

Annex 1 The Concept of Detection and Attribution

Armineh Barkhordarian

The main goal of ‘detection and attribution’ analysis in relation to climate change is to determine whether climate has changed as a direct result of human activity. The general philosophy of detection and attribution is first to assume that the variability in a dataset can be linearly decomposed into two components: variability that cannot be attributed to any known causes and variability that can in principle be attributed to externally forced climate change signals. When all physically plausible forcing mechanisms are considered to have been accounted for then the former will indicate the internal variability. This can be estimated either from control runs with global-scale climate models or from the observed record.

Detection and attribution is therefore a two-step process in which the first step – detection – establishes whether the observed change is significantly different from changes due to internal variability alone and the second step – attribution – identifies the mix of external forcing mechanisms that best explain the change detected. Detection can be phrased as a statistical test with the null hypothesis being that the observed change is due to internal variability alone. Attribution, in contrast, cannot be phrased as a statistical test but rather as a plausibility assessment.

Different statistical methods have been used in the detection and attribution context, such as correlation-based methods (Santer et al. 1993), Bayesian methods (Hasselmann 1998) and the optimal detection method (also known as ‘optimal fingerprinting’), which can be considered as a generalised multivariate linear regression (Hasselmann 1997; Hegerl et al. 1997). A short description of the optimal detection method (optimal fingerprinting) is now given.

A1.1 Standard approach

The standard detection model (e.g. Hasselmann 1997; Hegerl et al. 1997), assuming a noise-free model-simulated response pattern is a linear regression problem as follows:

$$y = \sum_{i=1}^m x_i a_i + u_0 = Xa + u_0 \quad (1)$$

where y is the observed record, which is expressed as a linear sum of m model-simulated response patterns, x_i , with scaling factor a_i plus internally generated variability u_0 . Matrix X contains the estimated response to the external forcings (signals) under investigation. Vector a is a vector of scaling factor that adjusts the amplitude of signal patterns by accounting for uncertainty in the amplitude of the external forcing and for the possibility that the amplitude of the climate model response to the forcing may not be correct.

Fitting the regression model (Eq. 1) requires an estimate of the climate’s natural (internal) variability u_0 . The observed record is not long enough to give a reliable estimate of internal variability. Another difficulty is that the observed data may also contain the effect of external forcings. Thus, long control integrations of global-scale climate models, which are pre-industrial control experiments with all forcings held constant, are typically used for this purpose. Therefore, the distribution of the scaling factor a_i is assessed from fits of the regression model to non-overlapping control run segments derived from long control simulations. In this approach an ordinary least squares (OLS) method is used to estimate the scaling factors a_i .

In the standard approach, detection of a climate change signal occurs if any of the scaling factors a_i are shown to be significantly different from zero. This is handled by testing the null hypothesis $H_{DE}: a=0$ (where 0 is a vector of zeros). If the null hypothesis H_{DE} is rejected, it indicates that the observed representation of climate change deviates significantly from internal variability, that is, the observed change y cannot be explained by internal variability u_0 alone.

Once detection has been established, attribution (consistency of observed changes with a combination of external forcing) is assessed by testing the null hypothesis $H_{AT}: a=1$ (where 1 denotes a vector of unit). When there is insufficient evidence to reject H_{AT} , the attribution of changes to the respective forcing is claimed.

A1.2 Accounting for noise in model-simulated response patterns

The standard approach to optimal fingerprinting assumes that the model-simulated response pattern is known exactly, that is, it is not subject to sampling uncertainty (repeating the experiment would yield an identical pattern). Thus, the response patterns are specified by averaging the response to a particular forcing from an ensemble of runs of a climate model (Hegerl et al. 1997). The problem is that even with a few member ensembles, sampling uncertainty is still far from negligible, particularly in weak signal-to-noise situations such as the analysis of the response to solar forcing (Allen and Stott 2002). Therefore, the regression model (Eq. 1) is extended by an additional error term $\mathbf{u}_i$ that accounts for internal variability in the response patterns. Thus, the simple statistical model (Eq. 1) is replaced by:

$$y = \sum_{i=1}^m (x_i - u_i) a_i + u_0 \quad (2)$$

where x_i is the i^{th} model-simulated response pattern estimated from a finite ensemble and therefore contaminated with sampling noise $\mathbf{u}_i$ and $\mathbf{u}_0$ is the noise in the observation.

As described by Allen and Stott (2002), the presence of this internal variability (noise) in the signal patterns may bias OLS estimates of scaling factor ($\mathbf{a}$) towards zero slope, particularly if only a small ensemble is available to estimate signals. Using a total least squares (TLS) algorithm solves this problem. TLS minimises the perpendicular distance between the scatter points and the best-fit line, not the vertical distance minimised by OLS. Therefore, the bias of best estimate towards zero in OLS is removed by using TLS in optimal fingerprinting.

A1.3 Signal-to-noise optimisation

As described by Hasselmann (1993), an optimal fingerprint method further requires a signal-to-noise optimisation. The scaling factors $\mathbf{a}_i$ are optimal estimators only if the internal variability is independent and identically distributed (iid). Therefore, the original observations and model data cannot be used in the regression model (Eq. 2) and a transformed version resulting in iid residuals is needed. A transformation is given by using a truncated set of the empirical orthogonal functions (EOFs) of the control simulations (weighted by the square root of their eigenvalues). Thus, the control simulations are used to present the original data in a dimension-reduced EOF space. A residual test proposed by Allen and Tett (1999) is usually performed to assess whether the control simulations adequately represent the variability of the data in the truncated space.

References

- Allen MR, Stott PA (2002) Estimating signal amplitudes in optimal fingerprinting, part I: Theory. *Clim Dynam* 21:477-491
- Allen MR, Tett SFB (1999) Checking for model consistency in optimal fingerprinting. *Clim Dynam* 15:419-434
- Hasselmann K (1993) Optimal fingerprints for the detection of time-dependent climate-change. *J Clim* 6:1957-1971
- Hasselmann K (1997) Multi-pattern fingerprint method for detection and attribution of climate change. *Clim Dynam* 13:601-611
- Hasselmann K (1998) Conventional and Bayesian approach to climate-change detection and attribution. *Q J Roy Meteorol Soc* 124:2541-2565
- Hegerl GC, Hasselmann K, Cubasch U, Mitchell JFB, Roeckner E, Voss R, Waszkewitz J (1997) Multi-fingerprint detection and attribution analysis of greenhouse gas, greenhouse gas-plus-aerosol and solar forced climate change. *Clim Dynam* 13:613-634
- Santer BD, Wigley TML, Jones PD (1993) Correlation methods in fingerprint detection studies. *Clim Dynam* 8:265-276