

European Neolithic Societies Showed Early Warning Signals of Population Collapse

Sean S. Downey, W. Randall Haas Jr., and Stephen Shennan

8 June, 2016

Abstract

Ecosystems on the verge of major reorganization-regime shift may exhibit declining resilience, which can be detected using a collection of generic statistical tests known as early warning signals (EWS). This study explores whether EWS anticipated human population collapse during the European Neolithic. It analyzes recent reconstructions of European Neolithic (8-4 kya) population trends that reveal regime shifts from a period of rapid growth following the introduction of agriculture to a period of instability and collapse. We find statistical support for EWS in advance of population collapse. Seven of nine regional datasets exhibit increasing auto-correlation and variance leading up to collapse, suggesting that these societies began to recover from perturbation more slowly as resilience declined. We derive EWS statistics from a prehistoric population proxy based on summed archaeological radiocarbon date probability densities. We use simulation to validate our methods and show that sampling biases, atmospheric effects, radiocarbon calibration error, and taphonomic processes are unlikely to explain the observed EWS patterns. The implications of these results for understanding the dynamics of Neolithic ecosystems are discussed, and we present a general framework for analyzing societal regime shifts using EWS at large spatial and temporal scales. We suggest that our findings are consistent with an adaptive cycling model that highlights both the vulnerability and resilience of early European populations. We close by discussing the implications of the detection of EWS in human systems for archaeology and sustainability science.

Significance Statement: This study explores whether archaeologically detectable declines in resilience precede the onset of large-scale human population collapses. Our case study is the European Neolithic—a period that began approximately 9,000 years ago when the introduction of agricultural technologies initiated phases of rapid population growth that were in many cases followed by demographic instability and dramatic collapse. Our study finds evidence that statistical signatures of decreasing resilience can be detected long before population decline begins. To our knowledge, this study is the first to find early warning signals of demographic regime shift among human populations. The results suggest that demographic information can potentially be used to monitor social and ecological vulnerability in human societies at large spatial and temporal scales.

1 Introduction

A 2012 Special Issue in PNAS debates how analysis of historical collapse in ancient societies can contribute to sustainability science [1]. Key themes include accounting for complexity and multi-causality in instances of collapse, modeling and predicting both short and long-term environmental change, and the importance of historical and archaeological case studies. While significant progress has been made in measuring ecosystem resilience and predicting collapse [e.g., 2], quantifying the resilience of human societies presents a major challenge for social science research [3, 4]. Further, the use of archaeological data and EWS methods to predict known periods of collapse in ancient human societies (i.e., *retrodiction* [see 5]) remains largely unexplored. *Resilience* as we use the concept here is defined as the ability of a system to absorb change and recover from disturbance while maintaining relationships between populations or state variables [6]. Recent developments in ecology point to a promising new direction that follows from the observation that ecosystem resilience tends to decrease in advance of *regime shifts*—major transitions among qualitatively distinct ecosystem states [7]. Theoretical and empirical studies of non-human systems reveal that decreasing resilience is detectable via time-series statistics termed *early warning signals* (EWS) [8]. Although regime shifts are well documented in human-dominated ecosystems [9, 10, 11, 12, 13], the degree to which EWS anticipate them remains largely unexamined due to data limitations at requisite spatial and temporal scales. However, recent advances in the integration of large-scale archaeological data [14, 15, 16] are narrowing the gap between theory and data. This study presents what we believe to be the first statistical evidence for EWS of regime shifts in human population dynamics.

Our case studies include 2,378 archaeological sites from 9 regions of Neolithic Europe, ca. 8-4 kya. Previous research has observed evidence of major demographic regime shifts in the form of large-scale boom-bust dynamics among many of these Neolithic cases [17]. Estimated population declines range from 20 to 60 percent in as little as a century. The population proxies that revealed these boom-bust dynamics are based on the temporal frequencies of radiocarbon-dated archaeological sites, which are represented as summed probability densities (SPDs). This site-based population proxy assumes that the temporal frequencies of occupied human settlements in a given region index relative human population size. Use of SPD-based approaches to inferring population change has been debated in the literature [18, 19, 20, 21, 22, 23, 24, 25, 26]. Critics have raised concerns about

confounding factors including atmospheric effects, sampling biases, taphonomic processes, or calibration error; however, the methods used here and elsewhere [e.g., 20] attempt to control for these sources of error by (a) correcting systematic biases in the data and (b) comparing the corrected empirical patterns to null SPD models that simulate exogenous processes. Thus, current SPD methods reflect a significantly more conservative version of the approach that Rick [27] originally proposed.

We analyze archaeological SPDs for two classes of EWS: *critical slowing down* (CSD) and *flickering*. CSD describes a general increase in the time it takes a system to recover from external shocks such as population loss due to disease, warfare, or crop failure. Flickering describes increasing directional bias in a system’s response rate to such perturbations, such as a society stuck in a *socio-ecological trap* where strong reinforcing behavior and a lack of innovation prevents adaptation [28, 29, 30]. Here, flickering would suggest increasing recovery time from population decline events relative to growth events prior to major collapse. Because a number of non-human biological populations exhibit critical slowing down and flickering in advance of regime shifts [31, 32, 33, 34, 35, 36, 37, 38], we consider the possibility that these EWS indicators could also be detected in long-term human population dynamics as viewed through archaeological proxies.

2 Results

2.1 CSD detection with a simulated demographic regime shift and an archaeological deposition process model

Figure 1 illustrates how CSD can be detected in a model system that transitions from a high growth regime to low growth. In the model, slowly decreasing human population growth rates and increasing variance precedes the transition to the low growth regime. These simulated growth rates are used to simulate an archaeological SPD including deposition processes, taphonomic decay, atmospheric effects, radiocarbon calibration, and isotopic counting error processes (see Materials and Methods). The SPD is then evaluated for EWS using the same statistical procedures we use in the empirical analysis. Declining resilience is indicated in the bottom panel by the increasing EWS indicator values up to the point of collapse. The EWS indicator increase is statistically significant using the individual tau observations (τ_s), and with a robust statistical test (τ_n). Empirically, any of a number of drivers such as changing climate, declining environmental productivity, disease, warfare or some combination thereof could produce these results.

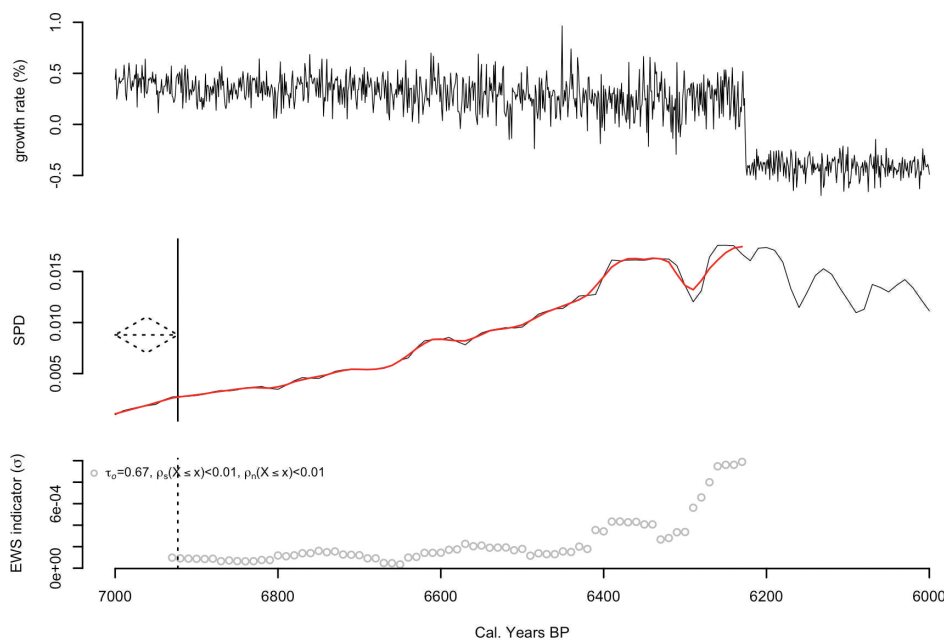


Figure 1: A regime shift model of population growth rate variability, radiocarbon date calibration, and EWS that demonstrates CSD in growth rates can be recovered from simulated SPDs.

2.2 Results for all European Neolithic regions

Qualitative EWS trends are evident in many cases (see Table S1 and SI Dataset), leading us to perform quantitative tests (see SI Text for an illustrative example). Table 1 presents the results of $\tau_{[o]}$ for the three EWS metrics, including their associated statistical strengths ($p_{[s]}$). All but two of the 27 tests produced statistically significant results ($p_{[s]} < 0.1$), indicating that statistical effects are unlikely to account for the observed EWS. However, these standard probability tests for $\tau_{[o]}$ do not distinguish between EWS produced by Neolithic demographic processes and those produced by the confounding factors, so we next compare the empirical tau values to a simulated null model. This comparison shows that in seven of nine regions, both AR[1] and σ exhibit increases that are significantly different from those produced by the null model ($p_{[n]} < 0.1$). Conversely, 1γ produces only one result out of nine that is significantly different from the Tau values produced by the null model. We conclude that confounding factors (a) are unlikely to account for the majority of CSD

patterns in the empirical data but (b) can account for the majority of skewness patterns in the data.

Table 1: EWS analysis results of summed probability distributions of eleven European regions.

region	AR[1]			σ			1γ		
	$\tau_{[o]}$	$p_{[s]}$	$p_{[n]}$	$\tau_{[o]}$	$p_{[s]}$	$p_{[n]}$	$\tau_{[o]}$	$p_{[s]}$	$p_{[n]}$
Southern Germany	0.54	0.00***	0.04**	0.89	0.00***	0.32	-0.78	0.00***	0.04**
Eastern Switzerland	0.59	0.00***	0.05**	0.13	0.11	0.47	-0.47	0.00***	0.14
England and Wales (w/o Wessex & Sussex)	0.49	0.00***	0.04**	0.96	0.00***	0.01***	-0.51	0.00***	0.14
Ireland	0.70	0.00***	0.01***	0.96	0.00***	0.03**	-0.37	0.00***	0.27
Paris Basin	0.65	0.00***	0.01***	0.97	0.00***	0.09*	-0.30	0.00***	0.18
Rhone-Languedoc	0.65	0.00***	0.09*	0.95	0.00***	0.04**	-0.79	0.00***	0.27
Scotland	0.49	0.00***	0.03**	0.95	0.00***	0.02**	-0.09	0.31	0.42
Southern England (Wessex & Sussex)	0.42	0.04**	0.17	0.68	0.00***	0.06*	-0.44	0.00***	0.13
Western France	0.48	0.02**	0.14	0.62	0.00***	0.04**	-0.36	0.02**	0.37

Significance levels: *** p 0.01; ** p 0.05; * p 0.10

Table 2 considers whether the multiple statistically significant results could be obtained by chance given $p_{[n]}$ values that are uniformly distributed between 0 and 1. For each set of nine $p_{[n]}$ outcomes for AR[1], σ , and 1γ , Fisher's Exact tests for AR[1] (7:2) and σ (7:2) show that the empirical results are unlikely to be explained by sampling effects ($p < 0.01$), while 1γ (1:8) can be explained by sampling effects ($p = 1$). In sum, the data from all regions reveal evidence of CSD but flickering does not appear to be a useful indicator of population collapse.

Table 2: Fisher's exact tests for multiple analyses. Pass/fail at $p < 0.1$.

Tau ($\tau_{[o]}$)	EWS	pass	fail	CI low	CI high	p
sample ($p_{[s]}$)	AR[1]	9	0	6.67	Inf	0.00***
	σ	8	1	3.54	4053.00	0.00***
	1γ	8	1	3.54	4053.00	0.00***
null model ($p_{[n]}$)	AR[1]	7	2	2.20	1854.15	0.00***
	σ	7	2	2.20	1854.15	0.00***
	1γ	1	8	0.02	117.49	1.00

[†]Expected ratio used in Fisher test is 1:8, pass:fail.

Significance levels: *** p 0.01; ** p 0.05; * p 0.10

3 Discussion

It is perhaps unsurprising that societies on the verge of collapse may exhibit warnings signs; yet it is difficult to demonstrate such phenomena empirically. Our results support the hypothesis that CSD was present in Neolithic Europe demographics and detectable in archaeological SPD curves, and that the EWS are not artifacts of sampling or confounding effects. This surprising finding encourages us to explore systemic relationships between human paleodemography and CSD, and to consider the implications for human ecosystem monitoring.

3.1 Regime shifts in human societies

Human population dynamics are known to exhibit multi-dimensional and non-linear processes, therefore regime shifts and EWS should also be expected. Equilibrium, multiple stable population points, and chaotic regimes are all known to emerge from even the simplest demographic models [39]. When density dependent population feedbacks, or Allee effects, interact with logistic growth and environmental perturbations, critical transitions may ensue [40, 6]. Allee effects and logistic growth processes have been observed or suspected in many biological populations including yeast, plants, shellfish, and humans [41, 42, 43]. Moreover, human systems involve cross-generational effects of past environments on the population levels of later generations [44, 45]. For example, dramatic environmental change, warfare, disease, or complex interactions among these mechanisms may lead to population collapse. Regime shifts and CSD should therefore be expected in at least some human population dynamics.

3.2 A framework for interpreting EWS among human populations at large spatial and temporal scales

Following a recent theoretical synthesis on resilience in socio-ecological systems [46], we consider the following three generic mechanisms that may offer insights into the general mechanisms of regime shifts in human systems over large spatial and temporal scales: (MI) a slowly changing driver to tipping point, (MII) interaction of fast and slow cycles, and (MIII) large but infrequent changes in external drivers.

M1 entails an external driver that slowly forces a system across a tipping point to a qualitatively different state. Such *critical transitions* entail feedback loops that create the context for non-linear responses to linear changes in the external driver [37]. As noted above, Allee effects can trigger critical transitions, and Janssen et al. [47, 37] propose that Neolithic population collapse in the US Southwest and among human societies more generally can be understood as critical transitions [see also 48]. As resource availability steadily declines, *sunk-cost effects* can generate adaptive feedback loops that artificially lock humans into maladaptive strategies such as remaining in established settlements until the exogenous driver eventually forces the system across a tipping point into an alternative adaptive regime. Continuing on such unsustainable courses in the face of steady resource decline ultimately leads to catastrophic failure. More recently, Lade et al. [49] show that common-pool resource management systems can create a context for critical transitions in human societies when external drivers slowly force systems to alternative states.

In M1 systems, the timing of regime shifts can be predicted at the point when variance and autocorrelation reach infinity, known technically as a bifurcation point (or tipping point). However, reaching a bifurcation point would only happen under stringent conditions, and stochasticity in real-world systems will tend to trigger regime shifts prior to the theoretical transition; for example, in early agricultural societies (Figure S3). Regardless, CSD should precede the regime shift [46]. Of the three mechanisms discussed here, only M1 causes true critical transitions with bifurcations. As a result, reducing or reversing the driver variable after a bifurcation will not return the system to the previous regime without resetting other system parameters (i.e., *hysteresis*). The other two mechanisms also cause regime shifts, but because recovery is possible and the changes are not permanent, these are not considered critical transitions [46].

MII describes interaction between slow and fast cycles that can cause dramatic regime shifts without bifurcation. For example, annual cultivation cycles in agricultural systems that involve forest clearing (i.e., swidden or slash-and-burn) are highly constrained by soil fertility and biomass recovery dynamics occurring over decades. Repeated and intensive forest cultivation can cause local resources to become depleted and ultimately trigger the abandonment of settlements; but eventually, forests and soils can recover and abandoned settlements may be reoccupied. Many types of environmental and human cycling exist that could lead to interactions and regime shifts, including environmental over-exploitation and recovery [50, 45], biogeochemical cycling [51], land surface change [2], climate cycling [52], epidemiology [53], human demography [13], and episodes of human violence [54, 55]. These dynamics are predicted to exhibit the statistical signatures of CSD [46].

In MIII systems a large but infrequent change in external conditions forces a system into another state. For example, catastrophic population losses could result from unusual natural disasters [56, 57], the emergence of genetically novel disease transmission vectors (e.g., airborne transmission of *Yersinia pestis* [58] and the Black Death [59]), social conflict at novel scales of severity (e.g., WW I), or extreme and rare climatic events such as volcanic eruptions (e.g., Pompeii). However, the demographic system itself is not trapped at low levels and may eventually recover. EWS is not expected in such cases [46].

3.3 Implications for understanding the causes of collapse during the European Neolithic

The EWS analysis of Neolithic Europe population dynamics tends to exhibit EWS in AR[1] and σ , suggesting that M1 and MII are plausible and that MIII is unsupported. However, determining which of the two potential regime shift mechanisms is the more plausible explanation for a pattern of collapse occurring independently at different times and in different regions throughout Europe cannot be determined from the EWS analysis alone. Instead, we use *wavelet analysis* to test the expectation of an MII slow-fast cycling mechanism (SI Materials and Methods). For example, the Wessex-Sussex region of England exhibits statistically significant cycling patterns in the SPDs (Figure 2). In this case the standard Gaussian white-noise filter identifies significant cycles at high frequencies that are probably associated with the radiocarbon date calibration process. However, when we use a noise filter based on the SPD null model distribution, the main statistically significant cycles remaining are around 5.5 kya with a frequency of between 400-1000 years. The pattern holds for each of the nine regions (see SI Dataset), and we find it to be in line with predictions from human demographic modeling [13]. A possible explanation is that interactions between fast human demographic cycles and slower ecosystem recovery cycles may explain the observed pattern of demographic collapse in Neolithic Europe. Similar cycling patterns have been observed among prehistoric agrarian populations in the U.S. Southwest [54, 55] and the hypothesis is consistent with recent paleoenvironmental research that finds correlations between deforestation and human population growth during the European Neolithic [60, 61, 62]. Using the framework introduced in the previous section, this deforestation hypothesis would suggest that Mechanism II cycle interactions between rapidly increasing human population levels and environmental dynamics during the early Neolithic may have contributed to the observed demographic collapses.

Importantly, fast-slow cycle interactions in human societies are more consistent with adaptive cycles [3, 45] than Lotka-Volterra interactions. In contrast to the latter, adaptive cycling emphasizes collapse and reorganization that may or may not reproduce previous ecological states. Instead, each collapse was generally followed by a level of recovery that sometimes exceeded pre-collapse levels and was often accompanied by distinct socioeconomic structure (e.g., Bronze Age systems). Recovery may be understood as the outcome of adaptive reorganization on fundamentally altered cultural and natural landscapes. Such reorganization is generally expected among human populations with

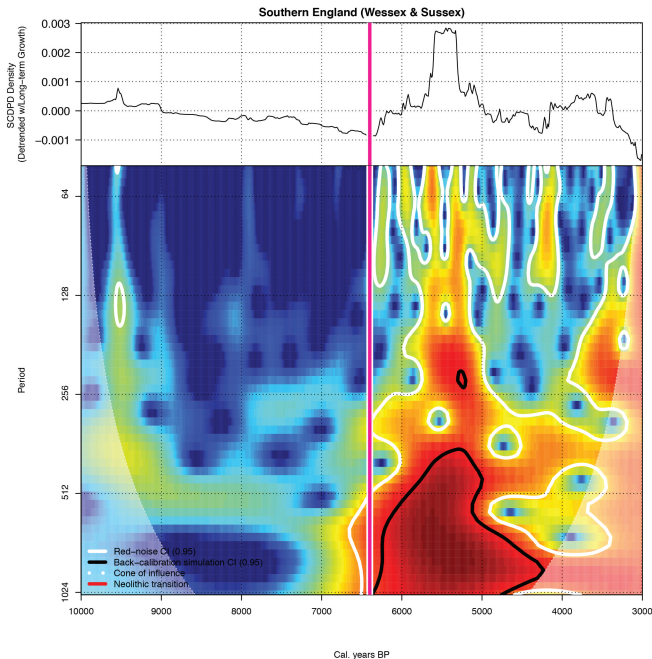


Figure 2: Wavelet analysis of Wessex Sussex showing Neolithic transition date (red line), periods of significant cycling and frequency using a gaussian noise (white highlight) and the SPD null model (black highlight).

cumulative culture. In some cases, endogenous expansion-growth dynamics may have temporarily increased societal resilience to changing rates of climate variability (e.g., during Linear Pottery Culture, see [63]). Therefore, on one hand, some societies tended to experience quantifiable loss of resilience that led to dramatic declines in population levels; on the other hand, other societies recovered to generate larger and presumably more resilient populations. Where recovery did not occur, other societies with new adaptations moved in, as ancient DNA studies are beginning to show [e.g., 64]. As such, our results highlight both the vulnerability and the resilience of early European societies.

