Causal Dynamic Time Lag (CDT)

Applications to Space Weather

Mandar Chandorkar

[1]Multiscale Dynamics
CWI, Amsterdam

[2]TAU Research Unit
INRIA, Paris-Saclay

25 September 2018

INVENTEURS DU MONDE NUMÉRIQUE

Centrum Wiskunde & Informatica

Outline

Outline

1. Motivation
   - Space Weather
   - Financial Markets
2. Causality in Time Series
   - Concepts
   - Existing Research
3. Problem & Model
   - Description
   - Proposed Solution
4. Applications
   - Benchmarks
   - Problem I
   - Problem II
   - Problem III

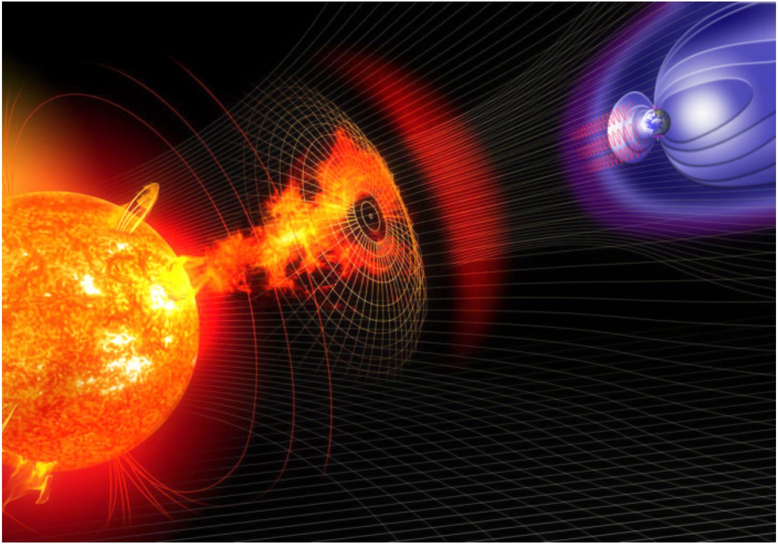

Figure: The Sun-Earth system

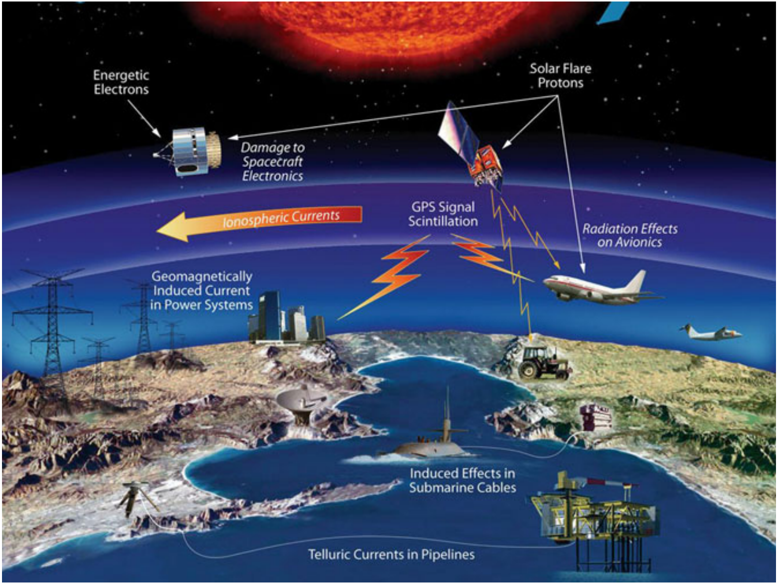

Figure: Effects of Solar Disturbances

Outline

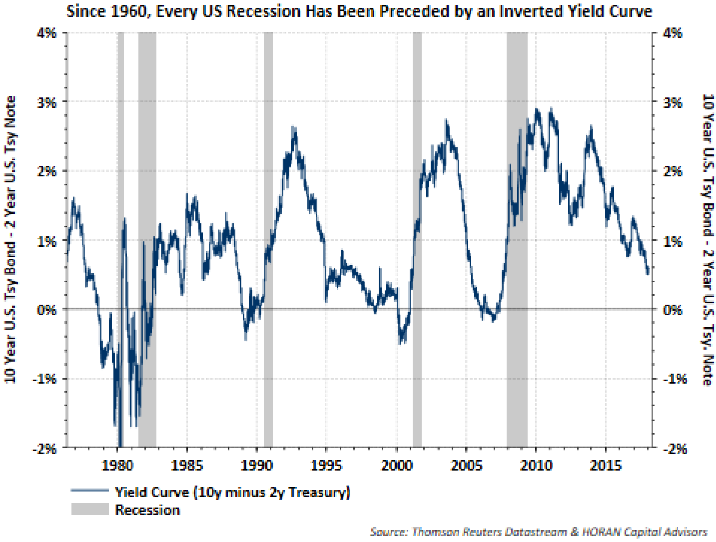

Figure: Yield curves and Recessions in the U.S. Economy

Outline

Granger Causality

1. The cause happens prior to its effect.
2. The cause has unique information about the future values of its effect.

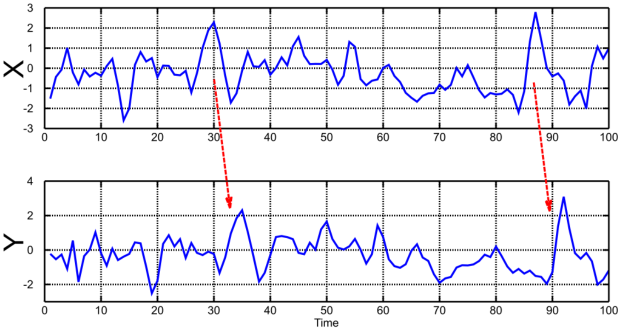

Figure: By BiObserver - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=33470670

Outline

Nonparametric learning of time-lag

[Zhou and Sornette, 2006] formulated the problem as minimisation of the *mismatch* between two time series.

Inputs: $X(t)$, $Y(t)$, two time series.

Learn: A mapping $\phi(t_1) = t_2$ which minimises

$$\epsilon(t_t, t_2) = |X(t_1) - Y(t_2)| \tag{1}$$

