Attribution of individual events based on data assimilation

Alexis Hannart
CNRS

1 February 2016
Colleagues

- **Argentina**
  - Sammy Metref – IFAECI, CNRS, Argentina
  - Manuel Pulido – Department of Physics, UNN, Argentina
  - Juan Ruiz – IFAECI, CONICET, Argentina

- **Europe**
  - Marc Bocquet – Ecole des Ponts, Paris Tech, France
  - Alberto Carrassi – Mohn-Sverdrup Center, NERSC, Norway
  - Philippe Naveau – LSCE, CNRS, France

- **US**
  - Judea Pearl – UCLA, USA
  - Michael Ghil – UCLA, USA and ENS, France
Funding

- **Project DADA**
- PI: Alexis Hannart
- Grant ANR-13-JS06-0007-01
- ANR (Agence Nationale de la Recherche, France)
- March 2014 - Sept. 2017

**Big Data Paris**
Award Winner 2015


Starting point

Is this individual event attributable to human influence?

Need for causal answers from scientists.
Motivations

- Conventional approach:
  - Define the event
  - Derive the probability of the event in two model ensembles:
    - factual probability: $p_1$
    - counterfactual probability: $p_0$
  - Derive the fraction of attributable risk:
    \[
    \text{FAR} = 1 - \frac{p_0}{p_1}
    \]

- Research questions:
  - View that “We can’t say anything for individual event, we can only say something about classes of event”
  - What is a cause? What is an event? What is the FAR? Are we doing it right?
  - Operationalization: real-time
Motivations

- Conventional approach:
  - Define the event
  - Derive the probability of the event in two model ensembles:
    - factual probability: $p_1$
    - counterfactual probability: $p_0$
  - Derive the fraction of attributable risk:
    \[ \text{FAR} = 1 - \frac{p_0}{p_1} \]

- Research questions:
  - View that “We can’t say anything for individual event, we can only say something about classes of event”
  - What is a cause? What is an event? What is the FAR? Are we doing it right?
  - Operationalization: real-time

Introduce definitions: causality + events
Develop a real-time methodology
Outline

- Causal theory

- Event definition

- Attribution using data assimilation
Causal theory (Pearl): causality has two facets

- Necessary Causation
- Sufficient Causation
- Nec. & Suf. Causation
Causal theory (Pearl): Probabilities of causation

- **Probability of necessary causation** = probability that the effect is removed when the cause is turned off, conditional on the fact that the effect and the cause were initially present.
Causal theory (Pearl): Probabilities of causation

- **Probability of sufficient causation** = probability that the effect appears when the cause is turned on, conditional on the fact that the effect and the cause were initially absent.
Causal theory (Pearl): Probabilities of causation

\[ PN = \max \{1 - \frac{p_0}{p_1}, 0\} \]

\[ PNS = \max \{p_1 - p_0, 0\} \]

\[ PS = \max \{1 - \frac{1 - p_1}{1 - p_0}, 0\} \]
Causal theory (Pearl): Probabilities of causation

\[ PN = \max\left\{ 1 - \frac{p_0}{p_1}, 0 \right\} \]

\[ PNS = \max\{ p_1 - p_0, 0 \} \]

\[ PS = \max\left\{ 1 - \frac{1 - p_1}{1 - p_0}, 0 \right\} \]
Causal theory (Pearl): Probabilities of causation

\[ PN = \max\{1 - \frac{p_0}{p_1}, 0\} \]

\[ PNS = \max\{p_1 - p_0, 0\} \]

\[ PS = \max\{1 - \frac{1 - p_1}{1 - p_0}, 0\} \]

For rare events, both \( p_1 \) and \( p_0 \) are small.

PS and PNS are always low, only PN may be high.

Specific events are attributable in a PN sense.
Specific events are attributable, in a necessary causation sense.

\[ PN = \max\{1 - \frac{p_0}{p_1}, 0\} \]

is the right quantity to look at.
Does FAR represent a fraction of attributable risk?

