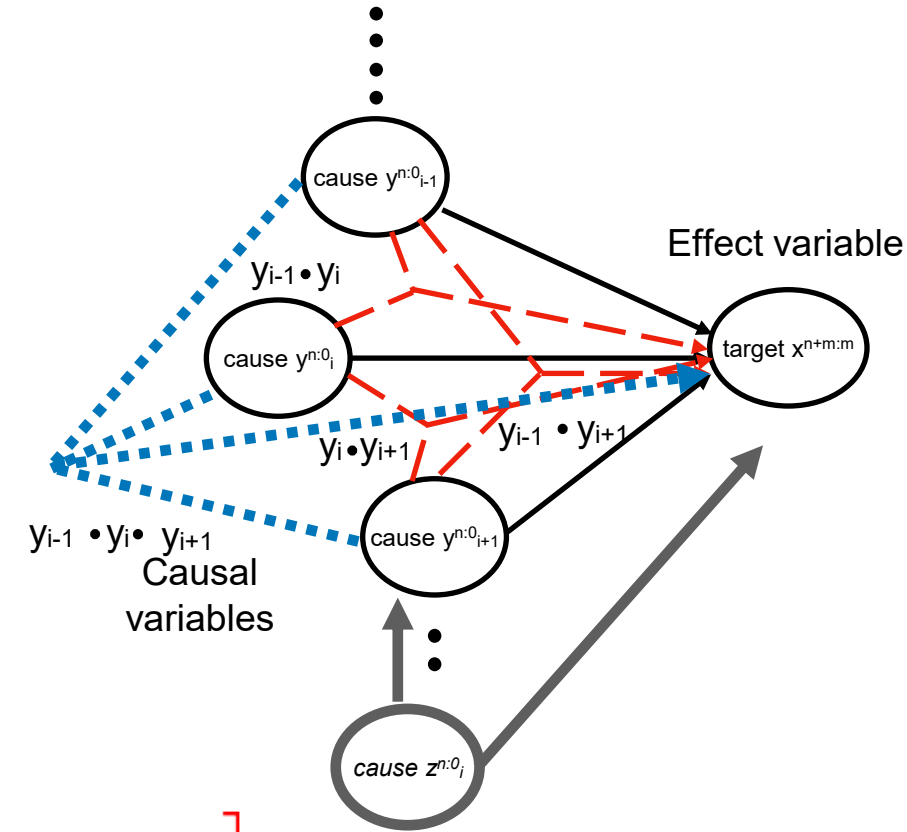
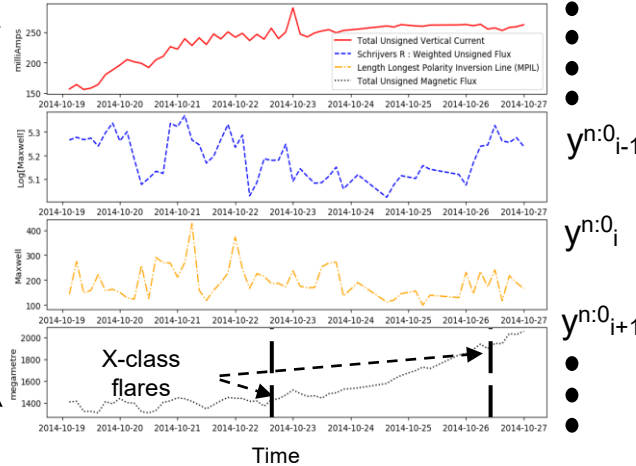
# Generalised, Hierarchical Information Graphs

## Limitations

- Non-linearity, non-Gaussianity
- No consensus on classes
- Black box ML/DL
- Ceiling in forecast performance



## Solar Outbursts



Information Theoretic  
Entropic Measures

$$(y \rightarrow x)_{\text{total}} = I(x; y|z) + \frac{1}{2} \left[ I(x; y) - I(x; y|z) \right]$$

$$(y_i \rightarrow x)_{\text{total}} = (y_i \rightarrow x)_{1\text{link}} + \frac{1}{2}(y_i \rightarrow x)_{2\text{links}} + \frac{1}{3}(y_i \rightarrow x)_{3\text{links}} + \dots + \frac{1}{N}(y_i \rightarrow x)_{N\text{links}}$$

# Cause-mic Universe : Causal Approaches probing Solar and Astrophysical Variability

Nachiketa Chakraborty<sup>1,2</sup> | Harriet Turner<sup>2</sup> | Mathew Owens<sup>2</sup> | Matthew Lang<sup>2,3</sup>

## Introduction

Astrophysical sources exhibit variability across diverse timescales, reflecting complexity and nonlinearity. Disentangling the underlying causal processes driving variability, often in-tandem is challenging. We illustrate how graphical causal models, based on information theory, can clarify these dynamics and provide new feature-sets for modelling in both solar and extragalactic contexts.

