Guidance note on best statistical practices for TOAR analyses

by Chang et al.

IDC 633 Presentations

14.04.2024

Review by Subham Das (MS20121)

IISER Mohali

Purpose of this article

The aim of this article [1, Chang] is to present recommendations on best statistical practices ensuring consistent communication of statistical analysis and associated uncertainity across TOAR (Tropospheric Ozone Assesment Report) publications.

Purpose of this article

The aim of this article [1, Chang] is to present recommendations on best statistical practices ensuring consistent communication of statistical analysis and associated uncertainity across TOAR (Tropospheric Ozone Assesment Report) publications.

Scope: Approaches for reporting trends, a discussion of strengths and weaknesses of common techniques, and calibrated language for the communication uncertainity.

Purpose of this article

The aim of this article [1, Chang] is to present recommendations on best statistical practices ensuring consistent communication of statistical analysis and associated uncertainity across TOAR (Tropospheric Ozone Assesment Report) publications.

Scope: Approaches for reporting trends, a discussion of strengths and weaknesses of common techniques, and calibrated language for the communication uncertainity.

Focus: Recommendations on trend analysis, numbered from **R1** to **R9**.

Purpose of this article

The aim of this article [1, Chang] is to present recommendations on best statistical practices ensuring consistent communication of statistical analysis and associated uncertainity across TOAR (Tropospheric Ozone Assesment Report) publications.

Scope: Approaches for reporting trends, a discussion of strengths and weaknesses of common techniques, and calibrated language for the communication uncertainity.

Focus: Recommendations on trend analysis, numbered from **R1** to **R9**.

Why trend analysis? : The purpose of trend analysis can be defined as *detecting and attributing the change and its uncertainty of the statistical properties in a time series of a predefined variable.*

R1: *Practitioners are recommended to provide quantitative trends and should avoid drawing conclusions when the magnitude of trends and associated uncertainties are not explicitly quantified*

R1: *Practitioners are recommended to provide quantitative trends and should avoid drawing conclusions when the magnitude of trends and associated uncertainties are not explicitly quantified*

Types of trend detection techniques:

- **Linear methods:** More quantitative in nature due to simpler calculations, rarer due to nonlinear nature of ozone assesments

R1: *Practitioners are recommended to provide quantitative trends and should avoid drawing conclusions when the magnitude of trends and associated uncertainties are not explicitly quantified*

Types of trend detection techniques:

- **Linear methods:** More quantitative in nature due to simpler calculations, rarer due to nonlinear nature of ozone assesments
- **Non-Linear methods:** Provides better qualitative insight, although hard to quantify uncertainity and trend values due to complex nature. Risks overfitting if underlying processes are not understood during quantitative assesment

R1: *Practitioners are recommended to provide quantitative trends and should avoid drawing conclusions when the magnitude of trends and associated uncertainties are not explicitly quantified*

Types of trend detection techniques:

- **Linear methods:** More quantitative in nature due to simpler calculations, rarer due to nonlinear nature of ozone assesments
- **Non-Linear methods:** Provides better qualitative insight, although hard to quantify uncertainity and trend values due to complex nature. Risks overfitting if underlying processes are not understood during quantitative assesment

Way-out: For short-term nonlinearity, incorporate relevant covariates into regression-based methods attributing data variability. For long-term nonlinear change, use piecewise linear trends for change point analysis due to explicit quantification.

R2: *Visualize data series and/or try non-linear trend detection methods to assess the validity of a linear trend model before reporting linear trends quantitatively.*

R3: *When time series exhibit clearly visible changes in the magnitude of a trend over time, we recommend the application of piecewise linear trends in combination with a change point detection algorithm.*

R4: *In TOAR-II, we recommend quantile regression (QR) for trend analysis.*

R4: *In TOAR-II, we recommend quantile regression (QR) for trend analysis.*

Quantile regression models the relationship between a set of predictor (independent) variables and specific percentiles (or "quantiles") of a target (dependent) variable, most often the median. It has two main advantages over Ordinary Least Squares regression:

R4: *In TOAR-II, we recommend quantile regression (QR) for trend analysis.*

Quantile regression models the relationship between a set of predictor (independent) variables and specific percentiles (or "quantiles") of a target (dependent) variable, most often the median. It has two main advantages over Ordinary Least Squares regression: **(1)** Quantile regression makes no assumptions about the distribution of the target variable.