FAR = 0.8
Does FAR represent a fraction of attributable risk?

FAR = 0.8

FAR2 = 0.2?
Does FAR represent a fraction of attributable risk?

- Anthropogenic emissions
- Intense SACZ
- Anticyclonic conditions

Heatwave

- PN1 = FAR1 = 0.8
- PN2 = FAR2 = 0.9*
- PN3 = FAR3 = 0.9*

FAR does not represent a share of causality. FARs do not sum up to one. PN could be used instead.

*stylized
Outline

- Causal theory
- Event definition
- Attribution using data assimilation
Event definition

- No matter which name is used, $1 - \frac{p0}{p1}$ is an appropriate metric: the main goal remains to derive $p0$ and $p1$.

- These probabilities are affected by the chosen event definition.
  - How general / specific?
  - Which variables?

**General definition**

Space time-average of a variable exceeds a threshold

**Specific definition**

Analogue sequence of several variables of interest over a space-time window
Event definition: formalization

- Denote $\mathbf{Y} \in \mathbb{R}^n$ a high-dimensional space-time random vector concatenating all observable variables

- $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ a scalar index and $E$ the event:

$$E = \{ \varphi(\mathbf{Y}) \geq u \}$$
Event definition: formalization

- Denote \( \mathbf{Y} \in \mathbb{R}^n \) a high-dimensional space-time random vector concatenating all observable variables.
- \( \varphi : \mathbb{R}^n \rightarrow \mathbb{R} \) a scalar index and \( E \) the event: \( E = \{ \varphi(\mathbf{Y}) \geq u \} \)
Event definition: formalization

- Denote $Y \in \mathbb{R}^n$ a high-dimensional space-time random vector concatenating all observable variables.
- $\varphi : \mathbb{R}^n \to \mathbb{R}$ a scalar index and $E$ the event: $E = \{\varphi(Y) \geq u\}$

![Diagram showing possible realizations of $Y$, observed realization $y$, and event $E$.](image)
Event definition: limitations

\[ E = \{ \varphi(Y) \geq u \} \]

- Very sensitive to the choice of \( \varphi \) and \( u \).
- Does not depend on the observation \( y \)
  - does not say much of the dynamic causal chain associated to the singular event \( y \).
  - table look-up approach: attribution performed beforehand for a list of pre-defined classes of events.
An illustration: the forced Lorenz model

dynamic equation

\[
\frac{dx}{dt} = \sigma(y - x) + \lambda_i \cos \theta_i, \quad \frac{dy}{dt} = \rho x - y - xz + \lambda_i \sin \theta_i, \quad \frac{dz}{dt} = xy - \beta z.
\]

factual: \( \lambda_1 = 30 \)

counterfactual: \( \lambda_0 = 0 \)
Difference in attractors in the forced Lorenz model
Event definition: limitations

\[ p_1 > p_0 \implies \text{the attribution conclusion for the chosen event definition corresponds to the position of the singular event.} \]
Event definition: limitations

\[ p_1 < p_0 \Rightarrow \text{the attribution conclusion for the chosen event definition does not correspond to the position of the singular event.} \]
Event definition: limitations

