cogent geoscience

CrossMark

ATMOSPHERIC SCIENCES | RESEARCH ARTICLE

Modelling rainfall trends in England and Wales

Terence C. Mills[1]*

*Corresponding author: Terence C. Mills, School of Business and Economics, Loughborough University, Loughborough, UK
E-mail: t.c.mills@lboro.ac.uk

Reviewing editor:
Nir Krakauer, City College of New York, USA

Additional information is available at the end of the article

Abstract: The monthly England and Wales precipitation (EWP) series (once power transformed to induce symmetry and to stabilise variance) may be characterised as having linear seasonal trends with a white noise error process superimposed. However, these trends are not stable, for they are interrupted by four regime shifts occurring in 1828, 1871, 1917 and 1976. If these shifts are ignored then the series is consistent with a trend pattern in which winters are becoming increasingly wet and summers drier. If only the last regime from 1976 is considered, then summers are still becoming drier but winters have no trend, with spring becoming wetter. The unusually wet winter of 2014 is seen to have been a consequence of very high January and February rainfall relative to that predicted, the conjunction of which is unprecedented during the two and a half centuries over which the EWP series has been available, during which time such pairs of values have been essentially uncorrelated.

Subjects: Geostatistics; Mathematical Statistics; Statistics; Statistics & Computing; Statistics & Probability

Keywords: rainfall modelling; time series analysis; structural breaks

1. Introduction

Mills (2005) modelled the monthly England and Wales precipitation series (EWP) from 1766 to 2002 using a variety of trend extraction techniques and provided evidence of a trend towards wetter winters, particularly in December and January, and drier summers, concentrated in June and July, although he found that it was difficult to come to firm conclusions about trends extracted from such a noisy data series. In general, however, Mills concluded that his findings were consistent with the earlier, essentially descriptive, analysis of Alexander and Jones (2001).

ABOUT THE AUTHOR

Terence C Mills is Professor of Applied Statistics and Econometrics at Loughborough University, UK. He has applied time series techniques to a wide range of disciplines, including economics, economic history, finance, health and climatology and meteorology, where he has published in the areas of rainfall and temperature modelling in journals such as the *Journal of Climate, Climatic Change, International Journal of Climatology, Atmospheric Science Letters* and *Meteorological Applications.*

PUBLIC INTEREST STATEMENT

Monthly rainfall in England and Wales may be characterised as having linear seasonal trends with an uncorrelated (white noise) error superimposed. However, these trends are not stable, for they are interrupted by four shifts occurring in 1828, 1871, 1917 and 1976. If these shifts are ignored then the series is consistent with a trend pattern in which winters are becoming increasingly wet and summers drier. If only the last regime from 1976 is considered, then summers are still becoming drier but winters have no trend, with spring becoming wetter. The unusually wet winter of 2014 is seen to have been a consequence of very high January and February rainfall relative to that predicted, the conjunction of which is unprecedented during the two and a half centuries over which rainfall data has been available, during which time such pairs of values have been essentially uncorrelated.

cogent •oa

✳️ **cogent** •• geoscience

In the decade since this publication, attention has focused on modelling high-frequency extreme "precipitation events" in England and Wales using simulation models: see Pearson, Shaffrey, Methven, and Hodges (2015), Rhodes, Shaffrey, and Gray (2015) and Otto et al. (2015). No further analyses of the monthly EWP series have been reported, but since the original Mills study was published various techniques have become available for detecting the possible presence of breaking trends in time series, a feature that could well have an impact on a series as long as the EWP. Furthermore, the unusually wet winter of 2014, which contained the wettest January on record, provides a further impetus to reanalyse the EWP series using these new techniques and an additional (almost) 12 years of observations. Consequently, Section 2 revisits and updates Mills' parametric model of monthly EWP for the period 1766–2014, extending it to allow for both deterministic and stochastic seasonal and trend behaviour. Section 3 subjects the model to various structural break tests to ascertain the presence of breaking trends and estimates the implied regime break points and associated trends. The unusually high winter rainfall of 2014 is investigated in Section 4, in which the monthly observations for December, January and February are compared to their forecasted values and the extent of their "unusualness" is calibrated. A summary of the results is offered in Section 5.