4 Conclusion: prospects and contributions to archaeology and sustainability science

During the Neolithic revolution new agricultural technologies initiated rapid demographic growth, followed by periods of devastating societal instability that we are only now beginning to understand. It remains unclear whether modern technological innovation can continue to outpace demand, and it is important for sustainability scientists to consider the possibility that generic mechanisms can contribute to demographic collapse in human societies as well as to develop ways to detect declining resilience. Here we present evidence that early warning signals preceded large-scale population collapse in the European Neolithic Period. To encourage further study of human ecodynamics, we include a framework for interpreting societal regime shift at large spatial and temporal scales that links three generic mechanisms that are known for causing social and ecological regime-shifts, with social processes such as growth and collapse, climate change, resource degradation, disease, and warfare. We suggest that complex interactions among social and natural factors, and emergent patterns such as Allee and sunk-cost effects, may slow the recovery of human societies during periods of decreased resilience. Further, distinct statistical signatures of declining resilience due to these these processes are detectable despite the complex depositional processes that confound archaeological proxy data. We suggest that the detection of EWS in human settlement patterns is a general finding that points to a need for EWS analyses of other types of archaeological data and other historical data sets. We believe our framework can provide a way to analyze complex dynamics in human ecosystems, and perhaps ultimately to monitor and prevent catastrophic consequences of societal regime shifts.

5 Materials & Methods

5.1 Archaeological radiocarbon database

The complete dataset includes archaeological radiocarbon dates comprising 2,759 sites for Mesolithic and Neolithic Europe, ca. 10.0-3.5 kya [65]. Nine regions were selected for the EWS analysis because they provide the clearest qualitative and quantitative shifts from high to low growth regimes (Table S2).

5.2 Statistical methods and SPD modeling

To minimize the effects of sampling bias, radiocarbon calibration effects [66] and taphonomic bias [21], we use the Bchron R package [67], IntCal13 [68], and Monte Carlo simulation to generate a

corrected SPD curve. The EWS analysis involves isolating subsets from each time series from the Neolithic transition to the point of collapse, detrending the resulting time series, and calculating three EWS statistics including autocorrelation, variance, and skewness following [69]. EWS patterns are assessed qualitatively (Table S1), and statistically using the Kendall’s Tau rank correlation test, using both an optimized Tau value ($\tau_{[o]}$) and the complete time series. The procedures for generating a null model follow [17] and the rigorous statistical test for examining the effects of confounding archaeological factors on the EWS analysis involves computing cumulative density functions from the null model, and calculating the probabilities of each observed EWS statistic using single-tailed tests. Fisher’s Exact tests are used to determine the probability of the significant EWS statistics τ_o and $\tau_{[o]}$ for the optimized time series (Table 2) and for the full time series (Table S4). A sensitivity analysis evaluates the effect of sliding window size and the length of the time series on the significance of the EWS statistics [70] (SI Dataset). These EWS methods, the simulation shown in Figure 1, and the wavelet analysis ([71, 72]) are described in further detail the SI Materials and Methods. All analyses are performed using R [73]. Code is available upon request.

6 Acknowledgements

The research was funded by the Anthropology Department, the College of Behavioral and Social Sciences at the University of Maryland, and by European Research Council Advanced Grant number 249390 to Shennan.

References

- [1] Butzer, K. W & Endfield, G. H. (2012) Critical perspectives on historical collapse. *Proceedings of the National Academy of Sciences* **109**, 3628–3631.
- [2] Streeter, R & Dugmore, A. J. (2013) Anticipating land surface change. *Proceedings of the National Academy of Sciences* **110**, 5779–5784.
- [3] Allen, C. R, Angeler, D. G, Garmestani, A. S, Gunderson, L. H, & Holling, C. S. (2014) Panarchy: theory and application. *Ecosystems* **17**, 578–589.
- [4] Redman, C. L. (2005) Resilience theory in archaeology. *American Anthropologist* **107**, 70–77.
- [5] Noble, I. R. (1996) *Global change and terrestrial ecosystems*, International Geosphere-Biosphere Programme Book Series, ed. Walker, Brian; Steffan, B. No. 2, pp. 173–183.
- [6] Holling, C. S. (1973) Resilience and stability of ecological systems. *Annual Review of Ecology & Systematics* **4**, 1–23.
- [7] Scheffer, M, Carpntenter, S, Foley, J. A, Folke, C, & Walker, B. (2001) Catastrophic shifts in ecosystems. *Nature* **413**, 591.
- [8] Scheffer, M, Carpenter, S. R, Dakos, V, & Nes, E. v. (2015) Generic indicators of ecological resilience: inferring the chance of a critical transition. *Annual Review of Ecology, Evolution, and Systematics* **46**, 145–167.
- [9] Tainter, J. A. (1988) *The Collapse of Complex Societies*. (Cambridge University Press, Cambridge, Cambridgeshire; New York).
- [10] Diamond, J. M. (2005) *Collapse: How Societies Choose to Fail or Succeed*. (Viking, New York).
- [11] May, R. M, Levin, S. A, & Sugihara, G. (2008) Complex systems: ecology for bankers. *Nature* **451**, 893–895.
- [12] Bocquet-Appel, J.-P. (2011) When the world’s population took off: the springboard of the Neolithic demographic transition springboard of the Neolithic demographic transition. *Science* **333**, 560–561.
- [13] Turchin, P. (2009) Long-term population cycles in human societies. *Ann N Y Acad Sci* **1162**, 1–17.
- [14] Kintigh, K. W, Altschul, J. H, Beaudry, M. C, Drennan, R. D, Kinzig, A. P, Kohler, T. A, Limp, W. F, Maschner, H. D. G, Michener, W. K, Pauketat, T. R, Peregrine, P, Sabloff, J. A, Wilkinson, T. J, Wright, H. T, & Zeder, M. A. (2014) Grand challenges for archaeology. *Proceedings of the National Academy of Sciences* **111**, 879–880.
- [15] van der Leeuw, S, Costanza, R, Aulenbach, S, Brewer, S, Burek, M, Cornell, S, Crumley, C, Dearing, J. A, Downy, C, Graumlich, L. J, Heckbert, S, Hegmon, M, Hibbard, K, Jackson, S. T, Kubiszewski, I, Sinclair, P, Sörlin, S, & Steffen, W. (2011) Toward an integrated history to guide the future. *Ecology and Society* **16**.
- [16] Dearing, J. A, Braimoh, A. K, Reenberg, A, Turner, B. L, & van der Leeuw, S. (2010) Complex land systems: the need for long time perspectives to assess their future. *Ecology and Society* **15**.

- [17] Shennan, S, Downey, S. S, Timpson, A, Edinborough, K, Colledge, S, Kerig, T, Manning, K, & Thomas, M. G. (2013) Regional population collapse followed initial agriculture booms in mid-Holocene Europe. *Nature Communications* **4**.
- [18] Williams, A. N. (2012) The use of summed radiocarbon probability distributions in archaeology: a review of methods. *Journal of Archaeological Science* **39**, 578–589.
- [19] Collard, M, Edinborough, K, Shennan, S, & Thomas, M. G. (2010) Radiocarbon evidence indicates that migrants introduced farming to Britain. *Journal of Archaeological Science* **37**, 866–870.
- [20] Timpson, A, Colledge, S, Crema, E, Edinborough, K, Kerig, T, Manning, K, Thomas, M. G, & Shennan, S. (2014) Reconstructing regional population fluctuations in the European Neolithic using radiocarbon dates: a new case-study using an improved method. *Journal of Archaeological Science*.
- [21] Surovell, T. A, Byrd Finley, J, Smith, G. M, Brantingham, P. J, & Kelly, R. (2009) Correcting temporal frequency distributions for taphonomic bias. *Journal of Archaeological Science* **36**, 1715–1724.
- [22] Ballenger, J. A. M & Mabry, J. B. (2011) Temporal frequency distributions of alluvium in the American Southwest: taphonomic, paleohydraulic, and demographic implications. *Journal of Archaeological Science* **38**, 1314–1325.
- [23] Torfing, T. (2015) Neolithic population and summed probability distribution of 14C-dates. *Journal of Archaeological Science* **63**, 193–198.
- [24] Timpson, A, Manning, K, & Shennan, S. (2015) Inferential mistakes in population proxies: A response to Torfing’s “Neolithic population and summed probability distribution of 14C-dates”. *Journal of Archaeological Science* **63**, 199–202.
- [25] Zahid, H. J, Robinson, E, & Kelly, R. L. (2016) Agriculture, population growth, and statistical analysis of the radiocarbon record. *Proceedings of the National Academy of Sciences* **113**, 931–935.
- [26] Downey, S. S, Bocaege, E, Kerig, T, Edinborough, K, & Shennan, S. (2014) The Neolithic demographic transition in Europe: Correlation with juvenility index supports interpretation of the summed calibrated radiocarbon date probability distribution (SCDPD) as a valid demographic proxy. *PLoS ONE* **9**, e105730.
- [27] Rick, J. W. (1987) Dates as data: an examination of the Peruvian Preceramic radiocarbon record. *American Antiquity* **52**, 55–73.
- [28] Boonstra, W. J & de Boer, F. W. (2014) The historical dynamics of social–ecological traps. *Ambio* **43**, 260–274.
- [29] Carpenter, S. R & Brock, W. A. (2008) Adaptive capacity and traps. *Ecology and society* **13**, 40.
- [30] Tschakert, P & Shaffer, L. J. (2014) in *Social-Ecological Systems in Transition*. (Springer), pp. 139–156.
- [31] Carstensen, J, Telford, R. J, & Birks, H. J. B. (2013) Diatom flickering prior to regime shift. *Nature* **498**, E11–E12.
- [32] Dakos, V, van Nes, E. H, & Scheffer, M. (2013) Flickering as an early warning signal. *Theoretical Ecology* **6**, 309–317.
- [33] Wang, R, Dearing, J. A, Langdon, P. G, Zhang, E, Yang, X, Dakos, V, & Scheffer, M. (2012) Flickering gives early warning signals of a critical transition to a eutrophic lake state. *Nature* **492**, 419–422.
- [34] Bailey, R. M. (2011) Spatial and temporal signatures of fragility and threshold proximity in modelled semi-arid vegetation. *Proceedings of the Royal Society B: Biological Sciences* **278**, 1064–1071.
- [35] Drake, J. M & Griffen, B. D. (2010) Early warning signals of extinction in deteriorating environments. *Nature* **467**, 456–459.
- [36] Scheffer, M, Bascompte, J, Brock, W. A, Brovkin, V, Carpenter, S. R, Dakos, V, Held, H, Nes, E. H. v, Rietkerk, M, & Sugihara, G. (2009) Early-warning signals for critical transitions. *Nature* **461**, 53–59.
- [37] Scheffer, M. (2009) *Critical Transitions in Nature and Society*. (Princeton University Press, Princeton, N.J.).
- [38] Hefley, T. J, Tyre, A. J, & Blankenship, E. E. (2013) Statistical indicators and state–space population models predict extinction in a population of bobwhite quail. *Theoretical Ecology* **6**, 319–331.

- [39] May, R. M. (1976) Simple mathematical models with very complicated dynamics. *Nature* **261**, 459–467.
- [40] Dai, L, Vorselen, D, Korolev, K. S, & Gore, J. (2012) Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science* **336**, 1175–1177.
- [41] Taylor, C. M & Hastings, A. (2005) Allee effects in biological invasions. *Ecology Letters* **8**, 895–908.
- [42] Steele, J. (2010) Radiocarbon dates as data: quantitative strategies for estimating colonization front speeds and event densities. *Journal of Archaeological Science* **37**, 2017–2030.
- [43] Steele, J. (2009) Human dispersals: mathematical models and the archaeological record. *Human Biology* **81**, 121–140.
- [44] Benton, T. G, Plaistow, S. J, & Coulson, T. N. (2006) Complex population dynamics and complex causation: devils, details and demography. *Proceedings of the Royal Society B: Biological Sciences* **273**, 1173–1181.
- [45] Gunderson, L. H & Holling, C. S, eds. (2002) *Panarchy: Understanding Transformations in Human and Natural Systems*. (Island Press, Washington, DC).
- [46] Dakos, V, Carpenter, S. R, Nes, E. H. v, & Scheffer, M. (2015) Resilience indicators: prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **370**, 20130263.
- [47] Janssen, M. A, Kohler, T. A, & Scheffer, M. (2003) Sunk-cost effects and vulnerability to collapse in ancient societies. *Current Anthropology* **44**, 722–728.
- [48] Hegmon, M, Peeples, M. A, Kinzig, A. P, Kulow, S, Meegan, C. M, & Nelson, M. C. (2008) Social transformation and its human costs in the prehispanic U.S. Southwest. *American Anthropologist* **110**, 313–324.
- [49] Lade, S. J, Tavoni, A, Levin, S. A, & Schlüter, M. (2013) Regime shifts in a social-ecological system. *Theoretical Ecology* **6**, 359–372.
- [50] Lee, C. T & Tuljapurkar, S. (2008) Population and prehistory I: food-dependent population growth in constant environments. *Theoretical Population Biology* **73**, 473–482.
- [51] Folke, C, Carpenter, S, Walker, B, Scheffer, M, Elmqvist, T, Gunderson, L, & Holling, C. (2004) Regime shifts, resilience, and biodiversity in ecosystem management. *Annual Review of Ecology, Evolution, and Systematics* pp. 557–581.
- [52] Chavez, F. P, Ryan, J, Lluch-Cota, S. E, & Niquen, M. (2003) From anchovies to sardines and back: multidecadal change in the pacific ocean. *Science* **299**, 217–221.
- [53] Althouse, B. M & Hébert-Dufresne, L. (2014) Epidemic cycles driven by host behaviour. *Journal of The Royal Society Interface* **11**.
- [54] Turchin, P & Korotayev, A. (2006) Population dynamics and internal warfare: a reconsideration. *Social Evolution & History* **5**, 112–147.
- [55] Kohler, T. A, Cole, S, & Ciupe, S. (2009) in *Pattern and Process in Cultural Evolution*, ed. Shennan, S. (University of California Press, Berkeley, CA), pp. 277–295.
- [56] Firestone, R. B, West, A, Kennett, J. P, Becker, L, Bunch, T. E, Revay, Z. S, Schultz, P. H, Belgia, T, Kennett, D. J, Erlandson, J. M, Dickenson, O. J, Goodyear, A. C, Harris, R. S, Howard, G. A, Kloosterman, J. B, Lechler, P, Mayewski, P. A, Montgomery, J, Poreda, R, Darrah, T, Hee, S. S. Q, Smith, A. R, Stich, A, Topping, W, Wittke, J. H, & Wolbach, W. S. (2007) Evidence for an extraterrestrial impact 12,900 years ago that contributed to the megafaunal extinctions and the Younger Dryas cooling. *Proceedings of the National Academy of Sciences* **104**, 16016–16021.
- [57] Kuijt, I. (2001) Reconsidering the cause of cultural collapse in the Lillooet area of British Columbia, Canada: a geoarchaeological perspective. *American Antiquity* **66**, 692–703.
- [58] Rasmussen, S, Allentoft, M. E, Nielsen, K, Orlando, L, Sikora, M, Sjögren, K.-G, Pedersen, A. G, Schubert, M, Van Dam, A, Kapel, C. M. O, et al. (2015) Early divergent strains of *Yersinia pestis* in Eurasia 5,000 years ago. *Cell* **163**, 571–582.
- [59] Benedictow, O. J. (2012) *The Black Death 1346-1353: The Complete History*. (Boydell Press), 1st edition.
- [60] Fyfe, R. M, Woodbridge, J, & Roberts, N. (2015) From forest to farmland: pollen-inferred land cover change across Europe using the pseudobiomization approach. *Global Change Biology* **21**, 1197–1212.
- [61] Woodbridge, J, Fyfe, R. M, Roberts, N, Downey, S, Edinborough, K, & Shennan, S. (2014) The impact of the Neolithic agricultural transition in Britain: a comparison of pollen-based land-cover and archaeological 14C date-inferred population change. *Journal of Archaeological Science*.

- [62] Lechterbeck, J, Edinborough, K, Kerig, T, Fyfe, R, Roberts, N, & Shennan, S. (2014) Is Neolithic land use correlated with demography? An evaluation of pollen-derived land cover and radiocarbon-inferred demographic change from central Europe. *The Holocene* **24**, 1297–1307.
- [63] Gronenborn, D, Strien, H.-C, Dietrich, S, & Sirocko, F. (2014) 'Adaptive cycles' and climate fluctuations: a case study from Linear Pottery Culture in western Central Europe. *Journal of Archaeological Science* **51**, 73–83.
- [64] Haak, W, Lazaridis, I, Patterson, N, Rohland, N, Mallick, S, Llamas, B, Brandt, G, Nordenfelt, S, Harney, E, Stewardson, K, et al. (2015) Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211.
- [65] Manning, K, Colledge, S, Crema, E, Shennan, S, & Timpson, A. (2016) The cultural evolution of Neolithic Europe. EUROEVOL dataset 1: Sites, phases and radiocarbon data. *Journal of Open Archaeology Data* **5**.
- [66] Brown, W. A. (2015) Through a filter, darkly: population size estimation, systematic error, and random error in radiocarbon-supported demographic temporal frequency analysis. *Journal of Archaeological Science* **53**, 133–147.
- [67] Parnell, A. (2015) *Bchron: radiocarbon dating, age-depth modelling, relative sea-level rate estimation, and non-parametric phase modelling*. R package version 4.1.2.
- [68] Reimer, P. (2013) IntCal13 and Marine13 radiocarbon age calibration curves 0-50,000 years cal BP. *Radiocarbon* **55**, 1869–1887.
- [69] Dakos, V, Carpenter, S. R, Brock, W. A, Ellison, A. M, Guttal, V, Ives, A. R, Kéfi, S, Livina, V, Seekell, D. A, van Nes, E. H, & Scheffer, M. (2012) Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS ONE* **7**, e41010.
- [70] Best, D & Gipps, P. (1974) Algorithm as 71: the upper tail probabilities of Kendall's tau. *Applied Statistics* pp. 98–100.
- [71] Torrence, C & Compo, G. P. (1998) A practical guide to wavelet analysis. *Bulletin of the American Meteorological society* **79**, 61–78.
- [72] Gouhier, T. (2014) *biwavelet: Conduct univariate and bivariate wavelet analyses*. (Version 0.17.4).
- [73] R Core Team. (2015) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria).

Main Paper Table and Figure Captions

- Table 1. EWS analysis results of summed probability distributions of eleven European regions.
- Table 2. Fisher’s exact tests for multiple analyses. Pass/fail at $p < 0.1$.
- Figure 1. A regime shift model of population growth rate variability, radiocarbon date calibration, and EWS that demonstrates CSD in growth rates can be recovered from simulated SPDs.
- Figure 2. Wavelet analysis of Wessex Sussex showing Neolithic transition date (red line), periods of significant cycling and frequency using a gaussian noise (white highlight) and the SPD null model (black highlight).

Supporting Information Table and Figure Captions

- Figure S1. EWS analysis of Neolithic Paris Basin showing statistical support for CSD in advance of population collapse ~ 6.2 kya. Panel B shows the overall population trajectory used in EWS analysis. Panel A shows the region of the curve prior to collapse that is analyzed for EWS. Panel C, E, and G plot the three EWS metrics and associated $\tau_{[o]}$ values with probabilities. Panels D, F, and H compare the empirical EWS results to simulated values given a null model defined by confounding factors. Overall, these results identify CSD in advance of collapse in Neolithic Paris Basin but it does not provide evidence for flickering. See SI Text 1 for further explanation and interpretation of these results.
- Figure S2. EWS method flowchart.
- Figure S3. A conceptual fold bifurcation model for human population collapse. We consider that bifurcation acts on population growth rates. Reading the graphic from left to right, early agricultural societies have initially high growth rates. Land degradation slowly drives the growth rates down while agricultural surpluses and cultural mechanisms maintain the growth rates at artificially high levels within growth regime A. As the system approaches the tipping point, it becomes increasingly unstable and susceptible to shifts into the low growth-rate regime (regime B). The system exhibits hysteresis insofar as the recovery of agricultural productivity does not necessarily result in recovery of the higher regime A growth rates.
- Figure S4. An example of a corrected site density distribution analyzed for EWSs.
- Table S1. Qualitative classification of EWS trends in all regions. Here we subjectively classify each EWS trend as increasing (+), decreasing (-), or equivocal (.) and describe the overall pattern including the relationship to interesting events in the underlying SPD.
- Table S2. Summary of 14C samples and Neolithic events for the European cases examined in this study.
- Table S3. EWS analysis results of summed probability distributions of nine European regions using the complete (non-optimized) subset.
- Table S4. Fisher’s exact tests for multiple analyses using the complete (non-optimized) subset.