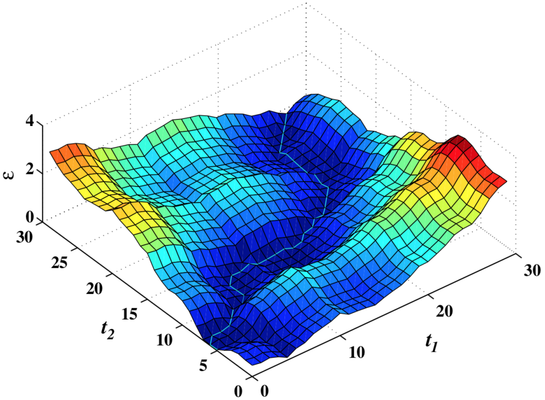

Figure: An example of energy landscape $E_{X,Y}$ given by (1) for two noisy time series and the corresponding optimal path wandering at the bottom of the valley similarly to a river. This optimal path defines the mapping $t1 \rightarrow t2 = \phi(t1)$.

Outline

Causal Dynamic Time-lag Inference (CDT)

Given an input (cause) and output (effect) time series, predict

- Magnitude of output signal. (What/How much)
- When the effect will be observed in the output signal (When)

CDT: Formal Definition

Input Signal (Cause)

$$t \in \mathbb{R}^+$$
$$x(t) \in \mathcal{X}$$

Output Signal (Effect)

$$f : \mathcal{X} \rightarrow \mathbb{R}$$
$$g : \mathcal{X} \rightarrow \mathbb{R}^+$$
$$\Delta(t) = g[x(t)]$$
$$y(t + \Delta(t)) = f[x(t)]$$

Outline

Model Specification

Input Patterns $x(t)$

Causal Time Window
Lower Limit: $\ell \in \mathbb{N} \cup 0$
Upper Limit: $\ell + h : h \in \mathbb{N}$

Targets $y(t+\ell), \cdots, y(t+\ell+h-1)$

Model Outputs
Predictions $\hat{y}(t+\ell), \cdots, \hat{y}(t+\ell+h-1)$
Time Lag Probabilities $\hat{p}(t+\ell), \cdots, \hat{p}(t+\ell+h-1)$

Training: Motivation

Balance two incentives

1. Generate accurate predictions for time window $y(t+\ell), \cdots, y(t+\ell+h-1)$
2. Learn time lag structure according to some intuition.