2. A model for monthly precipitation

Following Mills (2005), the basic model for the monthly EWP series takes the form

$$x_t^{(\lambda)} = \sum_{i=1}^{12} (\alpha_i s_{i,t} + \beta_i s_{i,t} t) + u_t \quad t = 1, 2, \ldots, T \tag{1}$$

In Equation (1), rainfall x_t is transformed by the Box and Cox (1964) power transformation

$$x_t^{(\lambda)} = \frac{x_t^\lambda - 1}{\lambda} \quad x_t^{(0)} = \log x_t$$

to ameliorate the skewness in the raw data, a consequence of x_t being bounded below at zero, and to induce linearity and constancy of variance. The estimate of 0.6 for the transformation parameter λ obtained by Mills (2005) is again used here. The $s_{i,t}$, $t = 1, 2, \ldots, T$, are "dummy" variables defined to take the value 1 in month i and 0 elsewhere (where $i = 1$ signifies January, etc.). Their inclusion allows a deterministic monthly pattern to be modelled. The presence of the $s_{i,t} t$ "interaction" variables allows for the possibility of different monthly linear time trends. The α_i and β_i parameters measure the intercept and slope of these trends, so that if $\beta_i \neq 0$ then the seasonal pattern for month i evolves linearly over time. As the sample period is January 1766–December 2014, there is a total of $T = 2,988$ observations: the data are taken from the UK Met Office (Hadley Centre) website at http://www.metoffice.gov.uk/hadobs/hadukp/data/monthly/HadEWP_monthly_qg.txt.

The error u_t can, in general, follow a seasonal ARMA process (see, e.g. Mills, 2014).

$$\phi(B)\Phi(B^{12}) u_t = \theta(B)\Theta(B^{12}) a_t \tag{2}$$

where

$$\phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$$

$$\theta(B) = 1 + \theta_1 B + \ldots + \theta_q B^q$$

$$\Theta(B^{12}) = 1 - \Phi_1 B^{12} - \ldots - \Theta_P B^{12P}$$

$$\Theta(B^{12}) = 1 + \Theta_1 B^{12} + \ldots + \Theta_Q B^{12Q}$$

are polynomials in the lag operator B defined such that $B^j a_t \equiv a_{t-j}$ and a_t is zero mean white noise with variance σ_a^2. The presence of the polynomials $\Phi(B^{12})$ and $\Theta(B^{12})$ allows the error to be seasonally (as well as non-seasonally) autocorrelated.

2.1. Deterministic and stochastic trends and seasonality

More general models result if unit roots are allowed in the $\phi(B)$ and $\Phi(B^{12})$ polynomials. If the non-seasonal autoregressive polynomial contains a unit root, so that it can be factorised as

$$(1 - B)(1 - \phi_1^* B - \ldots - \phi_{p-1}^* B^{p-1}) = (1 - B)\phi^*(B),$$

then Equation (1) becomes, with $\nabla = 1 - B$,

$$\nabla x_t^{(\lambda)} = \sum_{i=1}^{12} (\alpha_i \nabla s_{i,t} + \beta_i \nabla s_{i,t} t) + u_t^* \quad \phi^*(B) u_t^* = \theta(B)\Theta(B^{12}) a_t \tag{3}$$

Noting that $\nabla s_{i,t} = s_{i,t} - s_{i+1,t}$ and $\nabla s_{i,t} t = (12 + i)(s_{i,t} - s_{i+1,t})$, where it is taken that $\nabla s_{12,t} = s_{12,t} - s_{1,t+1}$, Equation (3) becomes

$$\nabla x_t^{(\lambda)} = \sum_{i=1}^{12} (\alpha_i + (12 + i)\beta_i)(s_{i,t} - s_{i+1,t}) + u_t^* \tag{4}$$

Thus, $x_t^{(\lambda)}$ contains a stochastic, random walk, trend with differing seasonal drifts, i.e. each month evolves as a random walk with its own drift.

Alternatively, suppose that the seasonal autoregressive polynomial contains a unit root:

$$(1 - B^{12})(1 - \Phi_1^* B^{12} - \ldots - \Phi_{P-1}^* B^{12(P-1)}) = (1 - B^{12})\Phi^*(B^{12}).$$

Equation (1) then becomes, with $\nabla_{12} = (1 - B^{12})$,

$$\nabla_{12} x_t^{(\lambda)} = \sum_{i=1}^{12} (\alpha_i \nabla_{12} s_{i,t} + \beta_i \nabla_{12} s_{i,t} t) + u_t^* \quad \Phi^*(B) u_t^* = \theta(B)\Theta(B^{12}) a_t \tag{5}$$

Now $\nabla_{12} s_{i,t} = 0$ and $\nabla_{12} s_{i,t} t = 12 s_{i,t}$, so that (5) becomes

$$\nabla_{12} x_t^{(\lambda)} = 12 \sum_{i=1}^{12} \beta_i s_{i,t} + u_t^* \tag{6}$$

and $x_t^{(\lambda)}$ now contains a stochastic seasonal random walk with differing seasonal drifts.