Supporting Information Dataset

- Dataset 1. EWS analysis for all regions.
- Dataset 2. Wavelet analysis for all regions.
- Dataset 3. EWS sensitivity analysis for all regions.

Supporting Information

SI Text

1 Illustrative example: early warning signals of collapse in Neolithic Paris Basin

Here we briefly step through a sample case in order to illustrate the procedure and interpretation of results. Figure S1 illustrates how EWS are detected in the Neolithic Paris Basin. Panel B shows the complete SPD curve including the region of the time series isolated for EWS analysis. The SPD curve serves as a temporal proxy for population change throughout the Paris Basin Neolithic Period. Following previous studies [17, 26, 20], we assign a regime-shift date, or “bust” point, at which time the system shifted from population growth to decline. For the Paris Basin example, the bust point was previously identified at 6225 cal. BP. We then qualitatively define a point-of-beginning for the Neolithic boom. In most cases, this is simply the point at which agriculture is thought to have entered the region. However, if the SPD exhibits growth prior to the currently accepted agricultural start date, we assign an earlier point-of-beginning. This decision is justifiable given that archaeological estimates of agricultural origins reflect minimum possible dates due to sample limitations. For the Paris Basin, we assign the beginning of the population boom to 7500 cal. BP, which precedes the currently accepted agricultural origins date by 300 years. Thus, the Paris Basin population boom isolated for EWS analysis ranges from 7500-6225 cal. BP.

Panel A shows the extracted time series, the time-series trend model, and a schematic representation of the EWS sliding window, all of which are used in the derivation of EWS metrics. Panels C, E, and G show the results of the EWS analysis for the three statistical indicators considered here, including lag-1 autocorrelation (AR[1]), standard deviation (σ), and skewness (1γ). AR[1] and σ are expected to increase (+), while skewness is expected to decrease (-) in advance of regime shift [37]. In the case of Neolithic Paris Basin, qualitative increases in AR[1] (+) and σ (+) are apparent after 6800 and 6300 cal. BP, respectively. In contrast, 1γ does not exhibit clear directionality, though some decrease may be present. Kendall’s Tau ($\tau_{[o]}$) statistics are used to quantify directionality and to assess statistical significance. The $\tau_{[o]}$ statistics support the qualitative observations at $p_{[s]} < 0.01$. That is, AR[1], σ , and 1γ exhibit significant departures from zero.

These statistics tentatively support the presence of CSD in the empirical data, but they cannot distinguish CSD due to systemic changes in human demography from CSD due to exogenous factors such as sampling error, atmospheric effects on radiocarbon, and taphonomic biases. Panels D, F, and H show the cumulative density distributions (*ECDF* ($\tau_{[e]}$)) for trend statistics generated from a null model simulating the effects of these confounding factors, and the probability of obtaining the empirical values given the null model (see Materials and Methods). The probability values for the $\tau_{[o]}$ trend statistics for AR[1] and σ reveal significant departures from the null model ($p_{[n]} < 0.1$), whereas 1γ does not. Therefore, the CSD trends are more likely to reflect demographic processes rather than exogenous factors. We conclude that CSD preceded the Neolithic Paris Basin collapse that occurred around 6225 BP. In contrast, the skewness results are inconclusive.

2 Theoretical considerations related to SPD-based EWS among human societies

While regime shifts and CSD may be expected, direct observation of such dynamics at requisite spatial and temporal scales is difficult. Archaeological data provide us with long-term, spatially extensive proxies for human demography [4]. The SPD method in this study does not directly measure human population levels but rather approximates the number of inhabited settlements in a given spatiotemporal unit, which is taken to index human population levels after applying appropriate corrections. Thus, SPDs directly measure of how settlement densities changed through time, but they do not directly measure population change. However, recent work has shown strong correlation between SPDs and a more direct demographic proxy based on skeletal measurements known as the juvenility index; this study also compared settlement plans and ethnographic censuses from contemporary farmers and foragers and found that differential settlement densities do not obscure demographic reconstruction using SPDs [26].

Because CSD is predicted to occur in human populations, a theoretical question that arises is how CSD should manifest in human settlement change, which may involve distinct dynamics. In order to explain how human settlement dynamics could come to reflect CSD in human populations, we hypothesize the relationships between archaeological settlement patterns in high and low growth regimes. In a high growth regime, such as occurred immediately after the introduction of agriculture in Neolithic Europe, there would have been a mix of young and old villages with new settlements forming from parent villages and rapid overall regional growth. In this regime, any abandoned settlements would have been rapidly repopulated because of high regional growth rates and high demand for resources. In a low growth regime, settlements would have been older on average with fewer new settlements appearing. Demographic variance and system shocks such as warfare, conflict, disease, or stochastic environmental effects would have had significant impacts, and failed settlements would have rarely repopulated because of the low overall demand for resources.

3 SI Materials and Methods

We present the complete methods in five parts. First, we describe the archaeological proxy data and the generation of a summed calibrated density distribution (SPD). Second, we describe the method for generating an SPD null model of the confounding factors: sampling error, taphonomic bias, and radiocarbon calibration effects. These procedures have been presented in previous publications [17, 26, 20], so we only briefly outline them here. Third, we describe the methods for calculating EWS metrics for the empirical data of each region, including the workflow for generating expected EWS results for each region given a null model. Figure S2 outlines the method and each step is cross-referenced in the text with curly braces (e.g., {1}). Fourth, we describe the simulation model used in Figure 1 in the main paper. Fifth, we describe the method for the wavelet analysis. All analyses are performed using R [73]. Code is available upon request.

3.1 Archaeological radiocarbon database

The publicly available dataset includes more than 13,658 archaeological radiocarbon dates comprising 2,759 sites for Mesolithic and Neolithic Europe, ca. 10.0-3.5kya [65]. Each site is assigned to a geographic region based on physiography, perceived spatial clustering, and prior knowledge of prehistoric cultural regions. Problematic dates including outliers and those with ambiguous stratigraphic associations are discarded. We select a subset of nine regions (Table S2) for the EWS analysis because they provide the clearest qualitative and quantitative shifts from high to low growth regimes—that is, population booms and busts [17], raising the possibility of EWS leading up to the bust events.

To minimize the effects of radiocarbon sampling bias within and among sites we use the following phase-based summing procedure [22, 24, 23]. Each date for a given site is assigned to a temporal site-phase and multiple phases are defined at a site when the minimum pairwise distances among all radiocarbon dates exceed 200^{14}C years. One ^{14}C assay is then randomly selected from each site phase and calibrated {1} using the Bchron R package [67] and IntCal13 atmospheric correction curve [68]. The resultant probability distributions are summed to unity. This procedure is repeated 20,000 times and each SPD is averaged {2}. The resultant SPDs are binned in 10-year periods. As a final step, we correct each probability density curve for radiocarbon calibration effects [66] and taphonomic bias [21] by estimating a null model using a Monte Carlo routine and subtracting it from the observed curve to generate a corrected SPD curve {9}.

3.2 SPD null model generation

In order to explore the possibility of observing EWS when they did not actually exist in the systemic context (Type I error), we construct a null model that simulates monotonic population growth, archaeological sampling, atmospheric effects on radiocarbon proxies, and taphonomic decay of radiocarbon proxies (Figure S4). We first fit a generalized linear model (GLM) to the empirical SPD {3}. This simultaneously models exponential growth of populations and decay of archaeological carbon [17]. Second, to model sampling, we randomly draw a probabilistic sample from a uniform distribution using the fitted GLM for probability weights, and a sample size equivalent to the empirical sample size {4}. Third, to model the effects of the radiocarbon calibration process, we “un-calibrate” each simulated date by passing it backwards through the IntCal13 radiocarbon calibration curve {5}. We then calibrate the uncalibrated values using them as ^{14}C means with standard error values sampled without replacement from the standard error values in the empirical sample {6} and sum the results {7}. This procedure is iterated 20,000 times and all calibrated probability distributions are summed to unity to produce a null SPD model {8}.

3.3 Calculating EWS metrics using SPD

The methods for calculating EWS statistics generally follow those outlined in [69]. For each of the nine population trajectories, we isolate a subset of the time series from the beginning of the population boom to the end, or the collapse date {10}. The beginning date takes into account both qualitative evidence of growth as observed in the SPDs and the estimated date of earliest local archaeological evidence for agriculture. The date of collapse in each region is determined qualitatively by inspecting the SPD plots. The SPD curves may include edge effects due to decreased sampling in later periods, however the collapse dates are earlier in time so this does not affect the EWS analysis results. Next, we de-trend the time-series to remove linear structure in the time series by fitting a LOESS model and calculating residuals. The LOESS fitting procedure uses local regression to predict SPD values along the x-axis ($\lambda = 2$; $\alpha = 0.15$), and residuals are calculated as the difference between the predicted and observed SPD values. We then calculate autocorrelation (AR[1]), variance (σ), and skewness (1λ) values for the detrended data using a sliding time-window {12}. The first window includes the first 25 percent of points in the time series and then iteratively slides forward one decade at a time until reaching the defined collapse date. The size of this sliding window follows the recommendation of [69] and considers the scale of human demographic processes. Additionally, we conduct a sensitivity analysis (see SI Dataset) for each region to evaluate the effect of sliding window size and the length of the time series on the significance of the EWS statistics [70]. Autocorrelation within the sliding window is one of the simplest measurements of critical slowing down. AR[1] measures the degree of correlation between pairs of temporally adjacent measurements of state variable values (i.e., SPD_t and SPD_{t+1}). As the sliding window approaches the collapse point, increasing AR[1] values indicate the SPD values are becoming increasingly similar, which is

an EWS for decreasing system resilience. This measurement is made here by fitting a lag-1 (AR[1]) ordinary least-squares regression model to each window of data and estimating the autoregression coefficient. More precisely, temporal autocorrelation is found as follows: $z_{t+1} = \alpha_1 z_t + \varepsilon_t$, where α is an autoregressive coefficient and ε_t is Gaussian white noise. Variance, reported here as a standard deviation (σ), quantifies the spread of values, which may indicate a system is approaching a regime shift. σ is found as follows: $\sigma = \frac{1}{n-1} \sum_{t=1}^n (z_t - \mu)^2$, where z is SPD density. Skewness (1γ) measures directional bias in the data as the standardized third moment around the mean of a

distribution, or $1\gamma = \frac{\frac{1}{n} \sum_{t=1}^n (z_t - \mu)^3}{\sqrt{\frac{1}{n} \sum_{t=1}^n (z_t - \mu)^2}}$ [69]. The shape of the distribution of system indicator variables

can change in systems that are undergoing CSD because as the system slows down, it will hover around a critical transition point. These values can also include ones that indicate alternative system states, a process sometimes called *flickering* [32]. This ambivalence in the system trajectory can yield increasingly skewed distributions with significant skewness statistics in either the positive or negative direction, depending on whether a systems is moving move towards larger or smaller values.

Once the summary statistics are calculated for each time increment, we qualitatively assess the data trends for EWS (Table S1); if a given system exhibits EWSs of a regime shift (collapse), theory predicts increasing autocorrelation and variance and decreasing skewness. Next, we assess these predictions for each region statistically using the Kendall’s Tau (τ_o) rank correlation test for the entire time-series subset. The τ statistic indicates the direction of the EWS statistic as either increasing ($0 < \tau < 1$) or decreasing ($-1 < \tau < 0$). We also calculate an *optimized Tau* value ($\tau_{[o]}$) by finding the maximum (or minimum) τ value from time series of varying duration that iteratively extend back from the collapse date beyond a minimum temporal window of 100 years before the collapse. This optimization routine is necessary for two reasons. First, we do not know a priori when resilience began to decline in these systems. The time series data under consideration may be much longer than the period of declining resilience. Second, the EWS:noise ratio is expected to be low and may preferentially obscure signals that are necessarily weaker earlier in time. Optimized tau values are calculated separately for each EWS statistic because we currently have no reason to expect different EWS metrics to change at proportionate rates and thus emerge through background noise at the same point in time. Rather, once a system begins to experience declining resilience, different EWS may change at different rates and thus emerge through background noise at different points in time. The EWS results for the Neolithic Paris Basin illustrate a situation in which (1) the quantitative method as applied to the full range of data is unable to capture qualitatively distinct AR[1] and σ EWS leading up to collapse, and (2) AR[1] and σ trends indicate differential sensitivities to the empirical SPD. Conversely, the optimization routine identifies these trends, and therefore, we report the $\tau_{[o]}$ in Figure 2 and Table 1 in the main paper. We follow standard approaches and report the probability values associated with τ_o {13} and $\tau_{[o]}$ {14} using single-tailed tests, p_s and $p_{[s]}$, respectively {13, 14}. For completeness, we report τ_o in Table S3 and both τ_o and $\tau_{[o]}$ in SI Dataset.

Recognizing the potential for spurious results given the complexity of these archaeological cases, we develop a rigorous statistical test to examine the effects of confounding factors. To do this, we analyze the SPD null model described in Section 3.2. We calculate EWS statistics for each of the 20,000 simulated density curves and combine them into distributions represented as $\{\tau_e\}$ and $\{\tau_{[e]}\}$ {15}. Cumulative density functions (CDFs) are computed {16, 17}, and the probabilities of τ_o and $\tau_{[o]}$ are determined for each EWS statistic, given a null model of exogenous processes {18, 19}. Thus, we define two “robust” probability values using single-sided tests, p_n and $p_{[n]}$ as $F(x) = Pr(X \geq x)$, where $F(x)$ is the CDF for $\{\tau_e\}$ and $\{\tau_{[e]}\}$, respectively.

Last, we examine the probability of obtaining the observed ratio of statistically significant:insignificant results. We apply a Fisher Exact test, which involves classifying the probability values for each EWS statistic of each region as either “passing” or “failing” given a one-tailed probability test $Pr(X \leq 0.1)$ to determine the likelihood that the significant EWS statistics τ_o and $\tau_{[o]}$ could have occurred by chance. Given the sample size of 9 cases, we use a 0.1 significance level in this classification. The number of observed cases out of 9 that pass or fail a given EWS test is given using an expected pass:fail ratio of 1:8. We report Fisher Exact test results for $p_{[s]}$ and $p_{[n]}$ in Table 2 in the main paper, and for p_s and p_n in Table S4.

3.4 A regime-shift model for human settlement densities at large spatial and temporal scales

The model shown in Figure 1 in the main paper uses a simple step function with a high value step decreasing from 0.4% to 0.2% and a low value step at -0.4% growth. The upper limit of 0.4% growth reflects the upper end of expected annual population growth rates in early agricultural populations [12]. To simulate increasing variance leading up to collapse, a randomization function is applied to the upper step values in which the mean is zero and the standard deviation increases from 0.1% to 0.2%. The variance in the lower step is set to be constant at 0.2%. The growth rate model is then used to simulate an SPD curve by assuming a starting population of 500 individuals and iteratively increasing the population size by the modeled growth rate value at a given time step. We model the effects of the radiocarbon date calibration process, including atmospheric effects and isotopic counting error, by back-calibrating a uniform sample of dates 7-5.5 kya (n=1000), adding

error sampled from a Poisson distribution ($\lambda = 50$), and calibrating these simulated dates with using the IntCal13 calibration curve [68]. In order to model the effects of archaeological sampling, a random sample of 1000 dates is drawn from the simulated population. The modified population curve is finally summed by the same binning procedure used on the empirical data. The result is a simulated SPD that results from a theoretical fold bifurcation in population growth rates, and we calculate the variance EWS metric on the SPD data assuming a collapse date at 6225 cal BP. We use a 77-year sliding window ($\sim 10\%$ of the 775 year duration, see [69]). Finally, we calculate τ_o and test its significance using p_s and p_n as described in section 5.3. We note that this model is developed to illustrate that EWS are detectable in archaeological SPDs despite the potentially confounding factors discussed in the main paper. The model explores a number of interactions, though it is not exhaustive. Other interactions, such as the relative strength of EWS that would be required for detecting EWS at varying points in the calibration curve or the possibility of CSD due to biological population growth may require additional investigation.

3.5 Wavelet analysis

A univariate continuous wavelet transform analysis [71] is run using the biwavelet package [72]. For each region we analyze a corrected and detrended SPD and plot two 95% significance contour intervals (SI Dataset). The CI plotted in white uses the built-in time-averaged filters. Plotting the second CI involves performing univariate wavelet transforms and calculating the 95% quantile on the distribution of power matrices from the simulated SPDs from the null model. The power matrix from the observed SPD is then divided by the quantiles from the SPD null simulations and the resulting 95% significance contour is plotted in black.

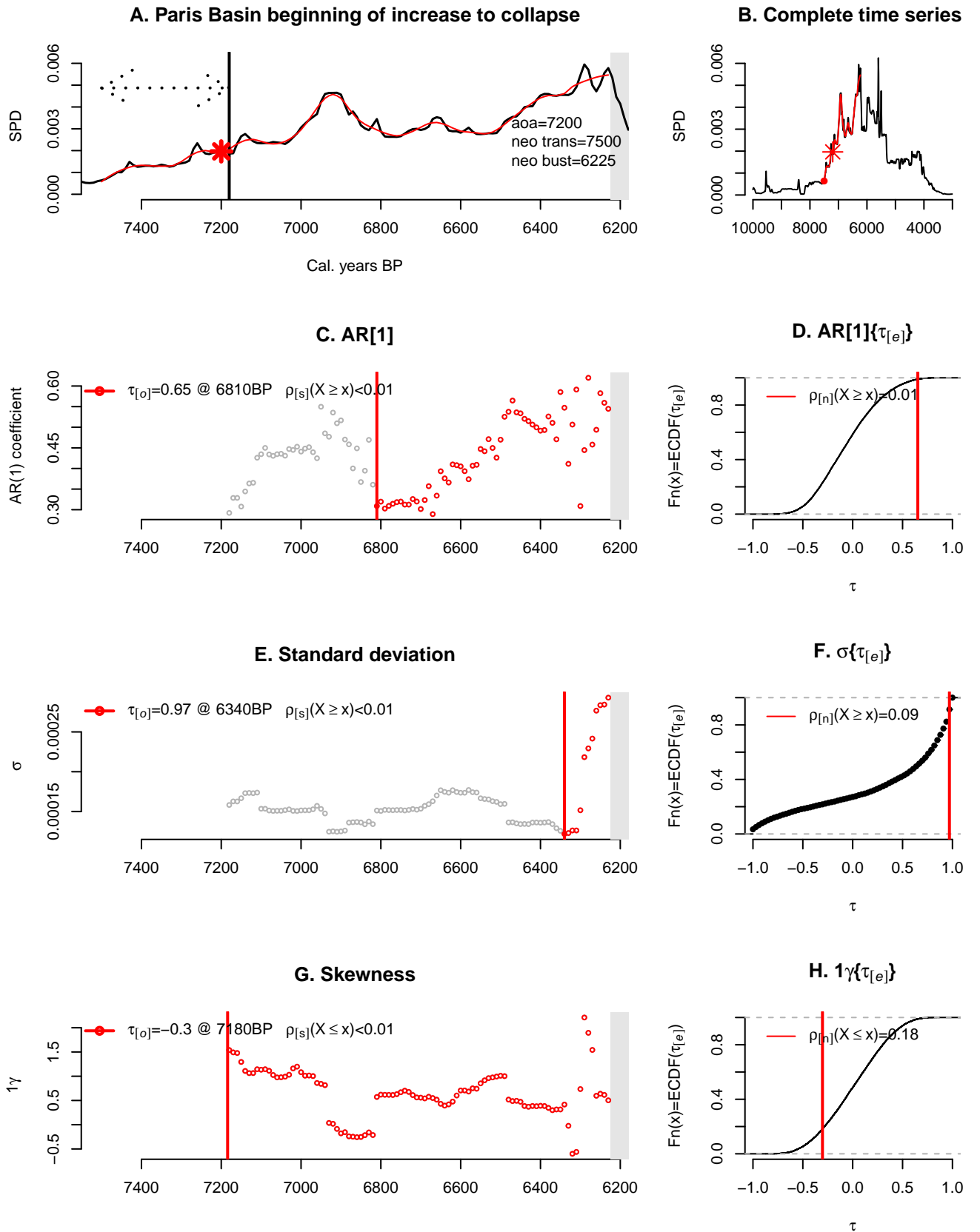


Figure S1: EWS analysis of Neolithic Paris Basin showing statistical support for CSD in advance of population collapse ~ 6.2 kya. Panel B shows the overall population trajectory used in EWS analysis. Panel A shows the region of the curve prior to collapse that is analyzed for EWS. Panel C, E, and G plot the three EWS metrics and associated $\tau_{[o]}$ values with probabilities. Panels D, F, and H compare the empirical EWS results to simulated values given a null model defined by confounding factors. Overall, these results identify CSD in advance of collapse in Neolithic Paris Basin but it does not provide evidence for flickering. See Section 1 for further explanation and interpretation of these results.