R4: *In TOAR-II, we recommend quantile regression (QR) for trend analysis.*

Quantile regression models the relationship between a set of predictor (independent) variables and specific percentiles (or "quantiles") of a target (dependent) variable, most often the median. It has two main advantages over Ordinary Least Squares regression: **(1)** Quantile regression makes no assumptions about the distribution of the target variable. **(2)** Quantile regression tends to resist the influence of outlying observations

Reasons:

- QR is a well-suited technique for detecting heterogeneous distributional changes **which is often the case for free tropospheric and surface ozone as they typically show diverse percentile trends**

Reasons:

- QR is a well-suited technique for detecting heterogeneous distributional changes **which is often the case for free tropospheric and surface ozone as they typically show diverse percentile trends**
- QR is a regression-based method that allows us to incorporate covariates for attributing data variability and piecewise trends for change point analysis.

Reasons:

- QR is a well-suited technique for detecting heterogeneous distributional changes **which is often the case for free tropospheric and surface ozone as they typically show diverse percentile trends**
- QR is a regression-based method that allows us to incorporate covariates for attributing data variability and piecewise trends for change point analysis.

Limitation : A larger sample size is required for a valid estimation of extreme variations, which means that datasets with small or insufficient sample sizes should be limited to analysis of the median trend.

	GLS	Sen-Theil	**Quantile regression**[1]
Inference statistics of a trend	Mean	Median	**Percentile/Quantile**
Robust to small sample size		✔	*[2]
Robust to outliers		✔	✔
Makes use of all available information	✔		✔
Inference with non-normal distribution on residuals		✔	✔
Unique solution	✔		
Adjustment for non-IID residuals (e.g. Incorporation of autocorrelation)	✔		✔
Incorporation of covariates	✔		✔
Extend to piecewise trends	✔		✔
Extend to adaptive nonlinear trends	*[3]		*[3]
Extend to extreme percentile estimate			✔[2]

R5: *Choose the specified quantiles in QR regression depending on the number of samples.*

No of samples	Recommended Quantiles
< 50	50%ile (median)
50 - 90	25%ile, 50%ile, 75%ile
91 – 120	10, 25, 50, 75, 90
> 120	5, 10, 25, 50, 75, 90, 95

R5: *Choose the specified quantiles in QR regression depending on the number of samples.*

No of samples	Recommended Quantiles
< 50	50%ile (median)
50 - 90	25%ile, 50%ile, 75%ile
91 – 120	10, 25, 50, 75, 90
> 120	5, 10, 25, 50, 75, 90, 95

To facilitate comparisons of trends and uncertainties across the TOAR analyses, it is desirable to define the default quantiles to be estimated. Table 2 provides rule-of-thumb recommendations on default percentile trends to be reported according to various ranges of sample size (i.e. the number of data points fitted in the trend model, not raw data). To avoid over-interpretation, we do not recommend reporting percentile trends in intervals narrower than 5%.

R6: *Seasonal adjustments are required for time series analysis.*

R6: *Seasonal adjustments are required for time series analysis.*

Ozone and many other atmospheric variables have distinct seasonal and/or diurnal cycles. These known variations should be included in the statistical model that is used to derive trend estimates or the data should be deseasonalized prior to the trend analysis.

R6: *Seasonal adjustments are required for time series analysis.*

Ozone and many other atmospheric variables have distinct seasonal and/or diurnal cycles. These known variations should be included in the statistical model that is used to derive trend estimates or the data should be deseasonalized prior to the trend analysis.

The authors recommend preserving as much information as possible by utilizing QR to detect heterogeneous changes which is otherwise lost during data agrregation. Such aggregations is simply to reduce heterogeneous data variability and meet the homoscedasticity assumption.

R7; *Practitioners should provide associated uncertainty for any statistical estimation, typically the 95% confidence interval*

R7; *Practitioners should provide associated uncertainty for any statistical estimation, typically the 95% confidence interval*

Factors affecting uncertainity of trend estimates: (1) the larger sample size, **(2)** the smaller data variability, and/or **(3)** the weaker autocorrelation and heteroskedasticity,

R7; *Practitioners should provide associated uncertainty for any statistical estimation, typically the 95% confidence interval*

Factors affecting uncertainity of trend estimates: (1) the larger sample size, **(2)** the smaller data variability, and/or **(3)** the weaker autocorrelation and heteroskedasticity,

The independent and identically distributed (IID) is generally strong assumption autocorrelation/heteroskedasticity is common in atmospheric time series. Thus the authors use the Circular (moving) block bootstrap methodology which wraps around the start and end points of the time series, so the data near the beginning and end of the record are not systematically under-sampled.