If $\Phi(B^{12}) = \Theta(B^{12})$ then there is only deterministic seasonality (see Pierce, 1978; Mills & Mills, 1992, for similar set-ups and additional analysis).

2.2. Estimating and testing the model

To determine the most appropriate form of the combined model (1) and (2), initial analysis using the information from the sample autocorrelation and partial autocorrelation functions along with residual diagnostic checks from fitted models established that the polynomial orders could be set at $p = P = Q = 1$ and $q = 0$, leading to the model

$$x_t^{(0.6)} = \sum_{i=1}^{12} (\alpha_i s_{i,t} + \beta_i s_{i,t} t) + \frac{(1 - \Theta B^{12})}{(1 - \phi B)(1 - \Phi B^{12})} a_t \tag{7}$$

The estimates of the error specification are $\hat{\phi} = 0.034 \pm 0.020$, $\hat{\Phi} = -0.951 \pm 0.006$ and $\hat{\Theta} = 0.989 \pm 0.002$, clearly showing the absence of both a stochastic trend (since $\phi < 1$) and

stochastic seasonality (since $\Phi \approx -\Theta$). The residuals from this model exhibit no autocorrelation and little evidence of non-normality. The question of whether the deterministic seasonal model contained a non-linear component was addressed by including additional quadratic trends, taking the form $s_{i,t}t^2$, but these were found to be insignificant.

Since the estimate of ϕ is small and only significant at approximately the 8% level, this parameter is set to zero in subsequent models, along with the restriction $\Phi = -\Theta$: the models thus have a white noise error $u_t = a_t$. Table 1 reports the ordinary least squares (OLS) estimates of Equation (7) under this error specification, along with a restricted model in which seven insignificant trends have been excluded. A likelihood ratio test of these seven zero restrictions has a marginal significance level of just 0.75 and imposing them has no effect on the remaining trend estimates but slightly improves the precision of the seasonal coefficients associated with the deleted trends. Although all the remaining estimated coefficients are highly significant, the R^2 statistic shows that the model explains just 13.6% of the variation in (transformed) rainfall with an accompanying regression standard error of 23.7, which is some 28% of the mean of the series.

Table 1. OLS estimates of Equation (7)		
	Unrestricted	**Restricted**
	Estimate (Standard error)	**Estimate (Standard error)**
α_1	77.305 (3.071)	77.305 (3.067)
α_2	75.522 (2.923)	77.014 (1.554)
α_3	67.211 (2.819)	67.211 (2.816)
α_4	71.279 (2.850)	72.840 (1.379)
α_5	78.433 (2.726)	77.156 (1.299)
α_6	81.581 (2.882)	78.495 (1.424)
α_7	96.112 (3.227)	96.112 (3.223)
α_8	90.844 (2.828)	90.453 (1.429)
α_9	95.183 (3.378)	95.183 (3.374)
α_{10}	97.551 (3.326)	98.889 (1.639)
α_{11}	94.087 (3.034)	97.393 (1.546)
α_{12}	83.416 (3.369)	83.416 (3.365)
β_1	0.00774 (0.00182)	0.00773 (0.00182)
β_2	0.00010 (0.00179)	–
β_3	0.00457 (0.00164)	0.00457 (0.00164)
β_4	0.00100 (0.00174)	–
β_5	−0.00086 (0.00153)	–
β_6	−0.00207 (0.00172)	–
β_7	−0.00561 (0.00174)	−0.00561 (0.00173)
β_8	−0.00026 (0.00170)	–
β_9	−0.00048 (0.00196)	−0.00448 (0.00196)
β_{10}	0.00090 (0.00196)	–
β_{11}	0.00290 (0.00180)	–
β_{12}	0.00668 (0.00188)	0.00668 (0.00187)
R^2	0.136	0.135
$\hat{\sigma}_a$	23.725	23.714

Note: Standard errors are robust to both autocorrelation and heteroskedasticity; rainfall scaled as mm × 10.