<table>
<thead>
<tr>
<th>$\phi_1$</th>
<th>$\phi_2$</th>
<th>$\phi_3$</th>
<th>$p_1$</th>
<th>$p_0$</th>
<th>$p_1/p_0$</th>
<th>$PN = \max{FAR, 0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.8</td>
<td>0.3</td>
<td>0.5</td>
<td>7E-05</td>
<td>1E-08</td>
<td>5E+03</td>
<td>0.9998</td>
</tr>
<tr>
<td>0.7</td>
<td>-0.6</td>
<td>0.4</td>
<td>6E-02</td>
<td>2E-01</td>
<td>4E-01</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.1</td>
<td>0.8</td>
<td>0.5</td>
<td>7E-04</td>
<td>1E-02</td>
<td>5E-02</td>
<td>0.0000</td>
</tr>
<tr>
<td>1.0</td>
<td>-0.1</td>
<td>-0.2</td>
<td>5E-03</td>
<td>6E-03</td>
<td>8E-01</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.4</td>
<td>-0.9</td>
<td>-0.2</td>
<td>7E-05</td>
<td>9E-06</td>
<td>8E+00</td>
<td>0.8755</td>
</tr>
<tr>
<td>-0.9</td>
<td>-0.3</td>
<td>0.4</td>
<td>9E-05</td>
<td>7E-08</td>
<td>1E+03</td>
<td>0.9992</td>
</tr>
<tr>
<td>0.6</td>
<td>-0.7</td>
<td>-0.3</td>
<td>9E-05</td>
<td>5E-05</td>
<td>2E+00</td>
<td>0.3768</td>
</tr>
<tr>
<td>0.3</td>
<td>-0.7</td>
<td>0.7</td>
<td>7E-05</td>
<td>2E-04</td>
<td>3E-01</td>
<td>0.0000</td>
</tr>
<tr>
<td>0.4</td>
<td>-0.9</td>
<td>-0.2</td>
<td>7E-05</td>
<td>7E-06</td>
<td>1E+01</td>
<td>0.9044</td>
</tr>
<tr>
<td>-0.9</td>
<td>0.5</td>
<td>-0.3</td>
<td>5E-05</td>
<td>3E-14</td>
<td>2E+09</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
Event definition: limitations - illustration

- Mr. A is affected by obesity and is a Coca-Cola drinker. Is Mr. A’s obesity caused by drinking Coca-Cola?

- “Class of events” approach: 
  - “obesity” = individuals with Body Mass Index > 35
  - $p_1 = \text{proba of being obese when drinking Coca-Cola}$
  - $p_0 = \text{proba of being obese when not drinking Coca-Cola}$
  - $p_1/p_0 = 1.6$ (Ludwig et al. 2001)

- “Individual event” approach
  - A medical expertise shows that Mr. A’s obesity is caused by an hormonal disease and is completely independent of his drinking Coca-Cola.
  - $p_1/p_0 = 1$.

The attribution conclusion for the chosen event definition (obesity) does not correspond to the situation of the singular event (Mr. A).
Event definition: analogues

\[ E = \{ \| Y - y \| \leq r \} \]

- same problem for \( r \) too large (mismatch),
- sampling becomes too costly for \( r \) too small
Event definition: conditional analogues

$$E = \{ \| Y - y \| \leq r \}$$

with small $r$ and conditioning to make sampling easier

but does not mean the same thing
A quote on the unicity of events

“You cannot step into the same river twice, for other waters are continually flowing in.”

Heraclitus of Ephesus
535-475 BC

Every event is unique.

“We suggest that a different framing is desirable which asks why such extremes unfold the way they do.”
(Trenberth et al. 2015)
Proposed definition: a singular event is a realization

\[ E = \{ Y = y \} \]

We want to identify all the features of this unique event that reveal the causal influence of a forcing, with no a priori.

These may be different every time and should be observation-dependent.
Causal implications

\[ E = \{ Y = y \} \]

The event is unique and has probability zero in both worlds:

\[ p_1 = p_0 = 0 \]

Hence the probability of sufficient causation is always zero:

\[ PS = \max\{ 1 - \frac{1 - p_1}{1 - p_0}, 0 \} = 0 \]

What happens to the probability of necessary causation?

\[ PN = \max\{ 1 - \frac{p_0}{p_1}, 0 \} \]
Causal implications

- Let us take the radius $r$ to zero. We have:

$$p_0 \approx f_0(y) \, dV_r$$
$$p_1 \approx f_1(y) \, dV_r$$

where $f_1(y)$ is the PDF of the random variable $Y$ in point $y$ in the factual world and $f_0(y)$ is the same quantity in the counterfactual one.

$$PN \rightarrow \max\{1 - \frac{f_0(y)}{f_1(y)}, 0\}$$
Deriving the likelihood of the event \( \{Y=y\} \) in both worlds

**Objectives:**
- comparing the PDF (or likelihood) of \( y \) in both worlds.
- identifying the features creating a gap.