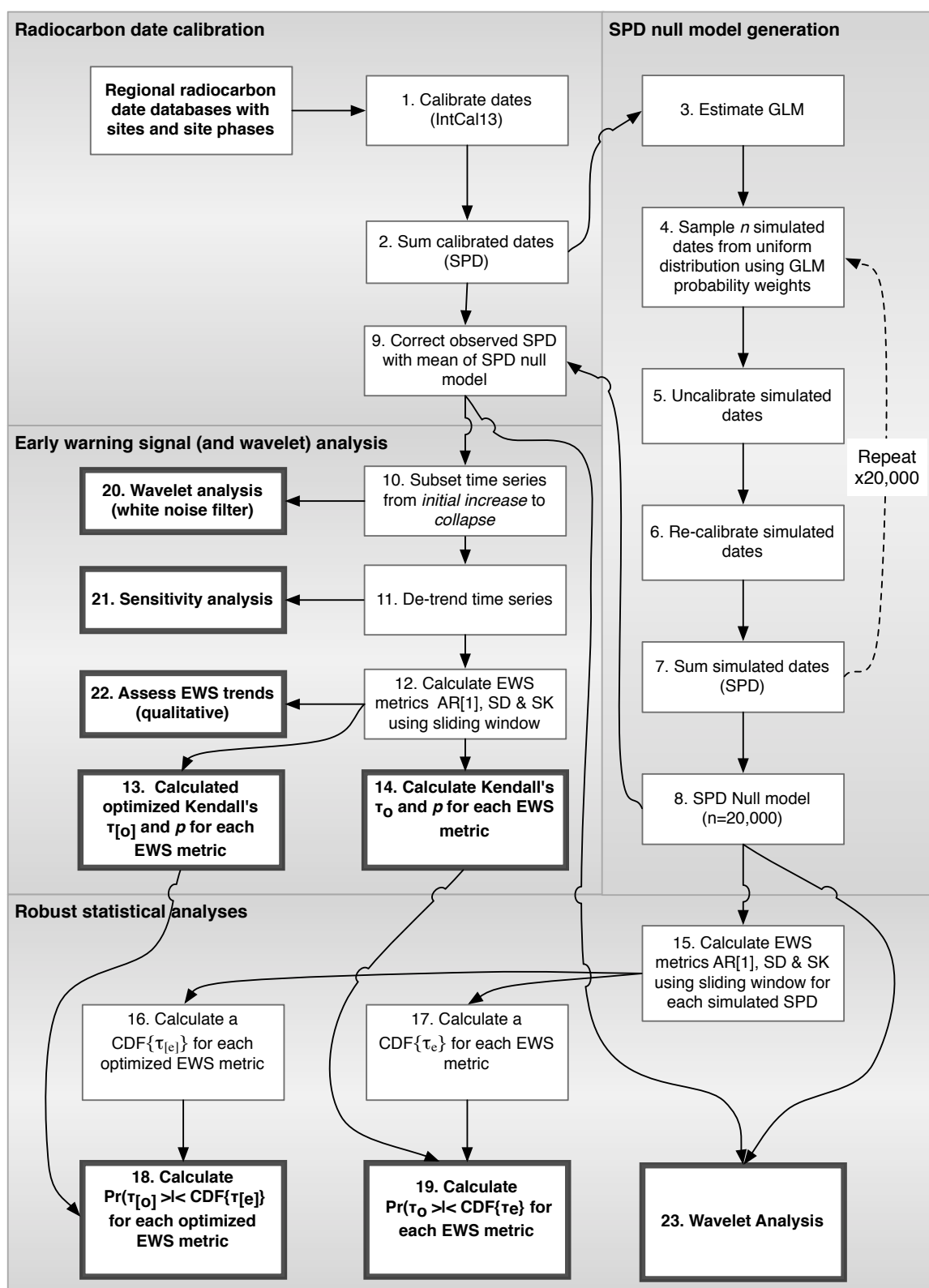


Figure S2: EWS method flowchart.

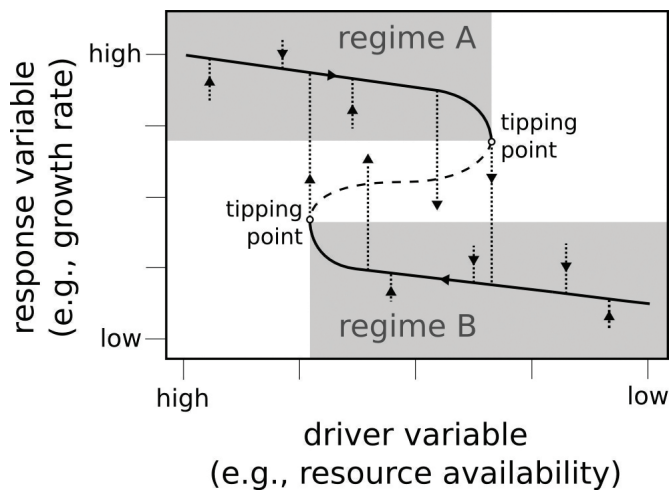


Figure S3: A conceptual fold bifurcation model for human population collapse. We consider that bifurcation acts on population growth rates. Reading the graphic from left to right, early agricultural societies have initially high growth rates. Land degradation slowly drives the growth rates down while agricultural surpluses and cultural mechanisms maintain the growth rates at artificially high levels within growth regime A. As the system approaches the tipping point, it becomes increasingly unstable and susceptible to shifts into the low growth-rate regime (regime B). The system exhibits hysteresis insofar as the recovery of agricultural productivity does not necessarily result in recovery of the higher regime A growth rates.

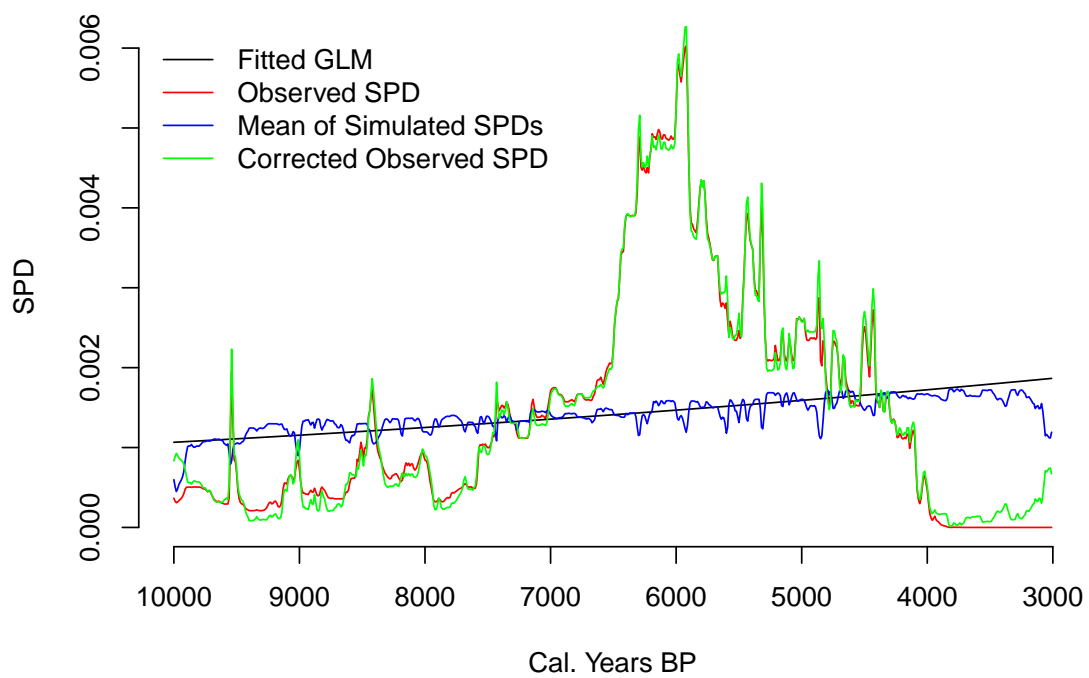


Figure S4: An example of a corrected site density distribution analyzed for EWSs.

Table S1: Qualitative classification of EWS trends in all regions. Here we subjectively classify each EWS trend as increasing (+), decreasing (-), or equivocal (.) and describe the overall pattern including the relationship to interesting events in the underlying SPD.

		EWS Trend			
	Region	AR[1]	σ	1γ	Summary
	Expectation	+	+	-	AR[1] and σ are expected to increase and 1γ is expected to decrease in advance of a regime shift to lower population levels.
1	Southern Germany	+	+	-	Long period of stability between 7-6 kya followed by increasing AR[1] and σ between \sim 6.4-5.6 kya and long-term decreasing trend in 1γ .
2	Eastern Switzerland	+	.	-	Increase in AR[1] beginning \sim 6 kya, and σ also increases around 6.0 kya but then declines \sim 5.8 kya against expectation. 1γ begins to decline just after 6.0 kya.
3	England and Wales (w/o Wessex & Sussex)	+	+	-	Increasing AR[1] beginning slightly after 5.8 kya and σ begins increasing at 5.7 kya. Decreasing 1γ beginning \sim 5.8 kya.
4	Ireland	+	+	-	AR[1] begins increasing slightly after 5.8 kya, and σ begins increasing at 5.7 kya. There is a weak long-term decreasing trend in 1γ .
5	Paris Basin	+	+	-	Increasing trend in AR[1] beginning \sim 6.8 kya and increasing σ \sim 6.3 kya. There is a weak long-term decreasing trend in 1γ .
6	Rhone-Languedoc	+	+	-	Some indication of increasing trends in AR[1] and σ beginning around 6.4 kya. 1γ begins weak decreasing trend after increase around 6.3 kya.
7	Scotland	+	+	.	AR[1] increase beginning around 5.8 kya, and σ increase beginning around 5.7 kya with no clear 1γ trend.
8	Southern England (Wessex & Sussex)	.	+	-	No clear trend in AR[1], but clear and strong increasing trend in σ beginning around \sim 6.0 kya. Very slight long-term decrease in 1γ .
9	Western France	+	+	.	Long-term increasing AR[1] beginning \sim 6.4 kya, and strong long-term increasing σ trend. Equivocal, or increasing trend in 1γ against expectation.

Table S3: EWS analysis results of summed probability distributions of nine European regions using the complete (non-optimized) subset.

region	AR[1]			σ			1γ		
	$\tau_{[o]}$	$p_{[s]}$	$p_{[n]}$	$\tau_{[o]}$	$p_{[s]}$	$p_{[n]}$	$\tau_{[o]}$	$p_{[s]}$	$p_{[n]}$
Southern Germany	-0.20	1.00	0.52	-0.01	0.60	0.57	-0.61	0.00***	0.03**
Eastern Switzerland	0.05	0.28	0.50	0.07	0.21	0.35	-0.07	0.21	0.46
England and Wales (w/o Wessex & Sussex)	-0.27	0.99	0.98	0.34	0.00***	0.18	-0.15	0.10*	0.25
Ireland	-0.08	0.76	0.81	0.32	0.00***	0.18	-0.25	0.01***	0.18
Paris Basin	0.33	0.00***	0.14	0.03	0.34	0.52	-0.30	0.00***	0.18
Rhone-Languedoc	-0.12	0.99	0.52	0.67	0.00***	0.01***	0.26	1.00	0.73
Scotland	0.08	0.28	0.36	0.00	0.51	0.41	0.17	0.91	0.79
Southern England (Wessex & Sussex)	0.15	0.03**	0.34	0.68	0.00***	0.06*	-0.15	0.02	0.31
Western France	-0.19	0.99	0.55	0.61	0.00***	0.04**	0.43	1.00	0.91

Significance levels: *** p 0.01; ** p 0.05; * p 0.10

Table S4: Fisher's exact tests for multiple analyses using the complete (non-optimized) subset.

Data subset	Test	EWS	pass at $p < 0.1$	fail at $p < 0.1$	Fisher CI [†]		Fisher p [†]
					low	high	
Tau for all data (τ_o)	sampling (p_s)	AR[1]	2	7	0.13	201.64	0.55
		σ	5	4	0.94	703.78	0.05**
		1γ	5	4	0.94	703.78	0.05**
	null model (p_n)	AR[1]	0	9	0.00	51.96	1.00
		σ	3	6	0.32	313.65	0.27
		1γ	1	8	0.02	117.49	1.00

[†]Expected ratio used in Fisher test is 1:8, pass:fail.

Significance levels: *** p 0.01; ** p 0.05; * p 0.10

Table S2: Summary of 14C samples and Neolithic events for the European cases examined in this study.

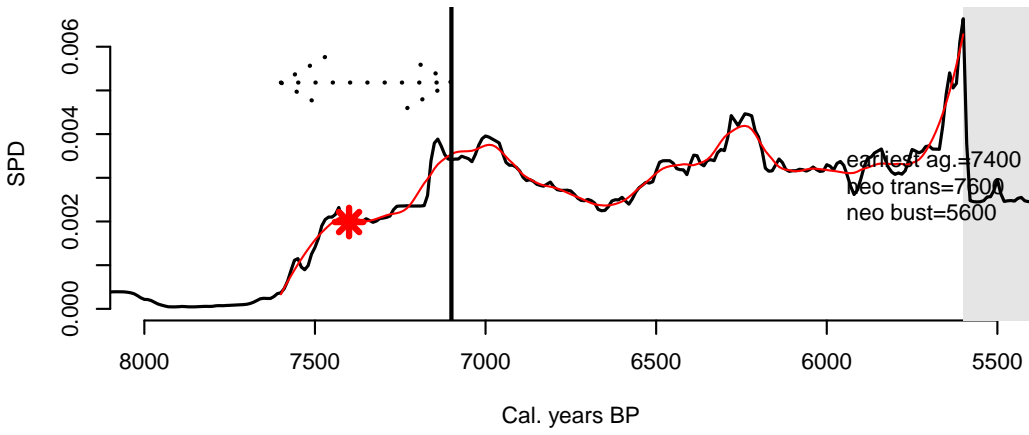
Region	Collapse date (cal. years BP)*	Neolithic transition	Neolithic bust	Duration	Complete time series		EWS analysis			
					Uncalibrated dates	Site-phases	Uncalibrated dates	Site-phases	Sites	
Southern Germany	7400	7600	5600	2000	391	119	80	185	51	35
Eastern Switzerland	6500	6500	5650	850	275	89	48	47	13	9
England and Wales (w/o Wessex & Sussex)	6000	6100	5600	500	1188	782	509	45	38	34
Ireland	6000	6100	5600	500	1721	1007	607	71	39	30
Paris Basin	7200	7500	6225	1275	570	300	176	90	57	51
Rhone-Languedoc	7800	7800	5800	2000	978	568	359	316	178	124
Scotland	6000	6000	5600	400	579	326	210	58	32	29
Southern England (Wessex & Sussex)	6000	6400	5400	1000	581	286	176	25	20	16
Western France	7000	7000	5940	1060	494	298	213	73	51	40
Totals:					6777	3775	2378	910	479	368

*collapse dates from [17].

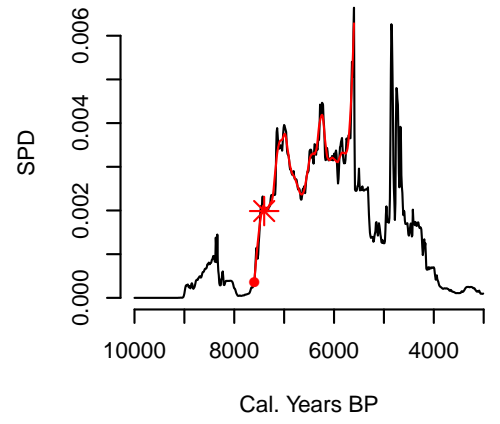
SI Dataset

EWS analysis for all regions.

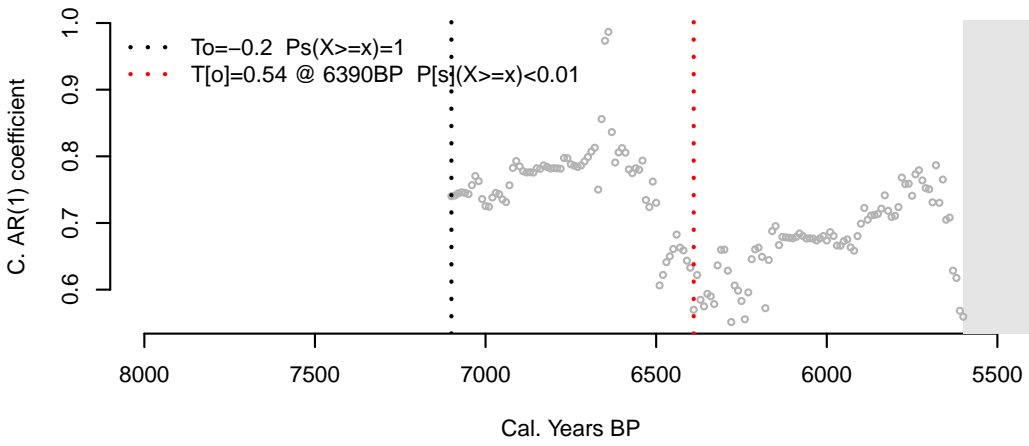
A. Southern Germany beginning of increase to collapse



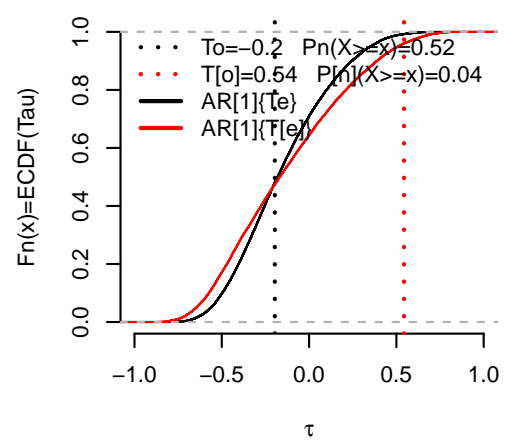
B. Complete time series



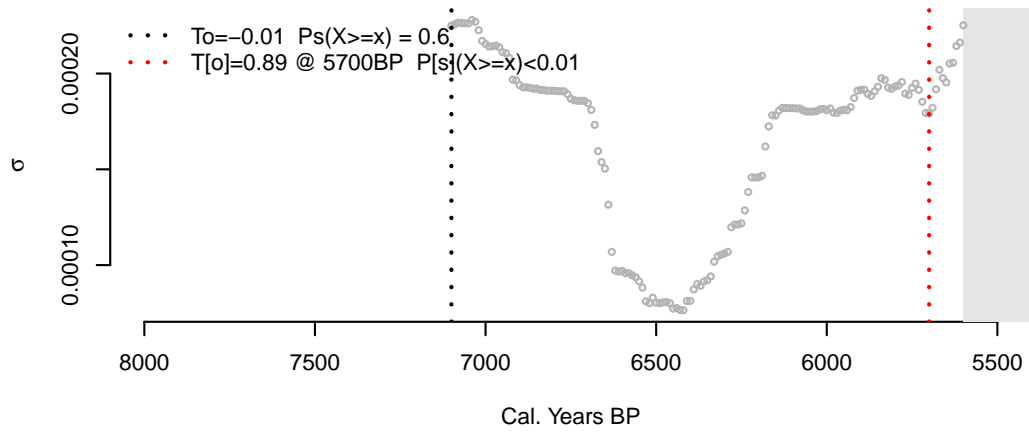
C. AR[1]