The trend estimates are significantly positive for January, March and December and significantly negative for July and September (the first four at marginal significance levels of 0.5% and smaller, the last at a level of 2.2%). Using these estimates, monthly trends may be calculated as

$$\tau_{i,t} = (0.6(\hat{\alpha}_i + 12\hat{\beta}_i) + 1)^{1.6667} \tag{8}$$

The average trend increases for January, March and December are 0.131, 0.070 and 0.121 mm, respectively, while the trend decreases for July and September are −0.095 and −0.070 mm. Thus, consistent with the findings of Mills (2005) and the earlier study of Alexander and Jones (2001), the evidence points to a tendency for increasingly wet winters and increasingly dry summers.

3. Testing for structural breaks

The trends calculated from Equation (8), when they are statistically significant, are essentially linear, so that the annual changes in trends for each month have been effectively constant (and possibly zero) throughout the almost two and half centuries of the sample period. The estimation of Equation (7) does, however, assume that the model has remained stable throughout this very long period, and this is an assumption that certainly needs checking.

Bai (1997), Bai and Perron (1998, 2003a, 2003b), and Perron (2006) develop and survey a variety of tests for examining the regression constancy assumption against the alternative that the model has m potential breaks (and hence $m + 1$ regimes), perhaps occurring at unknown times. A selection of these tests are available in EViews 8 (2013), which provides a convenient description of them, these being applied with a maximum of five breaks to Equation (7) and a "trimming percentage" of 15%, which sets the minimum length of a regime to be approximately 37 years. All tests produce a common set of four break points (defined as the *start date* of the new regime) at March 1828, January 1871, December 1917 and May 1976.[1]

A set of restricted models were then developed for each of the five regimes. Rather than report each estimated set of coefficients, the implied trends computed from Equation (8) are shown in Figure 1 aggregated into the four seasons and overlain on the "global seasonal trend" estimated from the coefficients of Table 1. As might be expected, the global seasonal trends average out the trends in each of the five regimes. However, the trends do shift across the regimes in interesting ways. Spring and autumn have had the most stable trends, with annual changes of 0.024 and −0.023 mm across the whole sample period. The spring trend had an essentially constant slope (0.024 ± 0.001 mm) until the final break in 1976, whereupon there has been a declining trend with a slope of −0.418 mm. The autumn trends were declining over the first three regimes, but for the last 100 years the trend has been flat. Winter has shown the largest trend increase of 0.084 mm per annum, with the regime trends being positive for the first three regimes but again essentially flat for the last 100 years. It is summer that has shown the most volatility in trend movements. An overall trend decline of −0.032 mm per annum masks much larger declines, of −1.124 and −0.526 mm, in the second and third regimes, and an increasing trend of 0.765 mm in the regime since 1976. Thus, if attention is only focused on this last regime, then winter and autumn rainfall trends have been constant, the spring trend has been negative and the summer trend has been positive.

4. How unusual was the winter of 2014?

The rainfall amounts for December 2013, January 2014 and February 2014 were 134.2, 184.6 and 136.7 mm, respectively, and during these months a great deal of flooding, with associated damage, occurred in parts of England and Wales. How unusual are these amounts? The December and February values have been exceeded by 7 and 6.25%, respectively, of the complete sample while the January value has been exceeded by just 0.5% (16 of the 2983 months for which data are available), but how extreme are these values given the models fitted to the data?

For the (restricted) global model of Table 1, the standardised residuals ($\hat{u}_t/\hat{\sigma}_a$) for these three months are 0.864, 2.105 and 2.031. Since these may be considered to be (independent) drawings