Training Loss

$$\mathcal{L}(y^{(1:M)}, \hat{y}^{(1:M)}, \hat{p}^{(1:M)}) = \lambda_1 \sum_{i,m} \frac{1}{2M}(y_i^{(m)} - \hat{y}_i^{(m)})^2(1 + \hat{p}_i^{(m)})$$

$$+$$

$$\lambda_2 \mathcal{J}(y^{(1:M)}, \hat{y}^{(1:M)}, \hat{p}^{(1:M)})$$

Training Loss Contd.

The term $\mathcal{J}(y^{(1:M)}, \hat{y}^{(1:M)}, \hat{p}^{(1:M)})$ penalizes the predicted probabilities $\hat{p}^{(1:M)}$, for deviation from some chosen *target probability*.

Target Probability

The *target probability* $\widetilde{p}$ for a time window $[t + \ell, t + \ell + h - 1]$ can be characterized by:

Conjecture: Causal Time Lag

The lagged output $y(t + i)$ which has greater predictability given $x(t)$, is a more likely causal link.

Target Probability Contd.

1. The target probability distribution for the time lag is,
$\widetilde{p}^{(m)} = softmax((y^{(m)} - \hat{y}^{(m)})^2 / T)$
2. The term $\mathcal{J}(y^{(1:M)}, \hat{y}^{(1:M)}, \hat{p}^{(1:M)})$ can be computed as the *Hellinger distance* between $\hat{p}$ and $\widetilde{p}$.

Outline

Need for Benchmarks

1. Labelled data sets are small in size.
2. Most real world data sets dont have explicit labels for causal time lag.

Benchmark Problems

$$x(t+1) = (1-\beta)x(t) + \mathcal{N}(0, \sigma^2)$$
$$y(t+\Delta(t)) = \alpha||x(t)||^2$$