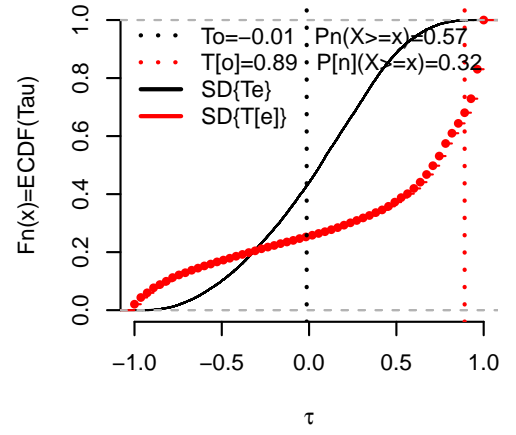
D. AR[1]{Te}



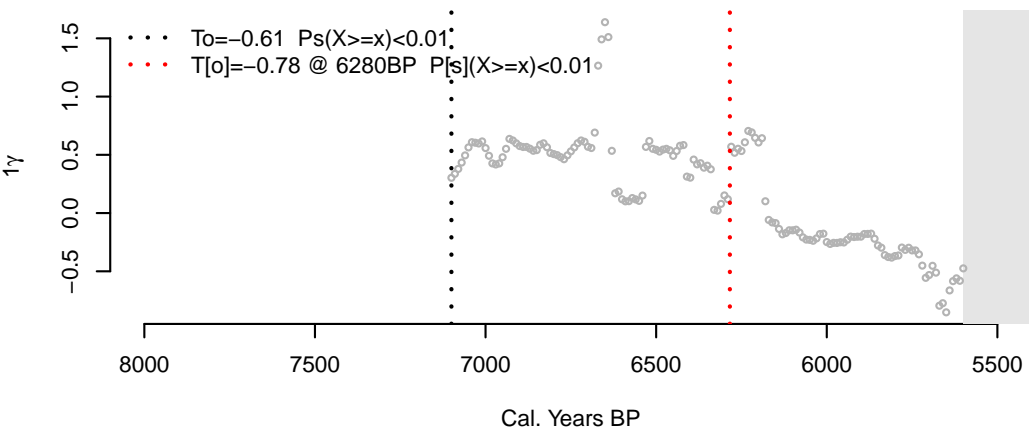
E. Standard Deviation



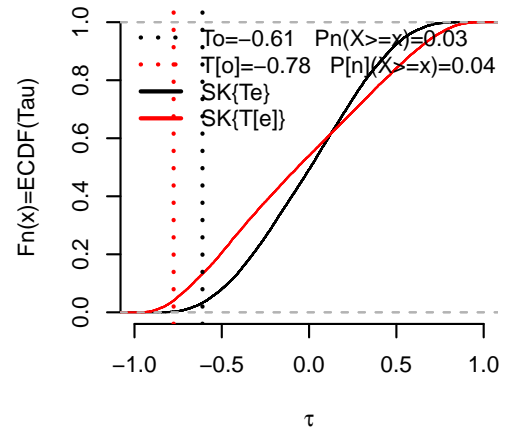
F. SD{Te}



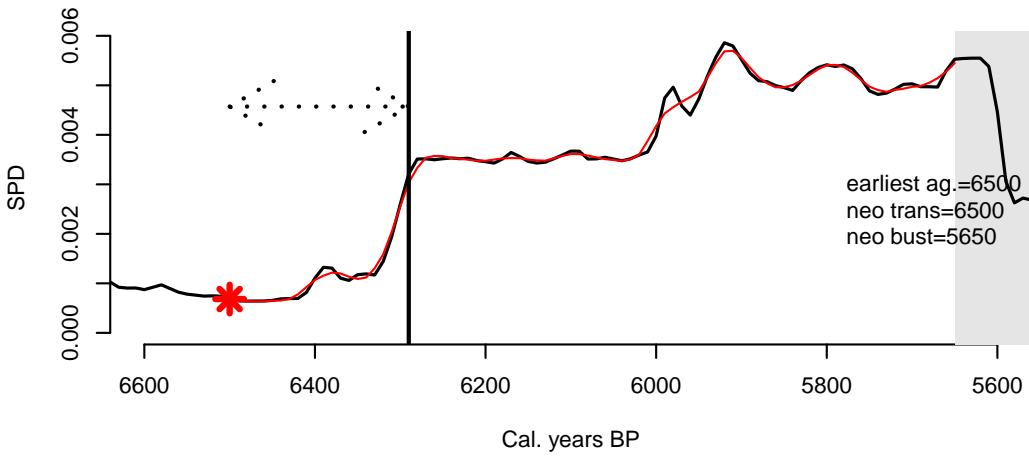
G. Skewness



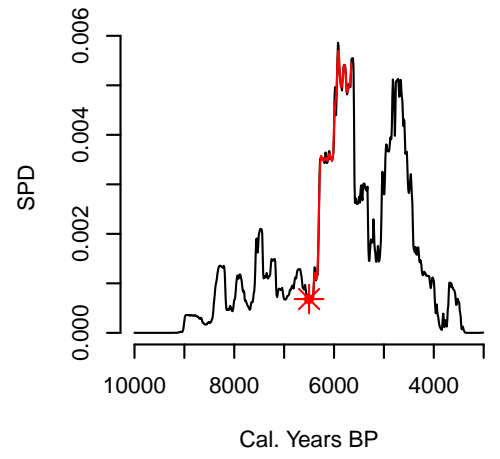
H. SK{Tau sim}



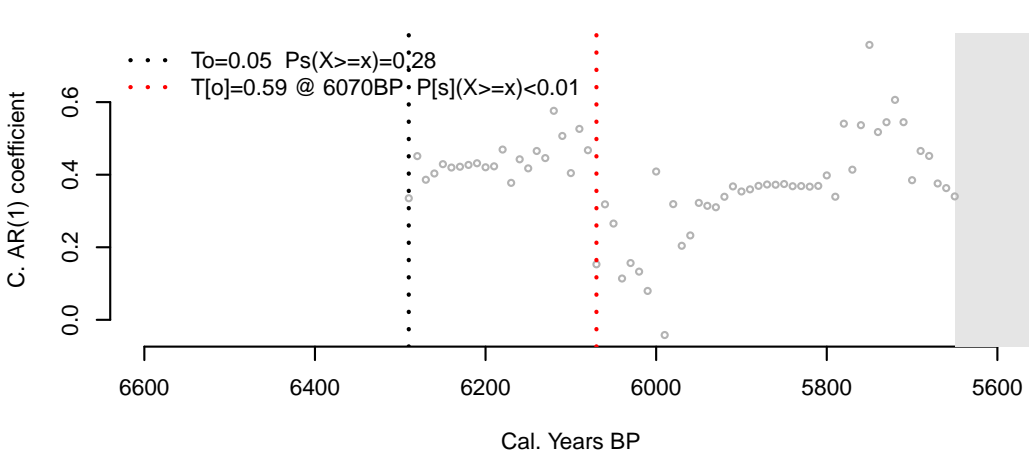
A. Eastern Switzerland beginning of increase to collapse



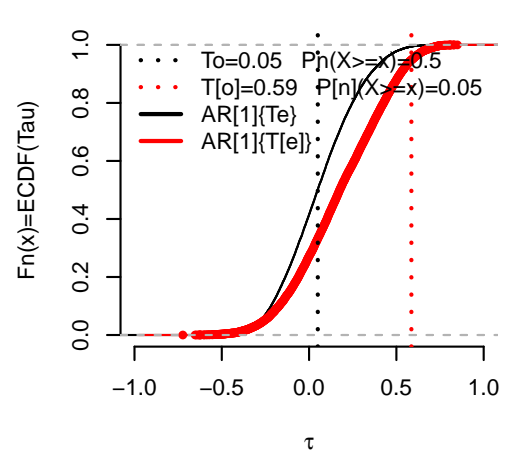
B. Complete time series



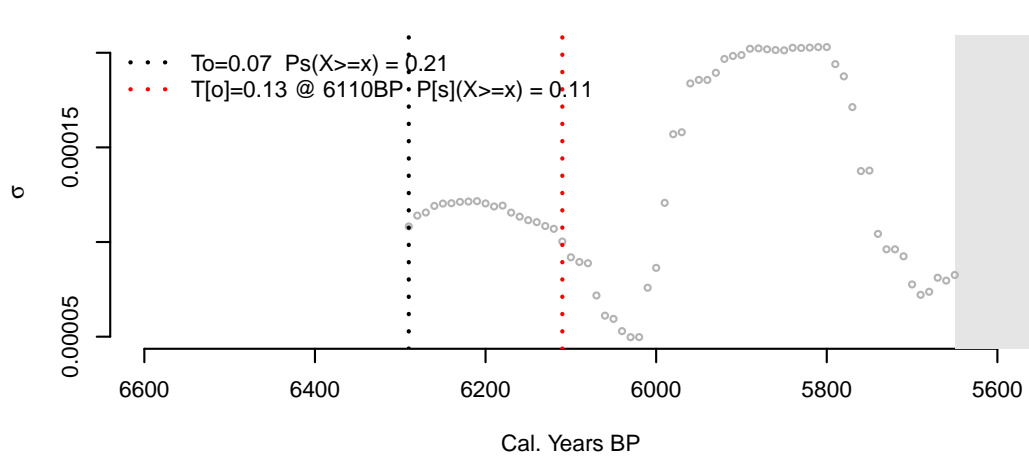
C. AR[1]



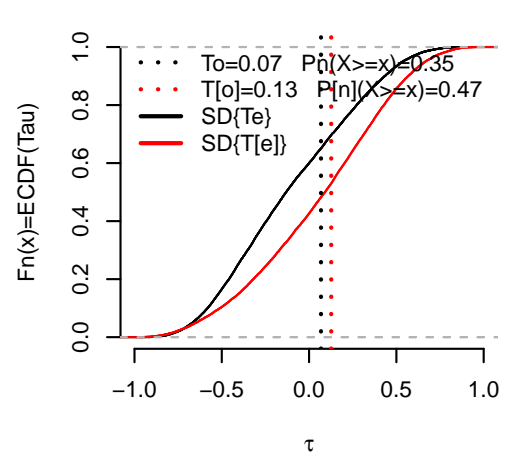
D. AR[1]{Te}



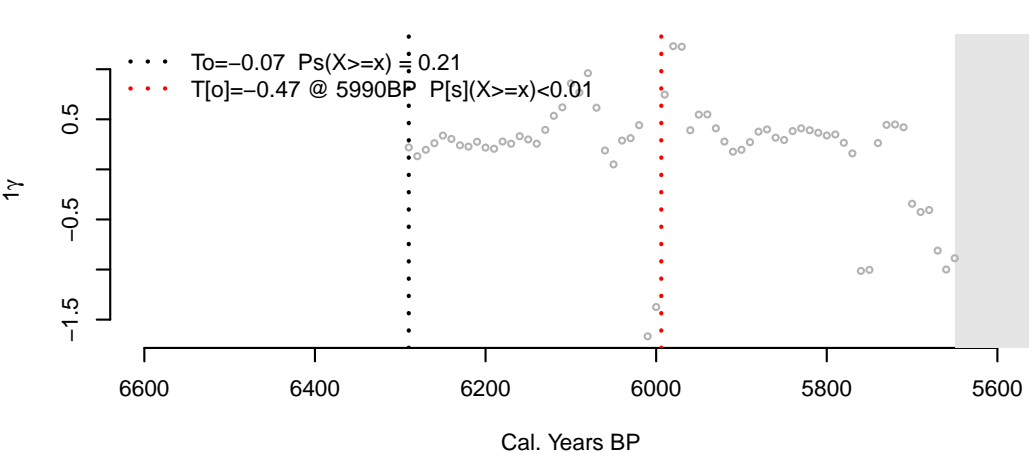
E. Standard Deviation



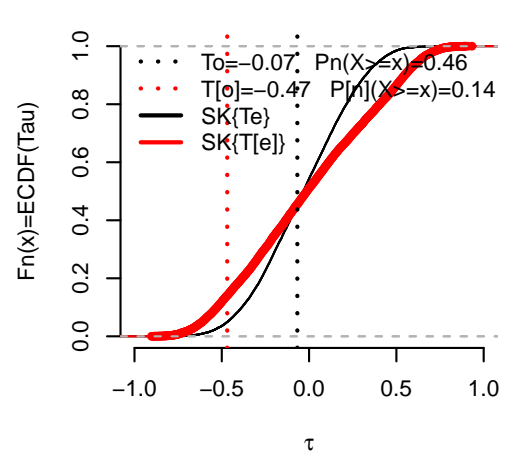
F. SD{Te}



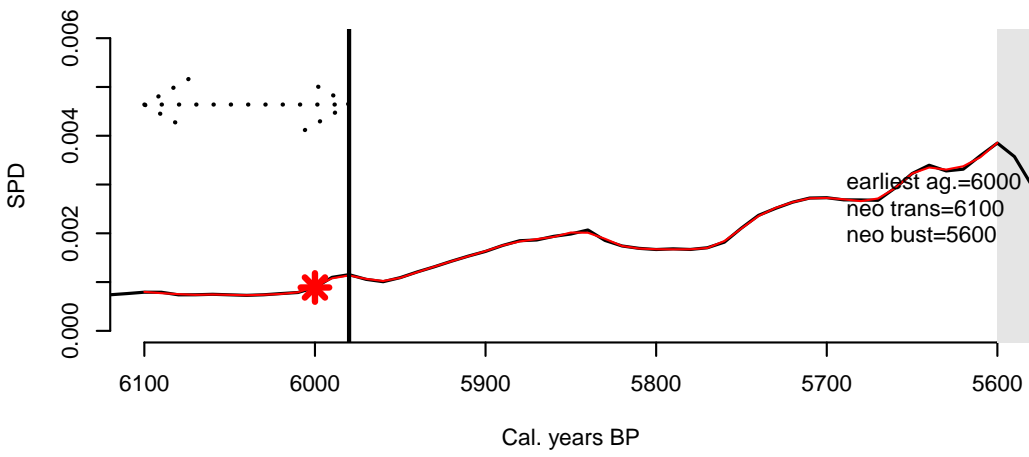
G. Skewness



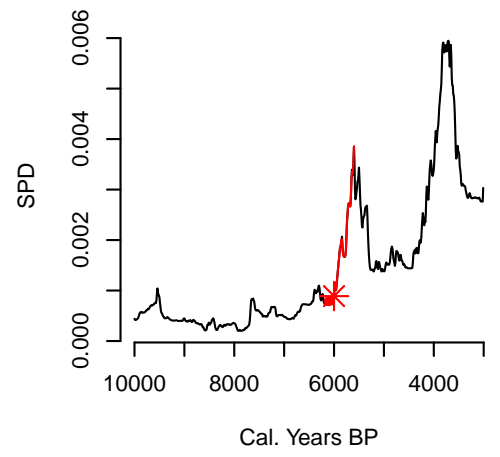
H. SK{Tau sim}



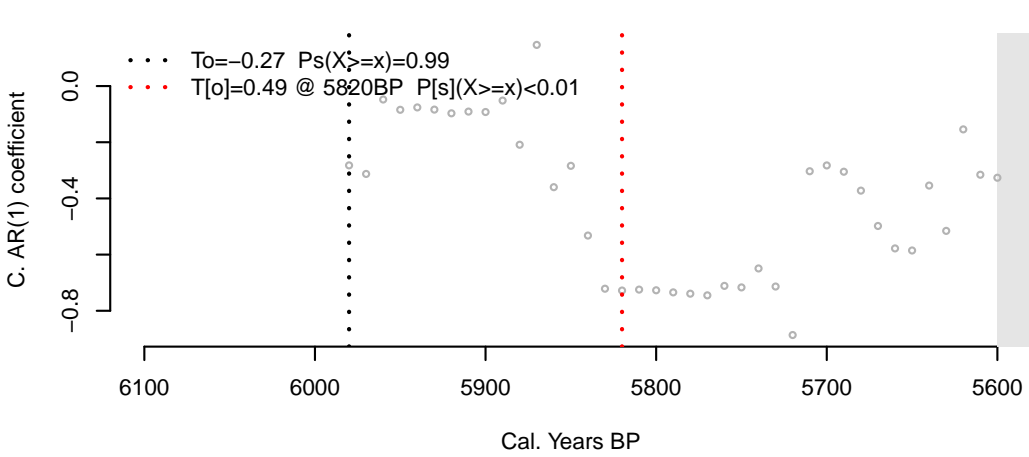
A. England and Wales (w/o Wessex & Sussex) beginning of increase to collaps



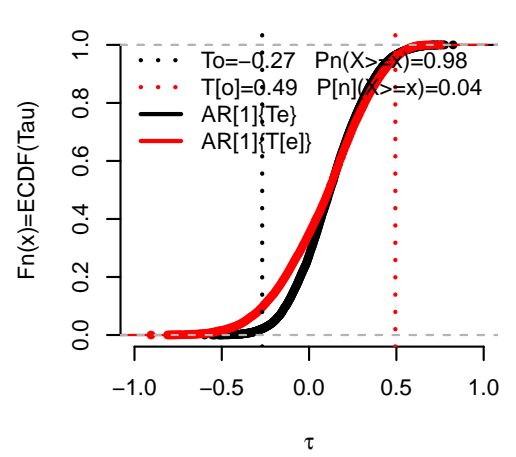
B. Complete time series



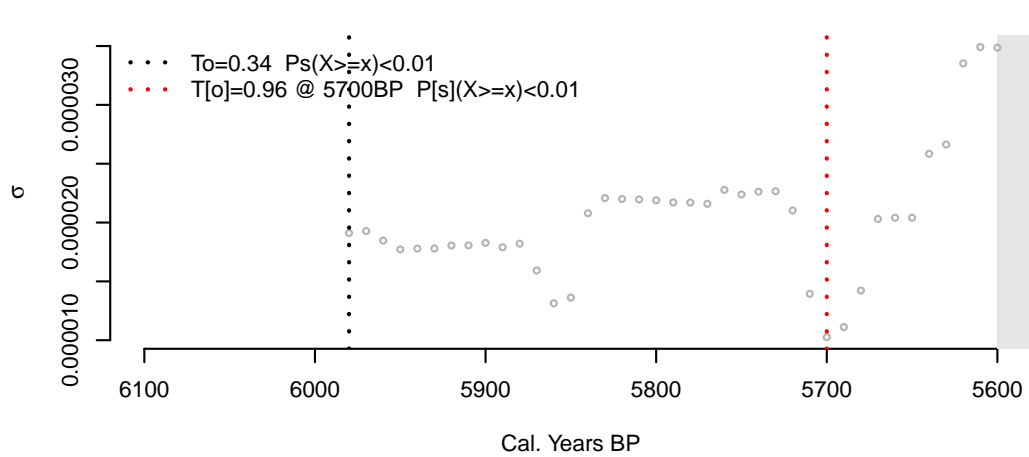
C. AR[1]