**Difficulties:**
- \( y \) is a very high-dimensional vector (\( N \sim 10^8 \)).
- the same event will never be randomly simulated.

**Proposed solution:** Data Assimilation.
Outline

- Causal theory
- Event definition
- Attribution using data assimilation
Outlook of Data Assimilation: origins

**Observations:** multiple sensors

**State vector:** atmospheric model

Numerical Weather Prediction requires to **initialize** the model every six hours with **new observations**.
**Outlook of Data Assimilation: evolution**

**Trend:** expansion towards new applications, general framework for interfacing large models and observations.

**Examples:**

- **initialization:**
  - weather forecast
  - climate prediction (seasonal to decadal)
  - nowcasting (storm, tornadoes)
- **reconstruction:**
  - reanalysis
  - paleoclimate
  - carbon cycle
  - Earth’s core
  - Oil reservoir
- **estimation:** model parameters
- **DADA:** real time causal analysis of weather events

**Observations:** multiple sensors

**State vector:** large physical model
Outlook of Data Assimilation: algorithms

Observations: multiple sensors

State vector: atmospheric model

Goal: deriving the PDF of $X$ conditional on $Y = y$

high dimensional Bayesian update.
The “primitive equations” of data assimilation

Assumptions:
Hidden Markov model

- Dynamic equation:

\[ X_{t+1} = M(X_t, F_t) + v_t \]

- Observational equation:

\[ Y_t = H(X_t) + w_t \]

- \( v_t \) and \( w_t \) Gaussian error terms with covariance \( Q \) and \( R \);
- \( M \) is the model with \( F_t \) external forcing;
- \( H \) is the observation operator.
The “primitive equations” of data assimilation

**Assumptions:**
Hidden Markov model

- **Dynamic equation:**
  \[ X_{t+1} = M(X_t, F_t) + v_t \]
- **Observational equation:**
  \[ Y_t = H(X_t) + w_t \]
- \( v_t \) and \( w_t \) Gaussian error terms with covariance \( Q \) and \( R \);
- \( M \) is the model with \( F_t \) external forcing;
- \( H \) is the observation operator.

**Solution:**
Gaussian linear approximation

- **Propagation equation:**
  \[ x_{t+1}^f = M x_t^a \]
  \[ P_{t+1}^f = M P_t^a M' + Q \]
- **Update equation:**
  \[ x_t^a = x_t^f + K(y_t - H x_t^f) \]
  \[ P_t^a = (I - KH) P_t^f \]
  \[ K = P_t^f H' (HP_t^f H' + R)^{-1} \]
The likelihood $f(y)$ is a by-product of data assimilation

Solution: Gaussian linear approximation

- Propagation equation:
  
  $x_{t+1}^f = Mx_t^a$
  
  $P_{t+1}^f = MP_t^a M' + Q$

- Update equation:
  
  $x_t^a = x_t^f + K(y_t - Hx_t^f)$
  
  $P_t^a = (I - KH)P_t^f$
  
  $K = P_t^f H'(HP_t^f H' + R)^{-1}$

By-product: PDF of observation $y$

Likelihood equation:

$$- \log f(y) = \sum_{t=0}^{T} \frac{1}{2} \log |\Sigma_t|$$

$$+ \frac{1}{2} d_t' \Sigma_t^{-1} d_t$$

with:

$$d_t = y_t - Hx_t^f$$

$$\Sigma_t = HP_t^f H' + R$$
Experiments in the forced Lorenz model
Experiments in the forced Lorenz model

- DADA likelihoods of each event’s trajectory are derived:
Experiments in the forced Lorenz model

- DADA likelihoods of each event’s trajectory are derived:

The factual reconstruction is usually slightly better than the counterfactual one.
Experiments in the forced Lorenz model

- The factual reconstruction is usually slightly better than the counterfactual one.
- Small local differences pile up into a large amount of causal evidence overall.
Experiments in the forced Lorenz model

Distribution of PN for several individual events sampled within a fixed class of events having PN = 0.7
Performance of an event attribution method: a proposal

- For a given class of events, simulate singular events $y$ using the factual model and then using the counterfactual model.