Figure 1. Seasonal trends.
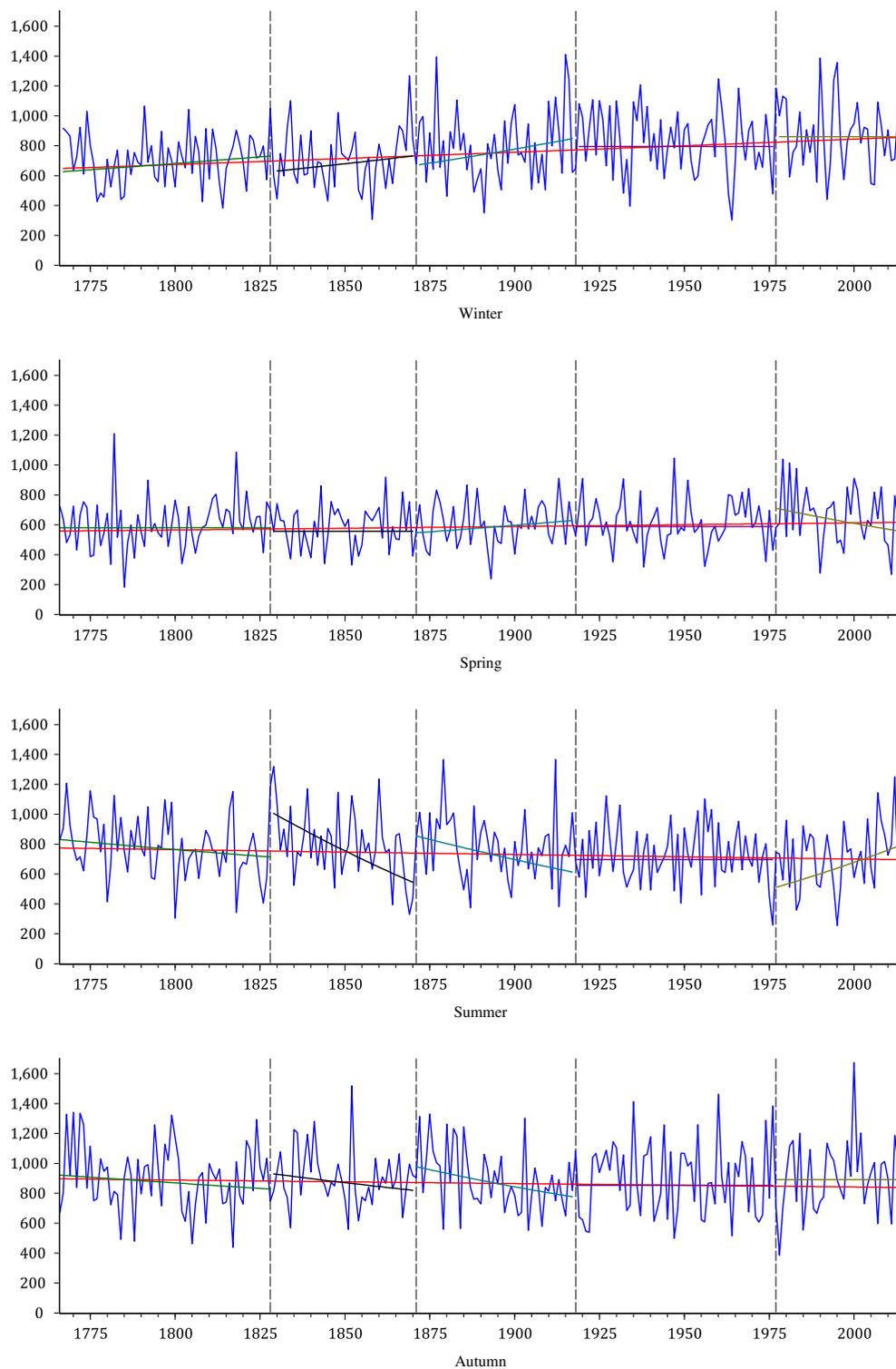
from a standard normal distribution, the probabilities of observing values at least as large as these are 0.194, 0.018 and 0.021, so that the January and February values, although relatively large, are by no means unusually extreme. For the model fitted to the last regime, the standardised residuals are 0.831, 2.209 and 1.855 with associated probabilities of 0.203, 0.014 and 0.032, reasonably consistent with the global model.

Figure 2. Scatterplot of standardised January and February residuals, 1766–2014.
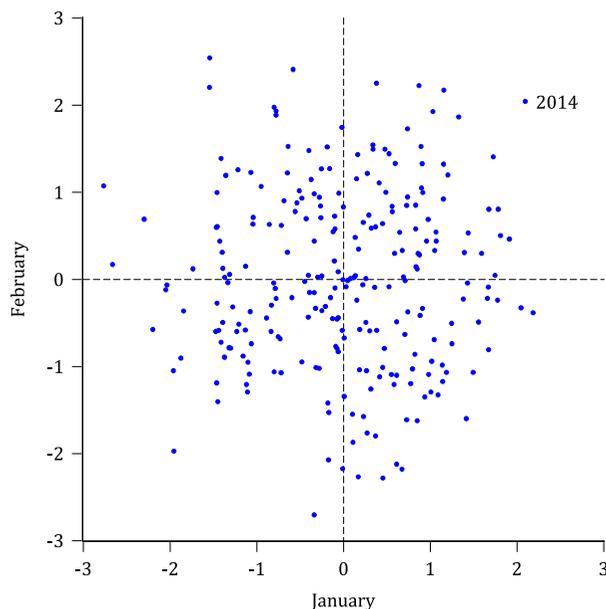
The correlation between the standardised January and February residuals over the entire period from 1766 to 2014 is just 0.02, so that the conjunction of two such large residuals in any single year, as has occurred in 2014, is unprecedented. This can be seen clearly from Figure 2, which presents a scatterplot of this pair of residuals for each year of the sample period.