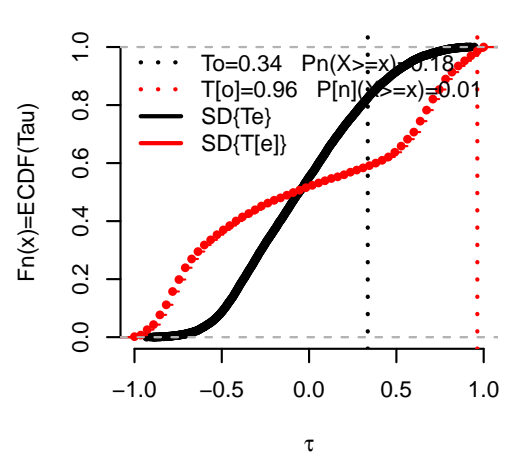
D. AR[1]{Te}



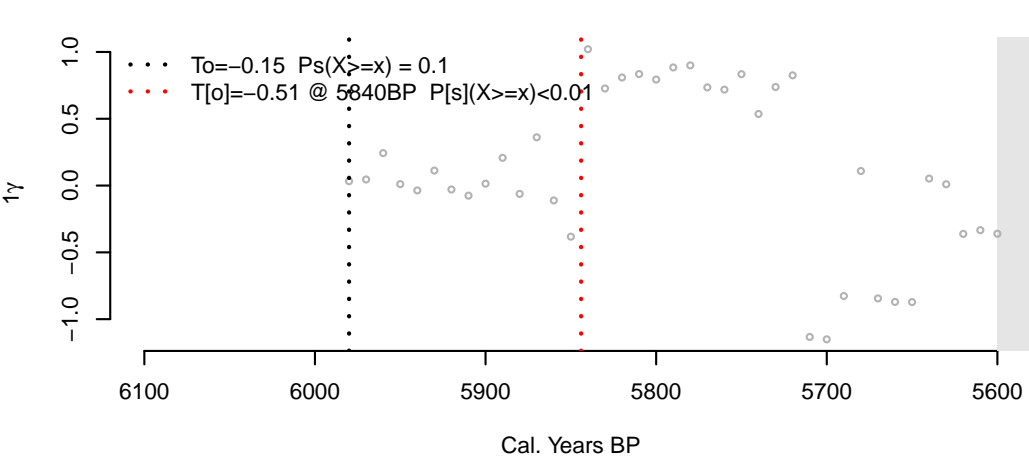
E. Standard Deviation



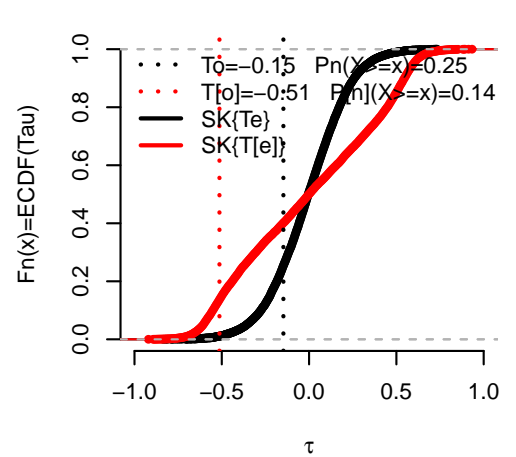
F. SD{Te}



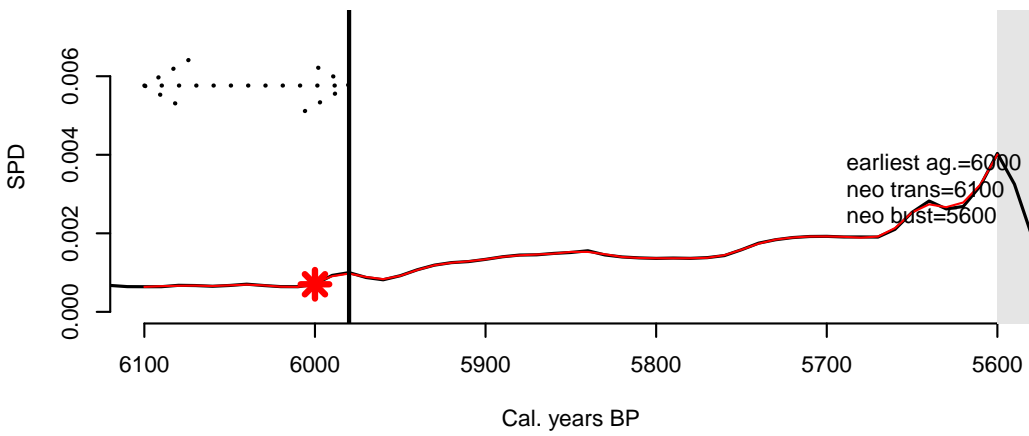
G. Skewness



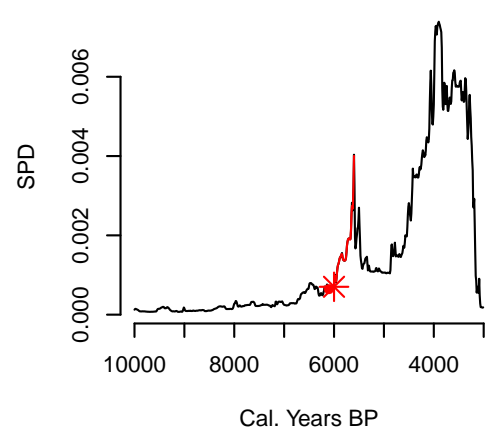
H. SK{Tau sim}



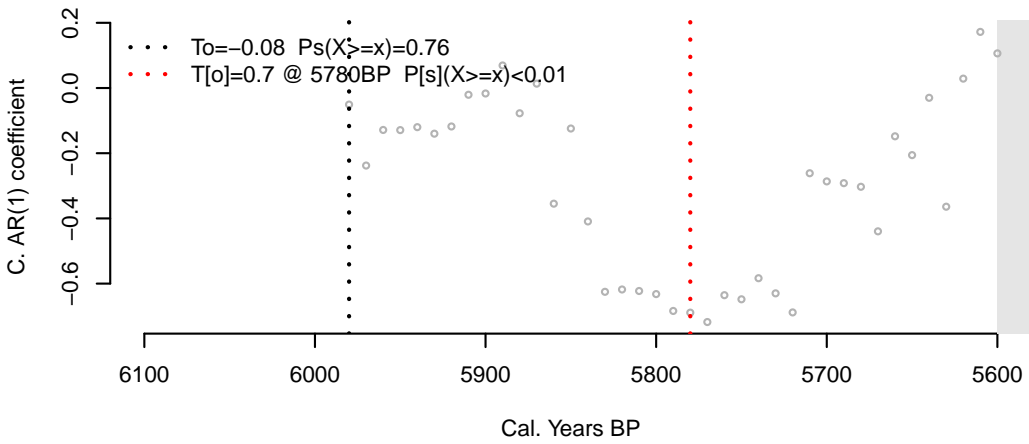
A. Ireland beginning of increase to collapse



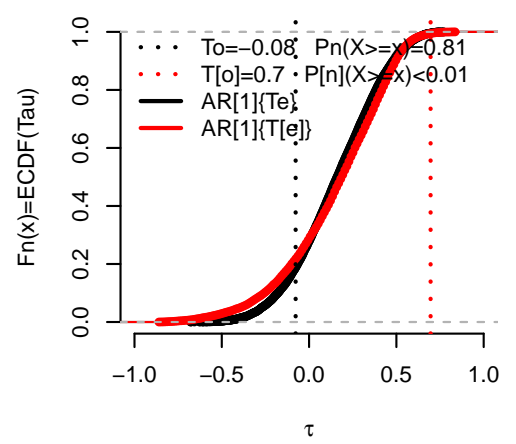
B. Complete time series



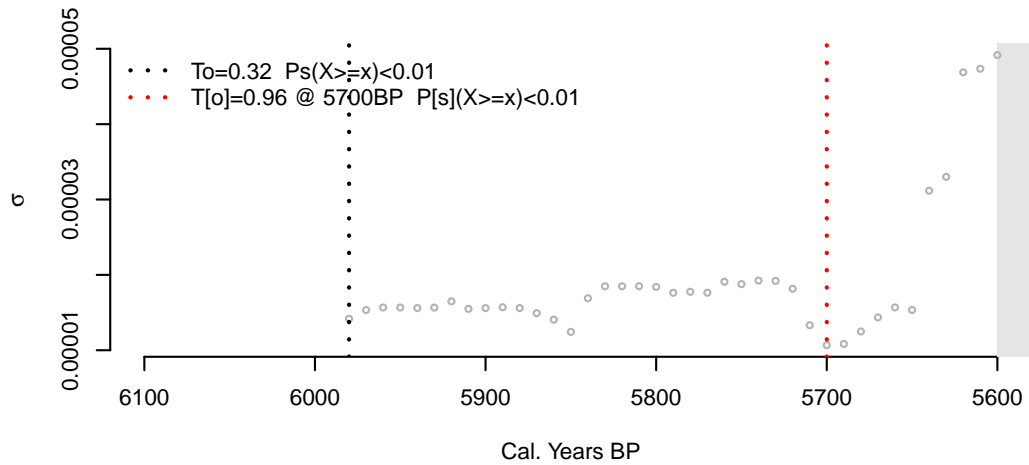
C. AR[1]



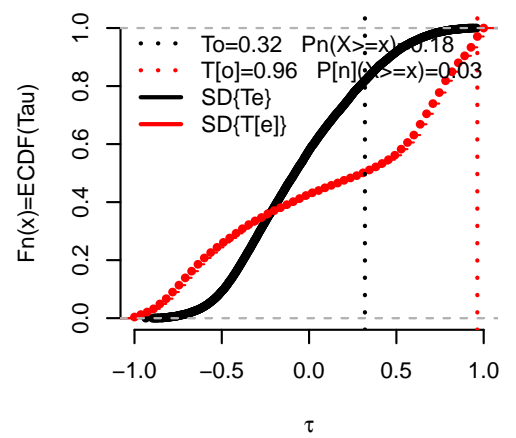
D. AR[1]{Te}



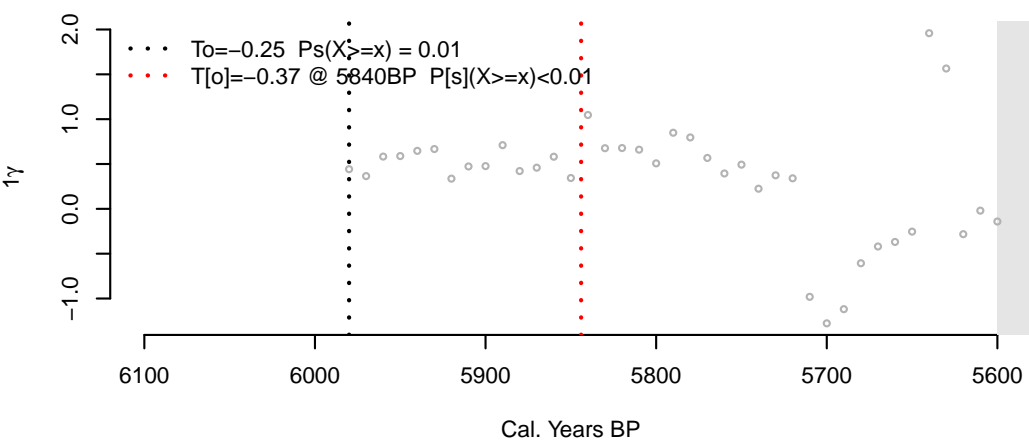
E. Standard Deviation



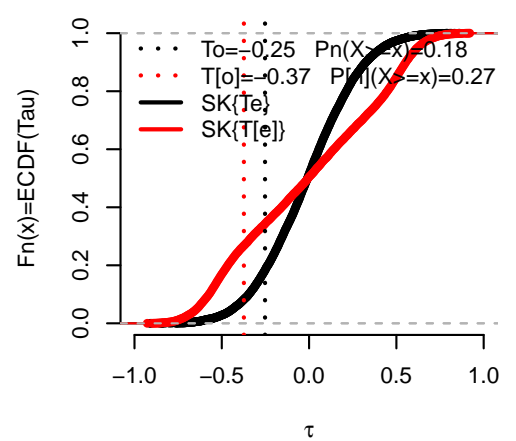
F. SD{Te}

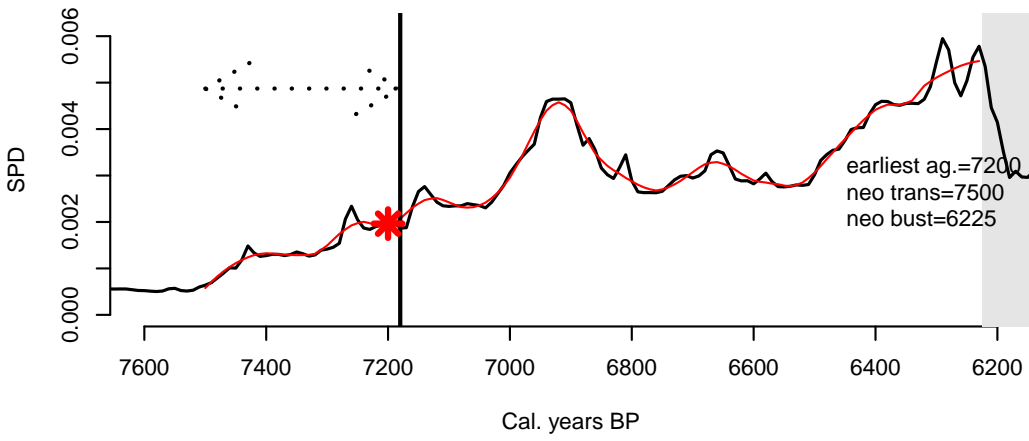
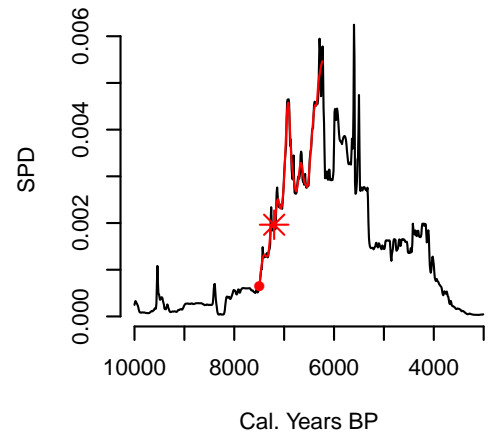
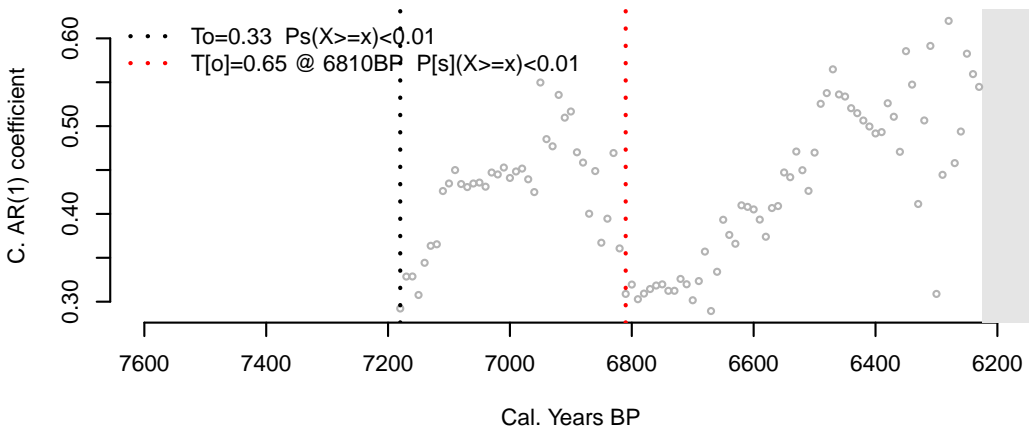
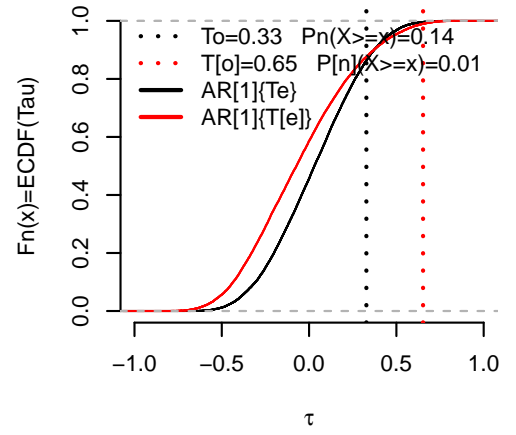
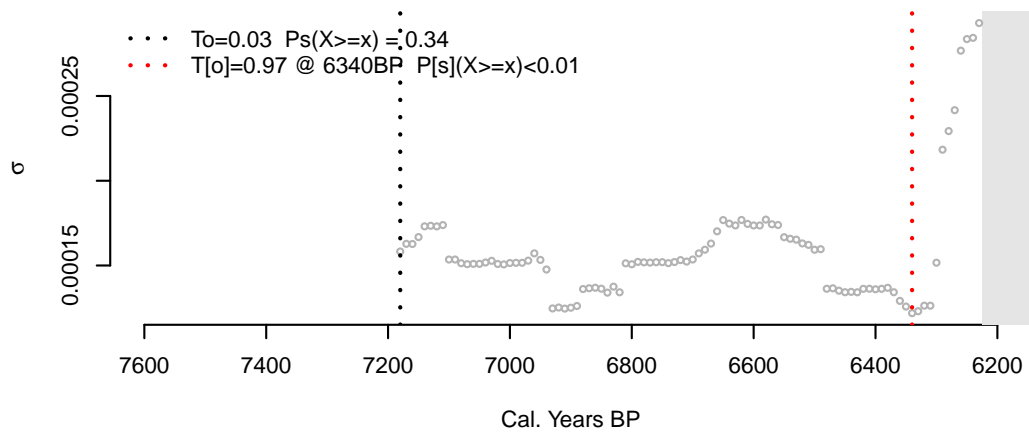
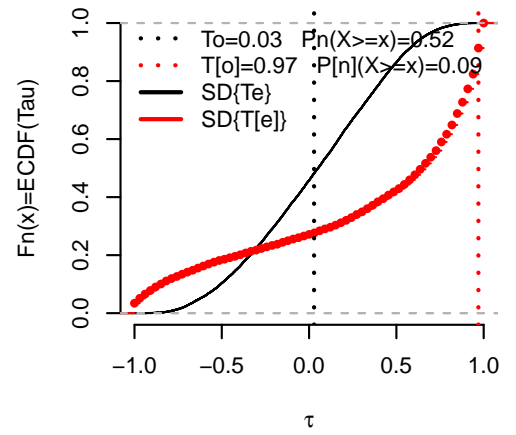
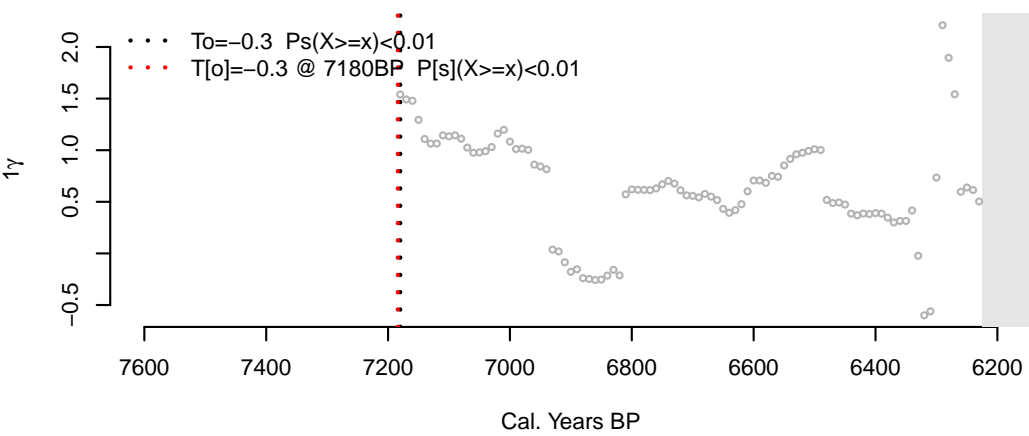
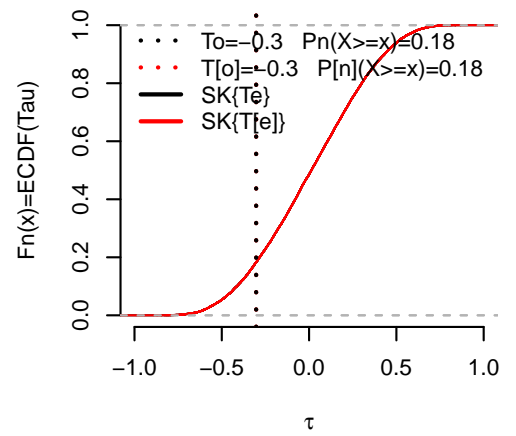