- Determine which model was used to simulate the observed event $y$ based on $p_1$ and $p_0$.

- Measure the discriminative performance (ROC curve and Gini index) of this procedure.
Experiments in the forced Lorenz model

- The performance of DADA is better overall (ROC curve, Gini index).
Experiments in the ICTP AGCM model
Experiments in the ICTP AGCM model

mse of analysis:

surface temperature

wind

geopotential height
Summary

- Causal theory can help.
- An individual event can be defined as a zero probability realization.
- Causal diagnostic of singular events may be implemented in an automated, real time and systematic way using DA.
- Synergy with existing infrastructure at NWP centers.
- The likelihood may also be used for the purpose of model evaluation.
- Research under way:
  - Experiments using larger models (ICTP AGCM, WRF)
  - Implementation on real case studies.
  - Theoretical developments for computing the PDF. (Carrassi et al. 2016)
Thank you

Founders of Dadaïsm, Zürich, 1915
Inference from numerical simulations

very rare events may never occur even in a large ensemble.

example: estimates of the return time of the 2003 European heatwave range from 500 years to 40,000 years.
Inference from numerical simulations – the issue

- Estimating a small probability through brute force direct simulation:

- Generate a large number of samples and count:

\[
\hat{p}_E = \frac{1}{N} \sum_{i=1}^{N} 1_E(Y^{(i)})
\]

- Variance and relative error:

\[
\begin{align*}
\text{Var}(\hat{p}_E) &= \frac{\hat{p}_E(1 - \hat{p}_E)}{N} \approx \frac{\hat{p}_E}{N} \\
\text{RE} &= \frac{1}{\sqrt{\hat{p}_E N}}
\end{align*}
\]

To estimate a return time of ~5000 years with a 10% error, one needs a 500,000 years of simulation.
Rare event simulation is a thematical field in statistics

- Main underlying idea: **importance sampling**

- Assume we can sample \( \tilde{Y} \) from the PDF:

\[
\tilde{f}(\tilde{Y}) = f(\tilde{Y}) \cdot L(\tilde{Y})
\]

- Then we can derive a new estimator:

\[
\tilde{p}_E = \frac{1}{N} \sum_{i=1}^{N} 1_E(\tilde{Y}^{(i)}) \cdot L^{-1}(\tilde{Y}^{(i)})
\]

- The variance may be substantially improved with an adequate \( L \):

\[
\mathbb{V}(\tilde{p}_E) \ll \mathbb{V}(\hat{p}_E)
\]
Rare event simulation is a thematical field in statistics

- **Main underlying idea:** importance sampling

\[ \tilde{f}(\tilde{Y}) = f(\tilde{Y}) \cdot L(\tilde{Y}) \]

- **For a stochastic process:** interacting particle algorithm
Rare event simulation – illustration in the L95 model

- Dynamic equation:

\[
\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F
\]

- A trajectory:
Rare event simulation – illustration in the L95 model

- The function $\varphi$ is the total energy.
- The interacting particle algorithm is applied.
Rare event simulation – illustration in the L95 model

- The function $\phi$ is the total energy.
- The interacting particle algorithm is applied.
- Probabilities of exceedances are accurately estimated at a much lower computational expense.
Take aways

- It is possible to use importance sampling in dynamical systems (e.g. interacting particle algorithm).

- For the deterministic L95 model, probabilities of order $10^{-16}$ can be sampled as easily as probabilities of order $10^{-3}$.

- This technique is promising for realistic climate models.

- Theoretical research and numerical experiments are needed.