It would thus appear that the extreme rainfall during the winter of 2014 is a consequence of unusually high, relative to predicted, January and February rainfall, the conjunction of which has been unprecedented in the last 250 years.

5. Summary

The monthly EWP series (once power transformed to induce symmetry and to stabilise variance) may be characterised as having linear seasonal trends with a white noise error process superimposed. However, these trends are not stable, for they are interrupted by four regime shifts occurring in 1828, 1871, 1917 and 1976. If these shifts are ignored, then the series is consistent with a trend pattern in which winters are becoming increasingly wet and summers drier. If only the last regime from 1976 is considered, then summers are still becoming drier but winters have no trend, with spring becoming wetter.

The unusually wet winter of 2014 is seen to have been a consequence of very high January and February rainfall relative to that predicted, the conjunction of which is unprecedented during the two and a half centuries over which the EWP series has been available, during which such pairs of values have been essentially uncorrelated.

Author details
Terence C. Mills[1]
E-mail: t.c.mills@lboro.ac.uk
[1] School of Business and Economics, Loughborough University, Loughborough, UK.

Note
1. The tests employed were the "Global *L* breaks versus none", the "Sequential *L* + 1 breaks versus *L*", the "Sequential test of all subsets" and the "*L* + 1 versus global *L*".

References
Alexander, L. V., & Jones, P. D. (2001). Updated precipitation series for the UK and discussion of recent extremes. *Atmospheric Science Letters, 1*, doi:10.1006/asle.001.0025
Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric Theory, 13*, 315–352. http://dx.doi.org/10.1017/S0266466600005831

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica, 66,* 47–78. http://dx.doi.org/10.2307/2998540

Bai, J., & Perron, P. (2003a). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics, 18,* 1–22. http://dx.doi.org/10.1002/(ISSN)1099-1255

Bai, J., & Perron, P. (2003b). Critical values for multiple structural change tests. *The Econometrics Journal, 6,* 72–78. http://dx.doi.org/10.1111/ectj.2003.6.issue-1

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B, 26,* 211–246.

EViews 8. (2013). *EViews 8 user guide* (Chapter 30, pp. 369–388). Irvine, CA: HIS Global.

Mills, T. C. (2005). Modelling precipitation trends in England and Wales. *Meteorological Applications, 12,* 169–176. http://dx.doi.org/10.1017/S1350482705001611

Mills, T. C. (2014). Time series modelling of temperatures: an example from Kefalonia. *Meteorological Applications, 21,* 578–584. http://dx.doi.org/10.1002/met.2014.21.issue-3

Mills, T. C., & Mills, A. G. (1992). Modelling the seasonal patterns in UK macroeconomic time series. *Journal of the Royal Statistical Society. Series A (Statistics in Society), 155,* 61–75. http://dx.doi.org/10.2307/2982669

Otto, F. E. L., Rosier, S. M., Allen, M. R., Massey, N. R., Rye, C. J., & Quintana, J. I. (2015). Attribution analysis of high precipitation events in summer in England and Wales over the last decade. *Climatic Change, 132,* 77–91. http://dx.doi.org/10.1007/s10584-014-1095-2

Pearson, K. J., Shaffrey, L. C., Methven, J., & Hodges, K. I. (2015). Can a climate model reproduce extreme regional precipitation events over England and Wales? *Quarterly Journal of the Royal Meteorological Society, 141,* 1466–1472. http://dx.doi.org/10.1002/qj.2015.141.issue-689

Perron, P. (2006). Dealing with structural breaks. In T. C. Mills & K. J. Patterson (Eds.), *Palgrave handbook of econometrics: Vol. 1, Econometric theory* (pp. 278–352). Basingstoke: Palgrave Macmillan.

Pierce, D. A. (1978). Seasonal adjustment when both deterministic and stochastic seasonality are present. In A. Zellner (Ed.), *Seasonal analysis of economic time series* (pp. 242–269). Washington, DC: US Department of Commerce.

Rhodes, R. I., Shaffrey, L. C., & Gray, S. L. (2015). Can reanalyses represent extreme precipitation over England and Wales? *Quarterly Journal of the Royal Meteorological Society, 141,* 1114–1120. http://dx.doi.org/10.1002/qj.2015.141.issue-689

Cogent Geoscience (ISSN: 2331-2041) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com