G. Skewness

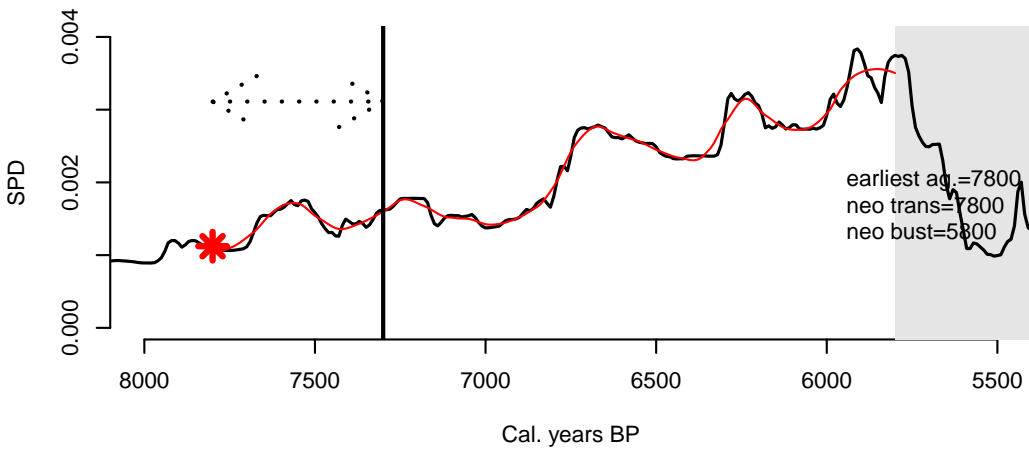


H. SK{Tau sim}

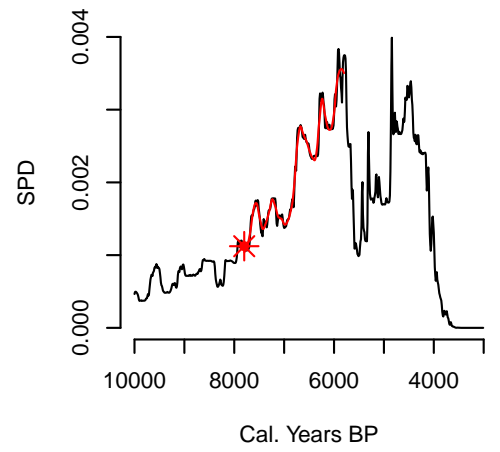


A. Paris Basin beginning of increase to collapse**B. Complete time series****C. AR[1]****D. AR[1]{Te}****E. Standard Deviation****F. SD{Te}****G. Skewness****H. SK{Tau sim}**

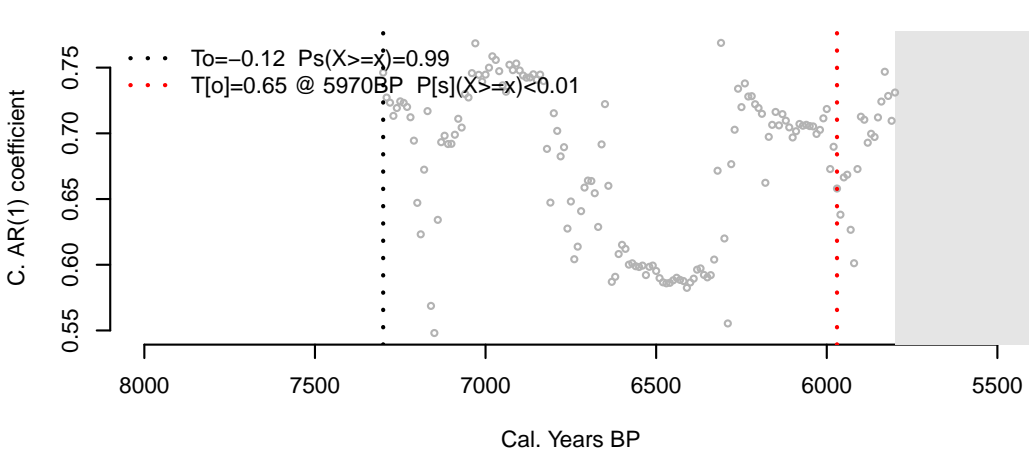
A. Rhone-Languedoc beginning of increase to collapse



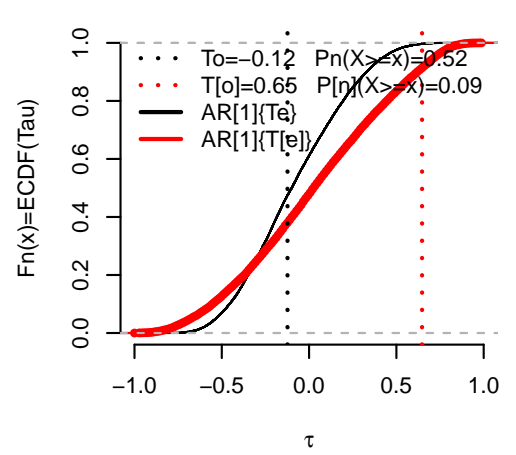
B. Complete time series



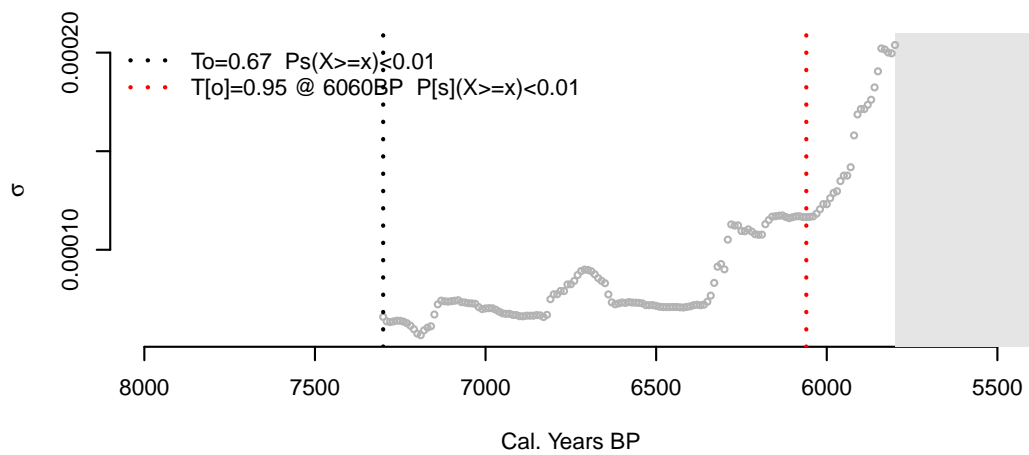
C. AR[1]



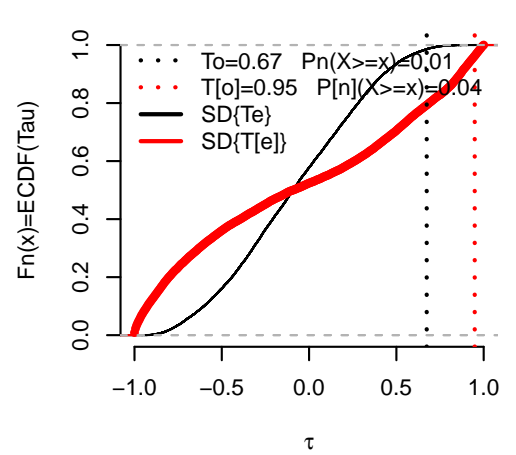
D. AR[1]{Te}



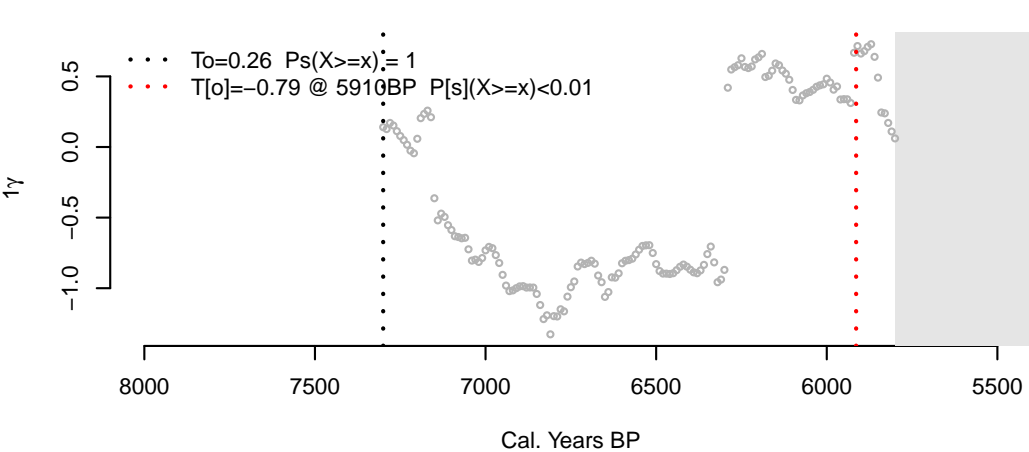
E. Standard Deviation



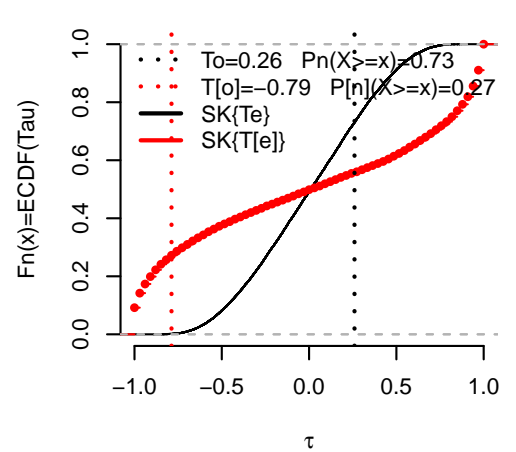
F. SD{Te}



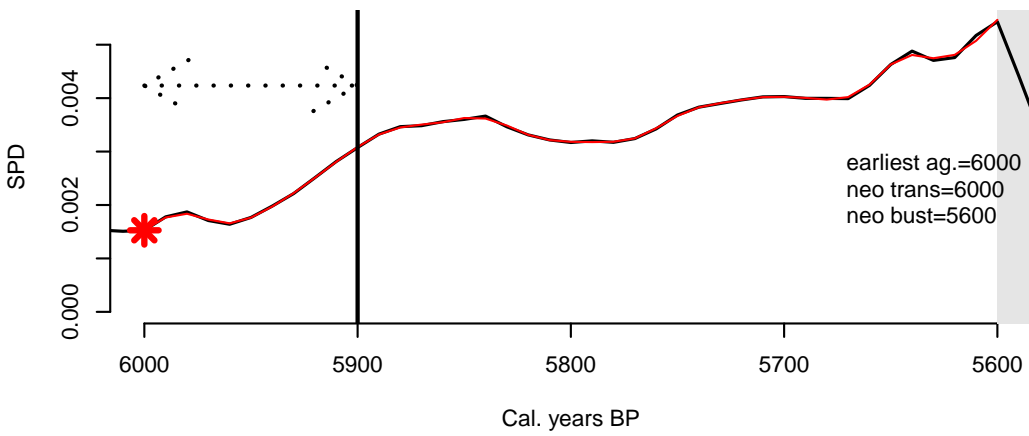
G. Skewness



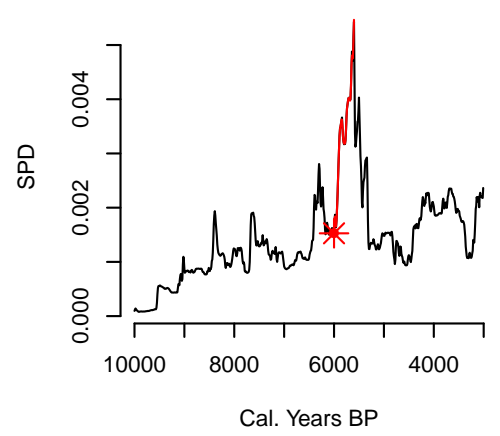
H. SK{Tau sim}



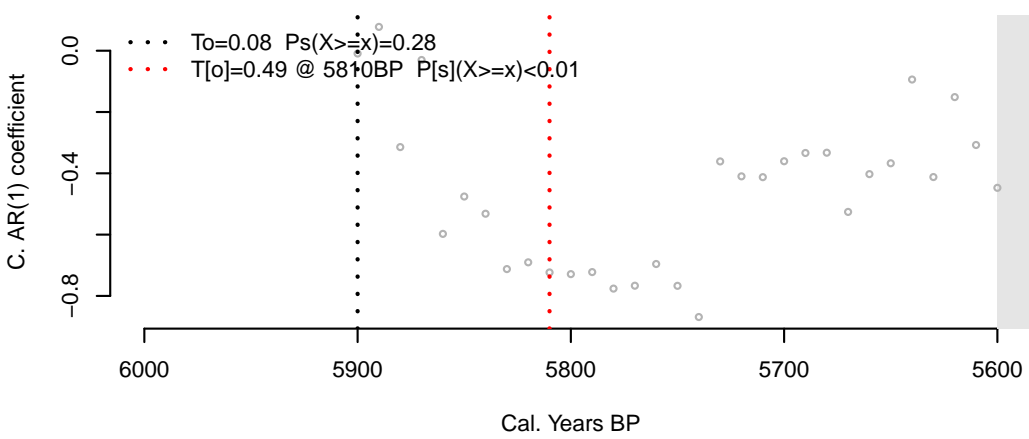
A. Scotland beginning of increase to collapse



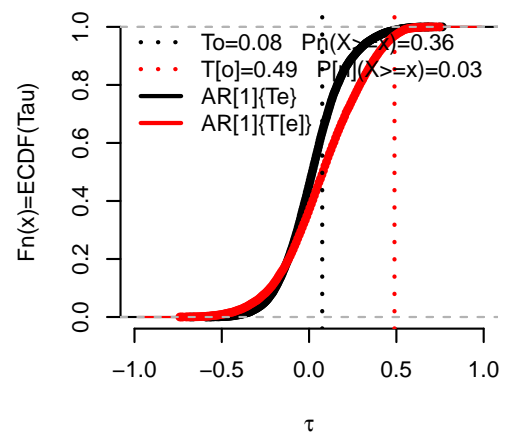
B. Complete time series



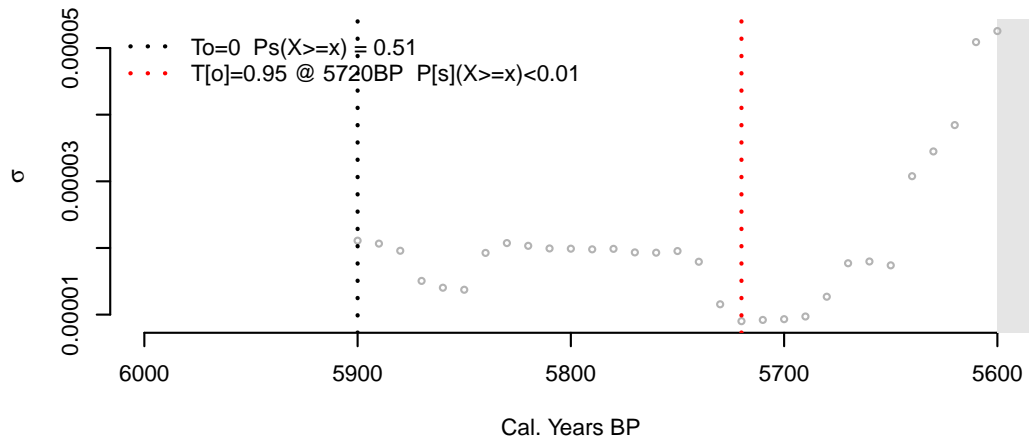
C. AR[1]



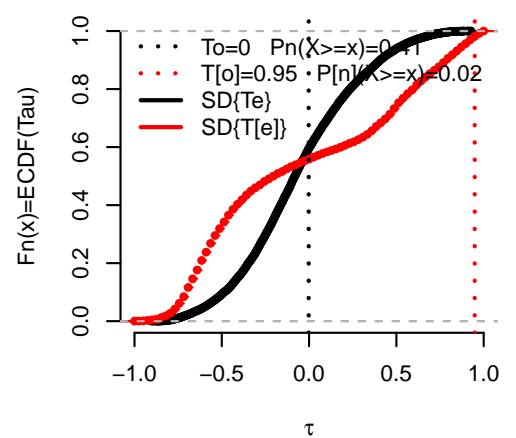
D. AR[1]{Te}



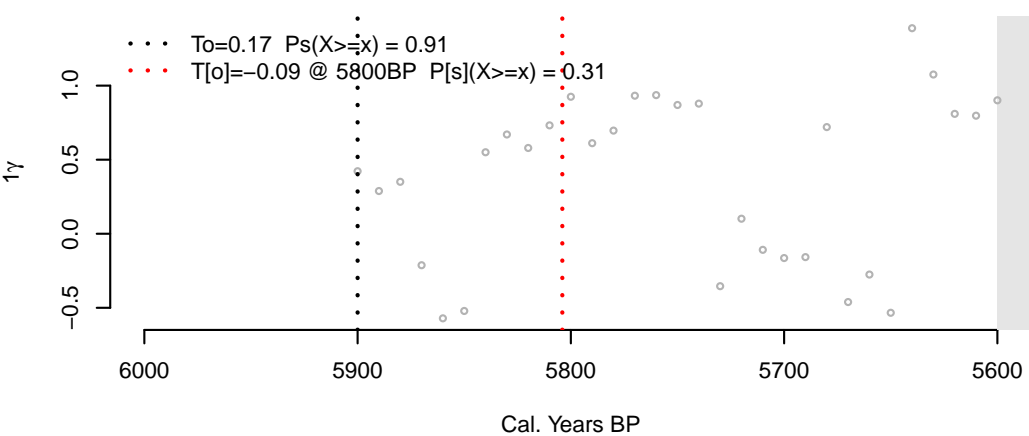
E. Standard Deviation



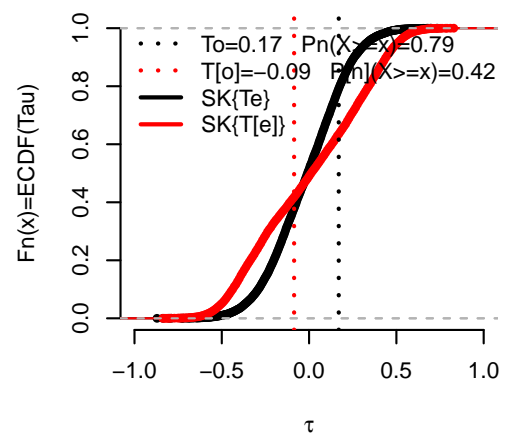
F. SD{Te}



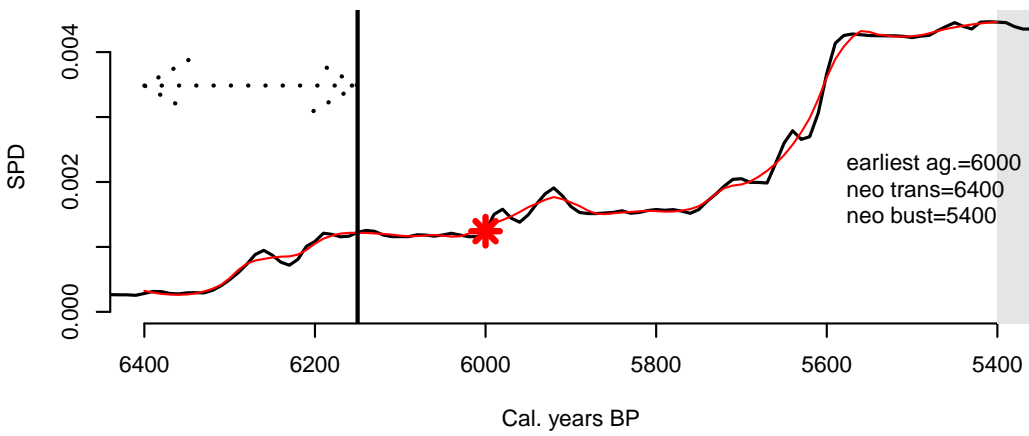
G. Skewness



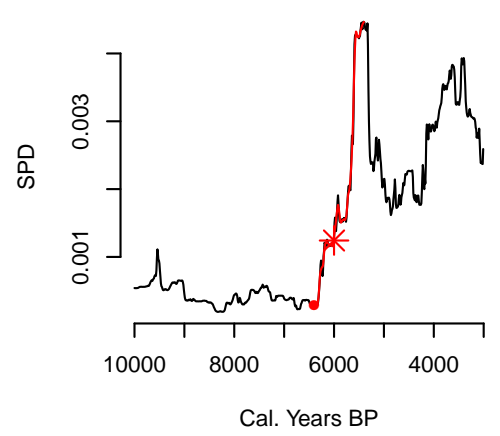
H. SK{Tau sim}



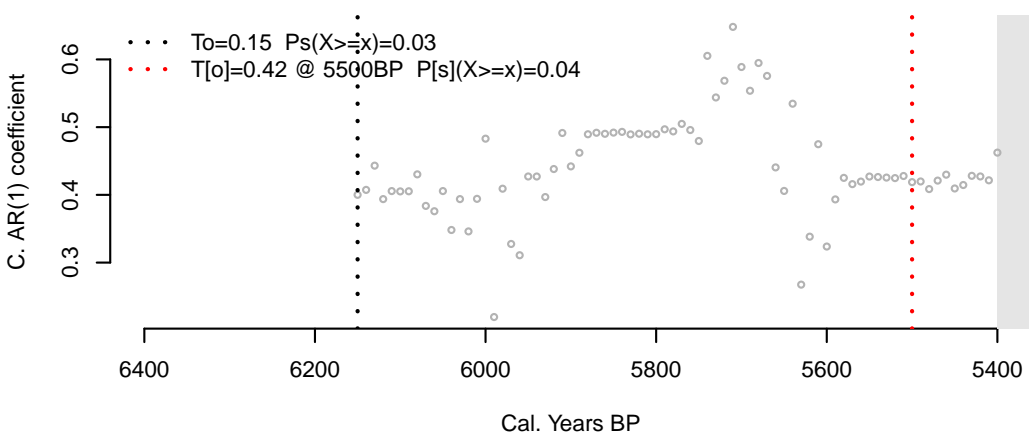
A. Southern England (Wessex & Sussex) beginning of increase to collapse