Problem I: Constant Lag

$\Delta(t) = k$

Problem II: Constant Velocity $||x(t)||^2$; Fixed Distance d

$\Delta(t) = d/(\alpha||x(t)||^2)$

Problem III: Constant Acceleration a; Fixed Distance d

$\Delta(t) = (\sqrt{\alpha^2||x(t)||^4 - 2ad} - \alpha||x(t)||^2)/a$

Outline

Data

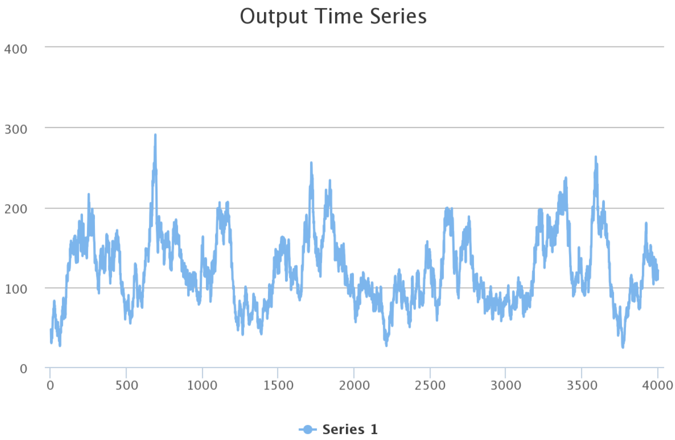

Figure: Generated Data

Errors

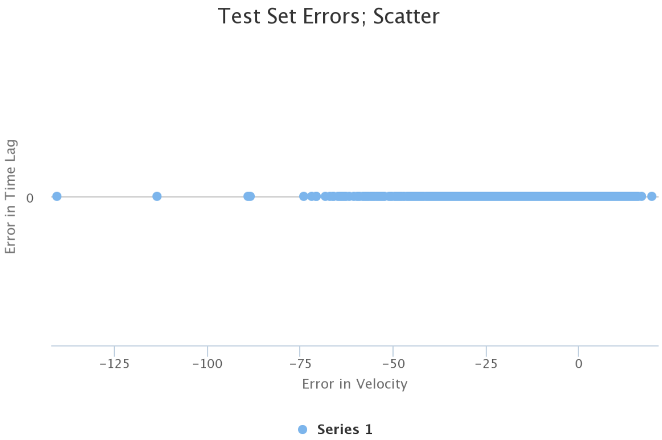

Figure: Error in Output vs Error in Time Lag

Predictions

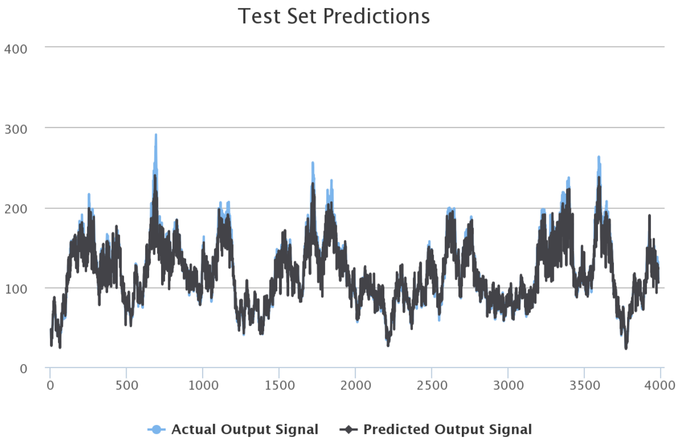

Figure: Test Set Time Series vs Predictions

Performance

Output

MAE: 8.602

Pearson Corr: 0.964

Spearman Corr: 0.999

Time Lag

MAE: 0

Pearson Corr: N.A

Spearman Corr: 1.0

Outline

Data

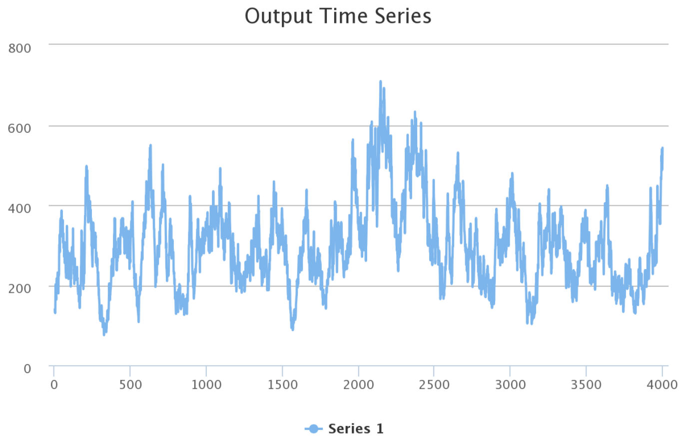