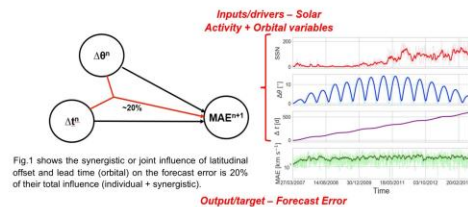
## Drivers of Solar Wind and Flare Forecast

- Solar wind and flares are complex, non-linear dynamical systems
- Together they impact space weather, necessitating timely forecasts
- Information theory provides non-linear measures of association : Shannon entropy (H), (Conditional) Mutual & Interaction Information (I)

$$I(x; y) = H(x) - H(x|y) (= 0 \text{ independence}) \quad I(x; y; z) = I(x; y) - I(x; y|z) \quad (1)$$

- Hierarchical information decomposition captures causal relations amongst variables in non-linear, dynamical systems – this includes **joint/synergistic influence** of multiple inputs on output – attribution to synergistic terms missed otherwise
- Solar Wind:** Satellite Observatories (STEREO+OMNI) measures wind speeds – forecast driven by solar activity + orbital variables namely latitudinal offset, lead time.
- Significant mixing of influence (**red term**)

$$(y \rightarrow x)_{\text{total}} = I(x; y|z) + \frac{1}{2} [I(x; y) - I(x; y|z)] \quad (2)$$



- Solar Flares:** Solar-flare index (FI) - total energy emitted in a given time (Kleczek, 1952). Indices Kp, Ap, Dst (Disturbance Storm Time) capture geomagnetic activity along with scalar magnetic field, Bscal.
- Applying Recurrent Neural Network (RNN) and Causal Information Graphs, we get features importance for forecast model with Kernel Shapley values and total information (direct+synergistic eqn.(2))
- Causal graphs (fig 2) reveal **significant synergy** between Ap/Dst and Bscal. Large part of influence of Ap/Dst is shared/mediated by Bscal.
- “Feature-set Importance”** (synergistic combination) missed by traditional networks and metrics

### Contact information

Department of Computer Science, University of Reading, Poly Vacher Building, RG6 6DH  
Email: [n.chakraborty@reading.ac.uk](mailto:n.chakraborty@reading.ac.uk) | [www.reading.ac.uk/staff/dr-nachiketa-chakraborty](http://www.reading.ac.uk/staff/dr-nachiketa-chakraborty)

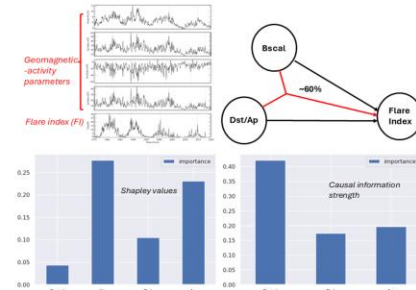


Fig.2 shows preliminary results for feature importance estimated using Shapley values for RNN model vs causal information strength using graphical causal models. The latter captures synergistic effects (60%) promoting scalar B to a higher rank. We can use this to design a feature-set including those capturing synergy between observables.

- Astrophysical Variability:** Sources like Active Galactic Nuclei (AGN), show variability across time-scales (minutes to years).
- Time-lags from (a) reprocessing from the torus and (b) inward propagating accretion disk fluctuations (scenario sketch in fig.3)
- UV and X-ray:** We investigate timescales of 0.1 to few days with multiwavelength light curves of AGN NGC 4593
- Mutual Information estimates detect **UV leading X-rays by ~1 day**—consistent with (b) and missed by traditional methods.

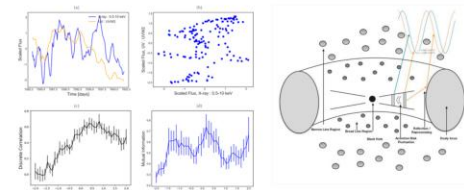


Fig.3 shows a simplified AGN schematic showing time lags between emissions from the black hole, accretion disk, dusty torus, and broad-line region (jet omitted for clarity). Torus reprocessing and inward-propagating fluctuations produce opposite lag directions. Both lag directions are detected accurately by mutual information.

## Conclusions: Hierarchical information graphs provide

- Causal information including **synergistic** component of influence due to multiple variables
- New comprehensive, **explainable “feature-set importance”** vital for both physics and space weather forecasting with ML
- Accurate time-lags** for non-linear relationships between observables and underlying non-linear physical mechanisms

### References

- N. Chakraborty, H. Turner, et al., *Solar Physics*, Vol. 298, article number 142, (2023) DOI: 10.1007/s11207-023-0232-4
- A. Ozguc, et al., *Solar Physics*, 297, article id 112, (2022), DOI: 10.1007/s11207-022-02049-7
- N. Chakraborty, P.J. van Leeuwen, *Phys. Rev. R.*, 4, 013036, (2022) DOI:10.1103

### Acknowledgements

- We acknowledge support from the Dept. of CS, Reading and past funding from the Horizon 2020 program, via CUNDA Project No. 694509, and STFC grant number ST/V000497/1..

# Conclusions:

Hierarchical information graphs provide

- Standalone inference or **diagnostics** for ML/DL
- Causal information including **synergistic influence** due to multiple variables
- New comprehensive, **explainable “feature-set importance”** vital for both physics and space weather forecasting with ML
- Accurate time-lags** for non-linearity

Poster Id : 63 (T1c)

Session A (Wednesday 12:00 - 15:00)

[n.chakraborty@reading.ac.uk](mailto:n.chakraborty@reading.ac.uk)