Block bootstrap inference to account for autocorrelation in the trends uncertainty

Moving block bootstrap algorithm
Data: a time series $y_1 \ldots y_N$
Input: block length B; number of iterations J
Output: an uncertainty of the trend estimate
1. **for** $1 \leq j \leq J$ **do**
2. generate N/B random blocks from data
3. fit a trend model to the random samples
4. extract the sampled trend value b_j
5. **end for**
6. **return** the standard deviation of $\{b_j\}$

linear trends and 95% CIs
for surface ozone at Mace Head, Ireland
[in units of ppb/decade]:
OLS: 1.06 [$\pm$0.37] $\rightarrow$ [$\pm$0.50]
GLS-AR1: 1.09 [$\pm$0.53]
WLS: 0.75 [$\pm$0.35] $\rightarrow$ [$\pm$0.46]
LTS: 0.75 [$\pm$0.34] $\rightarrow$ [$\pm$0.54]
Lasso: 1.05 [$\pm$0.20] $\rightarrow$ [$\pm$0.52]
Ridge: 1.03 [$\pm$0.21] $\rightarrow$ [$\pm$0.53]
Sen-Theil: 0.85 [$\pm$0.05] $\rightarrow$ [$\pm$0.53]
Siegel: 0.70 [$\pm$0.16] $\rightarrow$ [$\pm$0.61]
QR-50th: 0.71 [$\pm$0.44] $\rightarrow$ [$\pm$0.57]
M-estimator: 0.88 [$\pm$0.20] $\rightarrow$ [$\pm$0.50]

Code outline for the moving block bootstrap algorithm for deriving the trend uncertainty from autocorrelated time series, and its impact on the trend uncertainty for different trend techniques.

R8: *Gradated calibrated language should be used as a replacement for dichotomous “statistical significance”.*

R8: *Gradated calibrated language should be used as a replacement for dichotomous "statistical significance".*

The authors propose a scale to assess uncertainity similar to the scheme used in the first phase of TOAR (coming from the IPCC guidelines) but adds another gradation for "very high certainty", approximately corresponding to a p-value $\leqslant 0.01$ or $SNR \geqslant 3$, which allows us to investigate "the world beyond $p \leqslant 0.05$"

R9: *Apply change point detection to validate any suspicious changes of data structure.*

R9: *Apply change point detection to validate any suspicious changes of data structure.*

Change point analysis is highly relevant to the detection of trend change and attribution of intervention. A change point does not have a unique meaning in the literature, it can be a break or turning point connecting two data series that are considered to have different averages, variances, and/or trends

R9: *Apply change point detection to validate any suspicious changes of data structure.*

Change point analysis is highly relevant to the detection of trend change and attribution of intervention. A change point does not have a unique meaning in the literature, it can be a break or turning point connecting two data series that are considered to have different averages, variances, and/or trends

Factors affecting the detection of trends (e.g. seasonality and autocorrelation) are relevant to the detection of change points, so deseasonalization and incorporation of necessary components in the regression model are suggested. The change point algorithms are generally not robust to small perturbations in the data

R9: *Apply change point detection to validate any suspicious changes of data structure.*

Change point analysis is highly relevant to the detection of trend change and attribution of intervention. A change point does not have a unique meaning in the literature, it can be a break or turning point connecting two data series that are considered to have different averages, variances, and/or trends

Factors affecting the detection of trends (e.g. seasonality and autocorrelation) are relevant to the detection of change points, so deseasonalization and incorporation of necessary components in the regression model are suggested. The change point algorithms are generally not robust to small perturbations in the data

Given ozone's high temporal and spatial variability, it is recommended to consider the change point analysis of trends only when data are sufficiently long.

References

Kai-Lan Chang, Martin G. Schultz, Gerbrand Koren, and Niklas Selke.

Guidance note on best statistical practices for toar analyses, 2023.