Figure: Generated Data

Time Lags

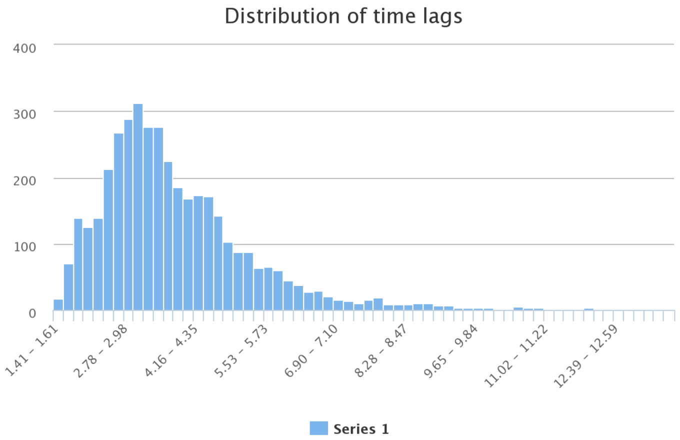

Figure: Time Lags

Test Data Distribution

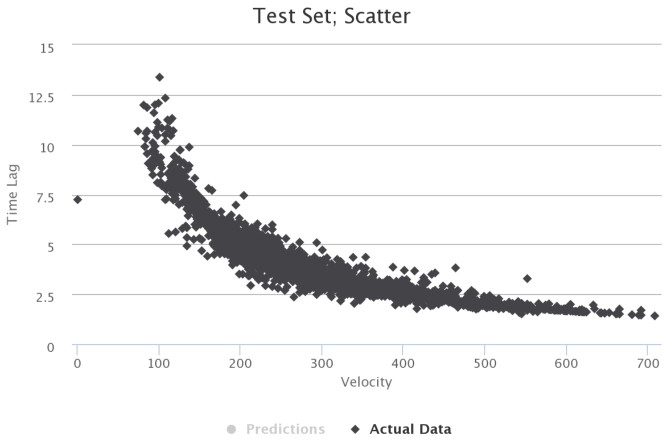

Figure: Output vs Time Lags

Predictions

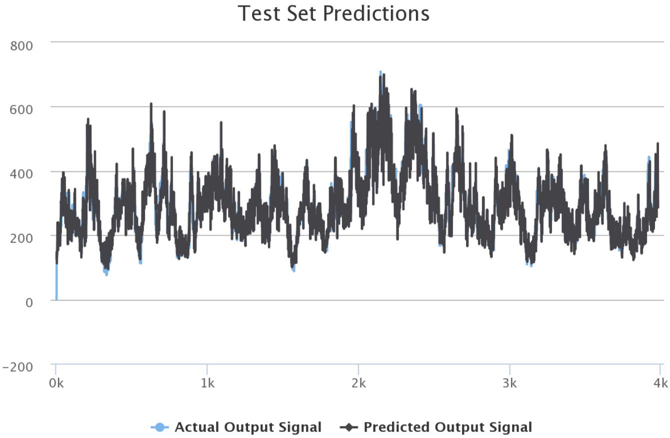

Figure: Test Set Time Series vs Predictions

Errors

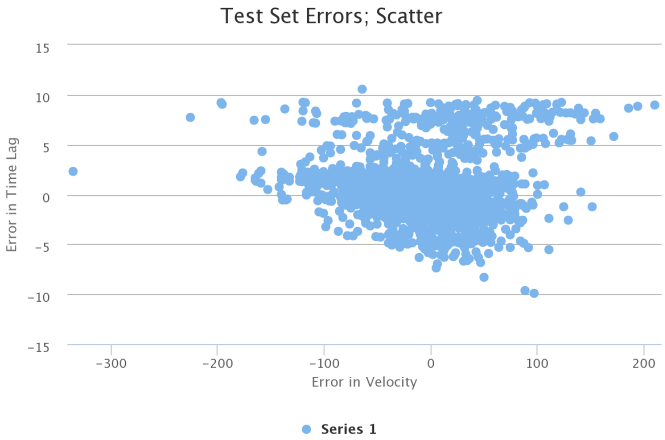

Figure: Error in Output vs Error in Time Lag

Performance

Output

MAE: 30.902
Pearson Corr: 0.918
Spearman Corr: 0.999

Time Lag

MAE: 1.593
Pearson Corr: 0.337
Spearman Corr: 0.999

Outline

Data

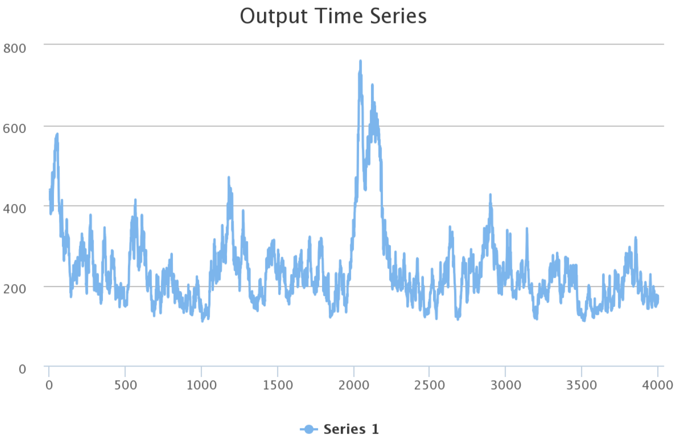

Figure: Generated Data

Time Lags

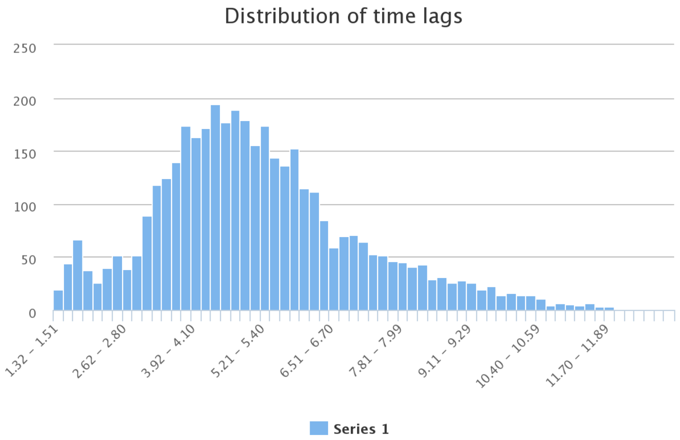

Figure: Time Lags

Test Data Distribution

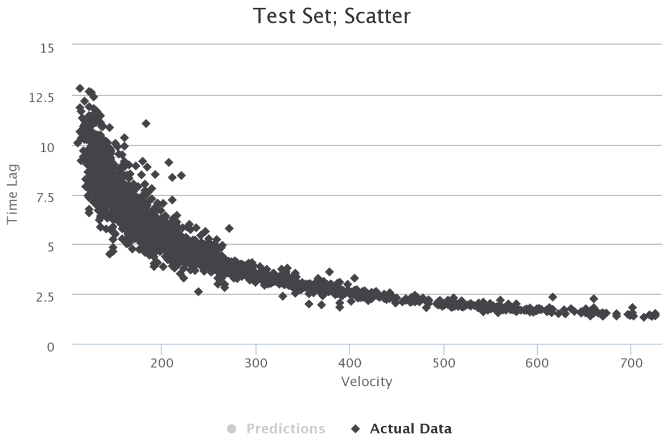

Figure: Output vs Time Lags

Predictions

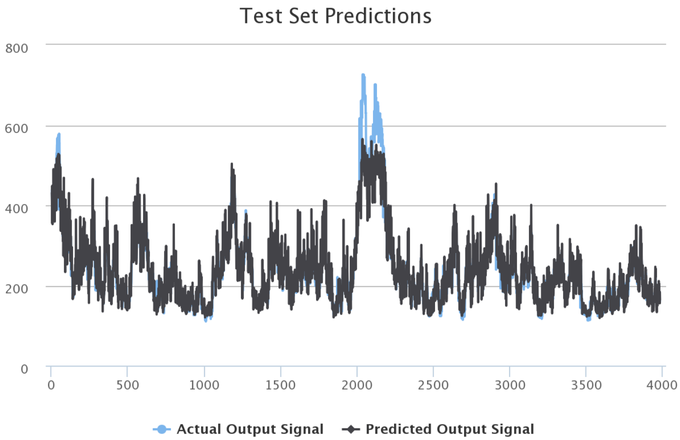

Figure: Test Set Time Series vs Predictions

Errors

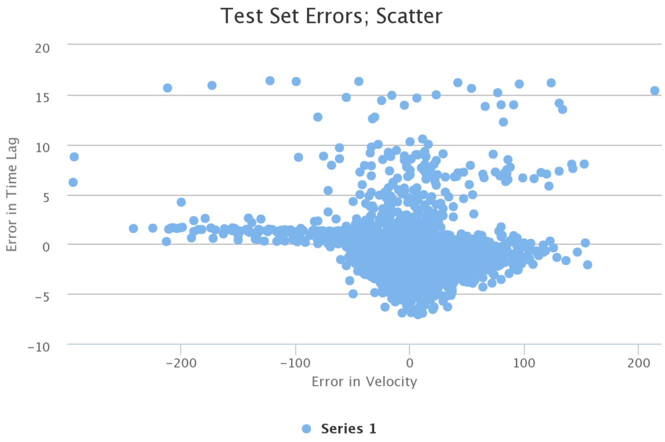

Figure: Error in Output vs Error in Time Lag

Performance

Output

MAE: 24.385
Pearson Corr: 0.928
Spearman Corr: 0.999

Time Lag

MAE: 1.758
Pearson Corr: 0.415
Spearman Corr: 0.999

MDG-TAU Team

Website `http://mlspaceweather.org`

References I

Zhou, W.-X. and Sornette, D. (2006).
Non-parametric determination of real-time lag structure between two time series: The optimal thermal causal path method with applications to economic data.
Journal of Macroeconomics, 28(1):195 – 224.
Nonlinear Macroeconomic Dynamics.