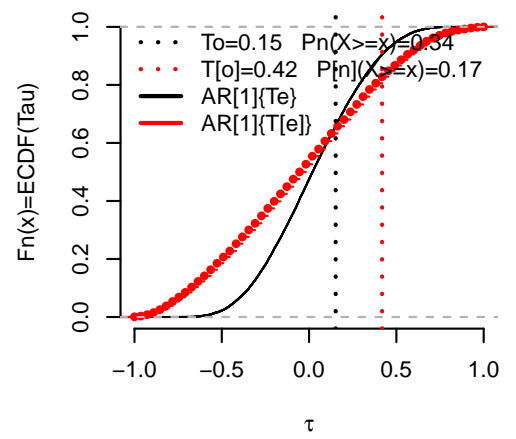
B. Complete time series



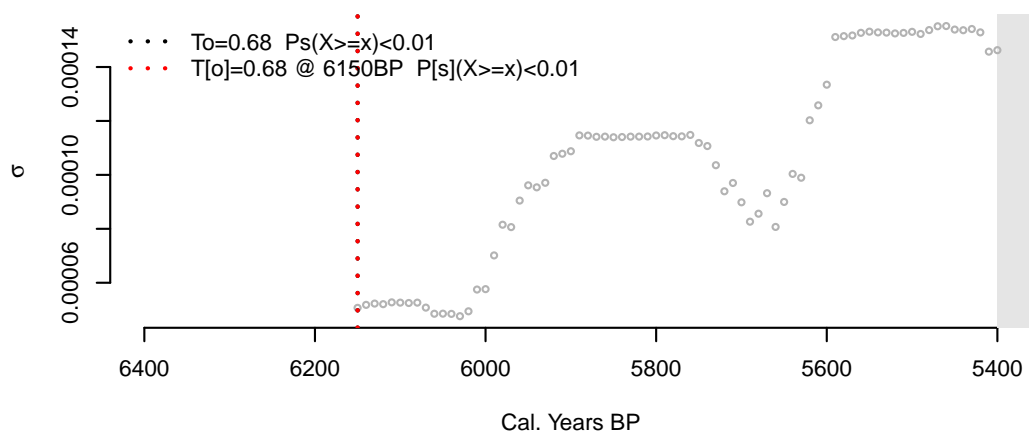
C. AR[1]



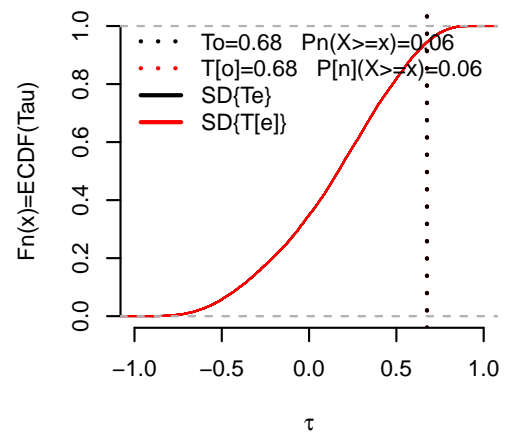
D. AR[1]{Te}



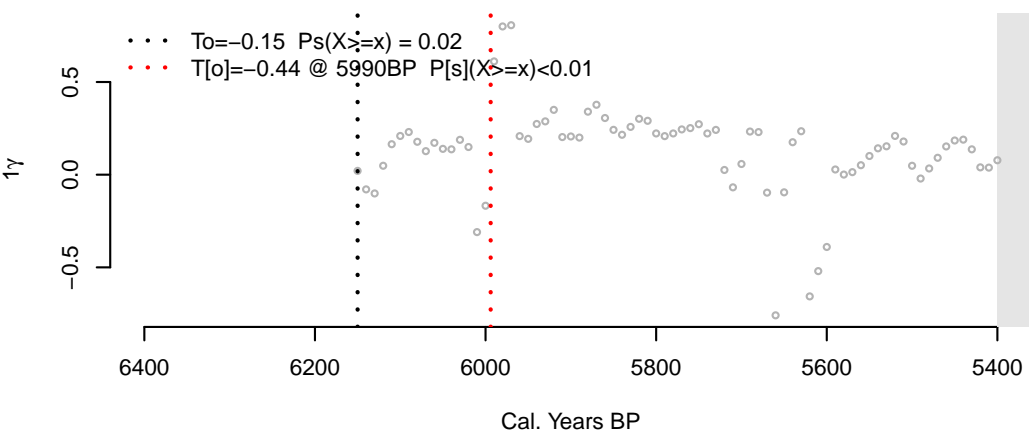
E. Standard Deviation



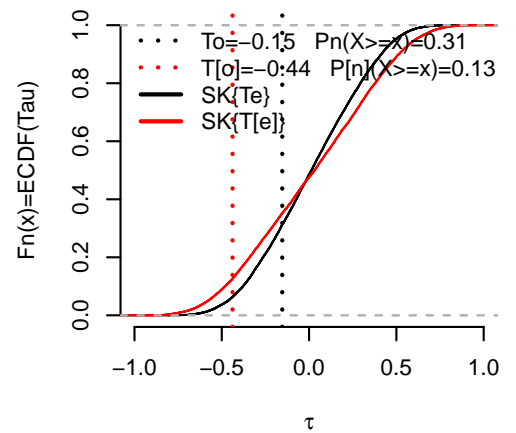
F. SD{Te}



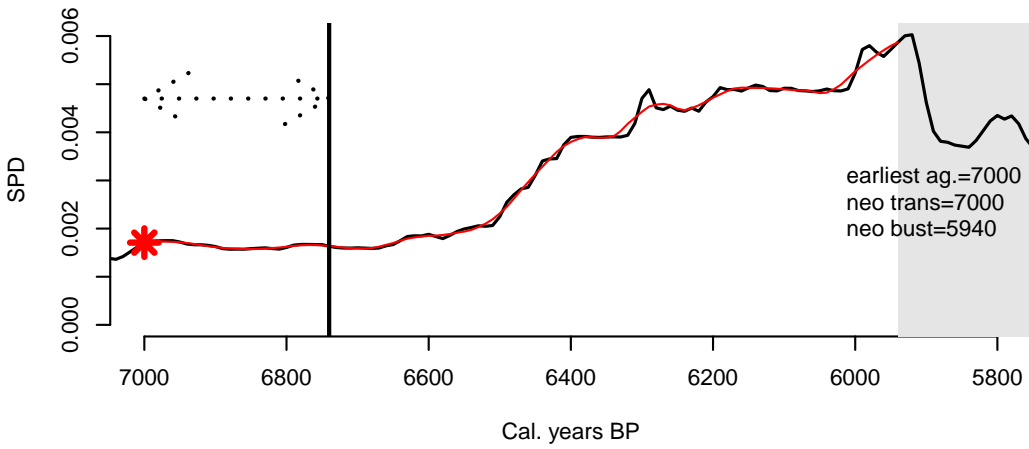
G. Skewness



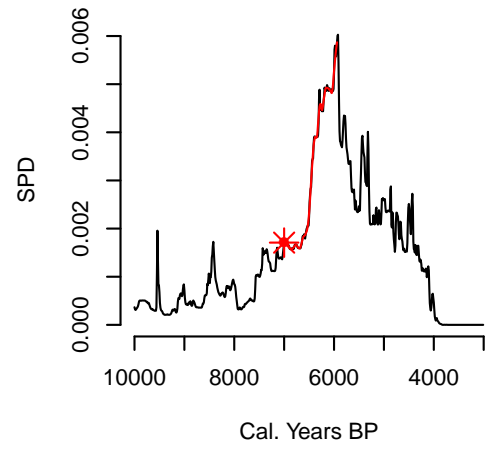
H. SK{Tau sim}



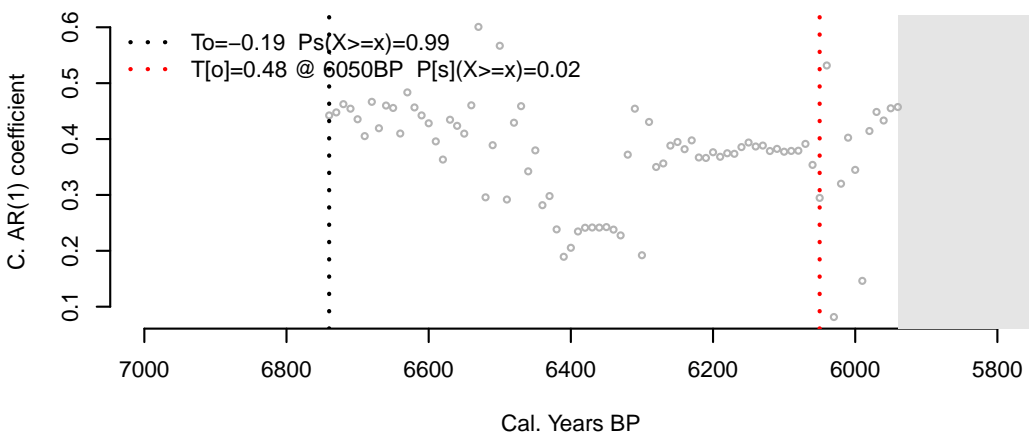
A. Western France beginning of increase to collapse



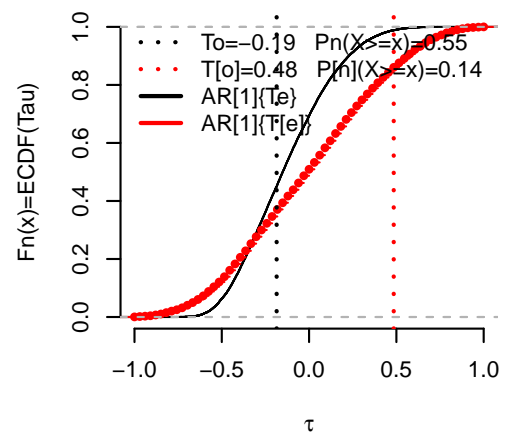
B. Complete time series



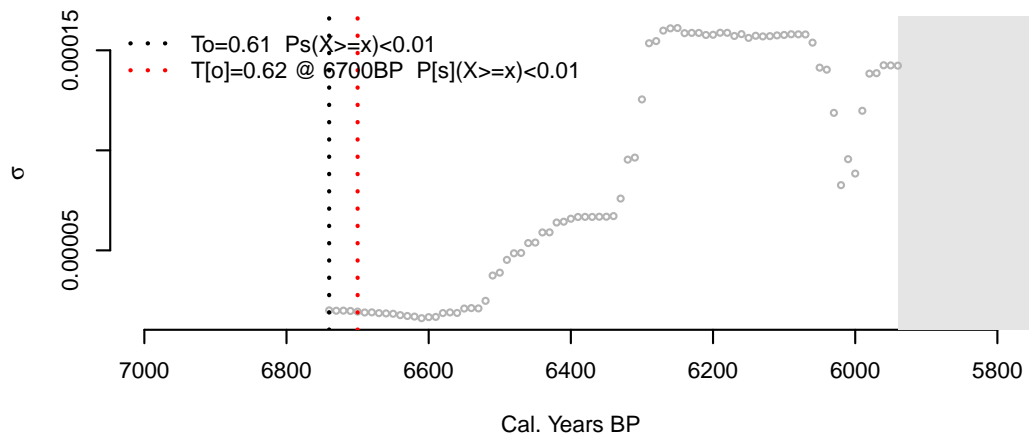
C. AR[1]



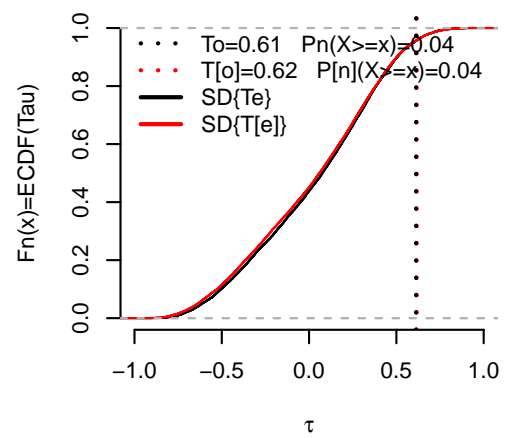
D. AR[1]{Te}



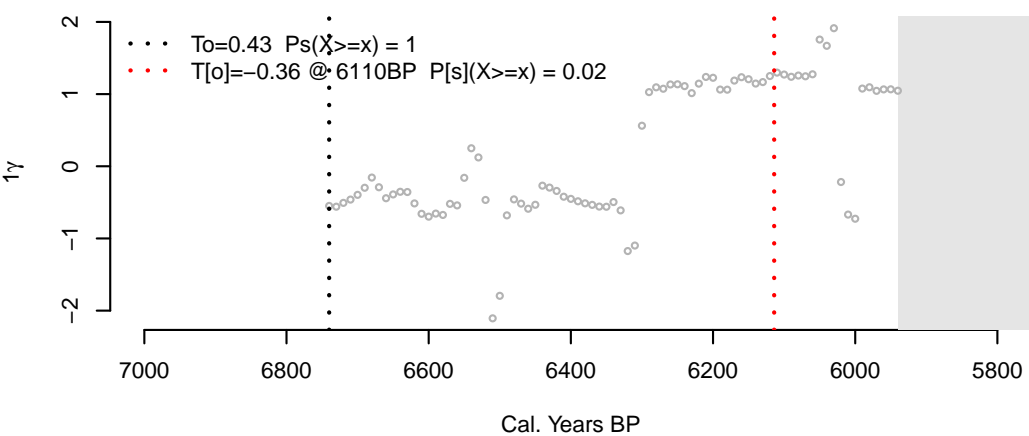
E. Standard Deviation



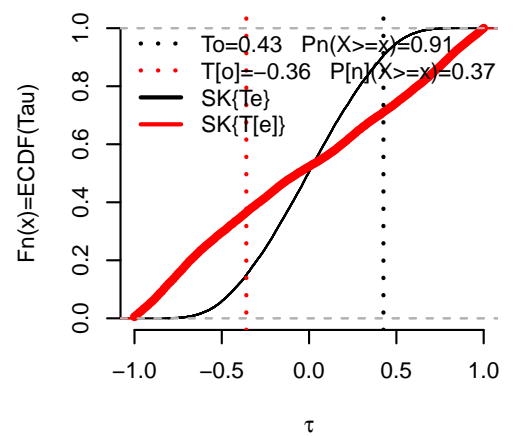
F. SD{Te}



G. Skewness

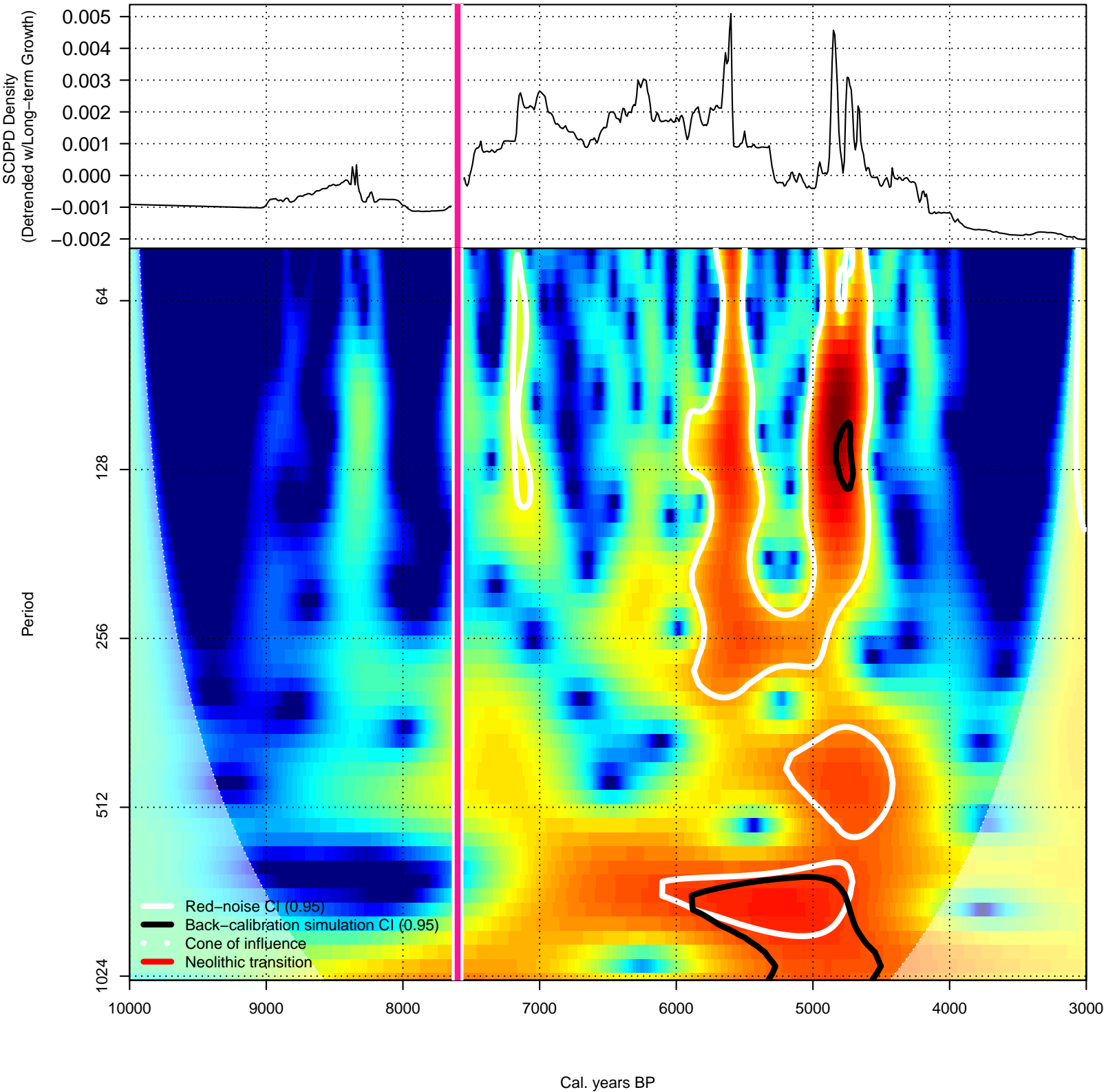


H. SK{Tau sim}

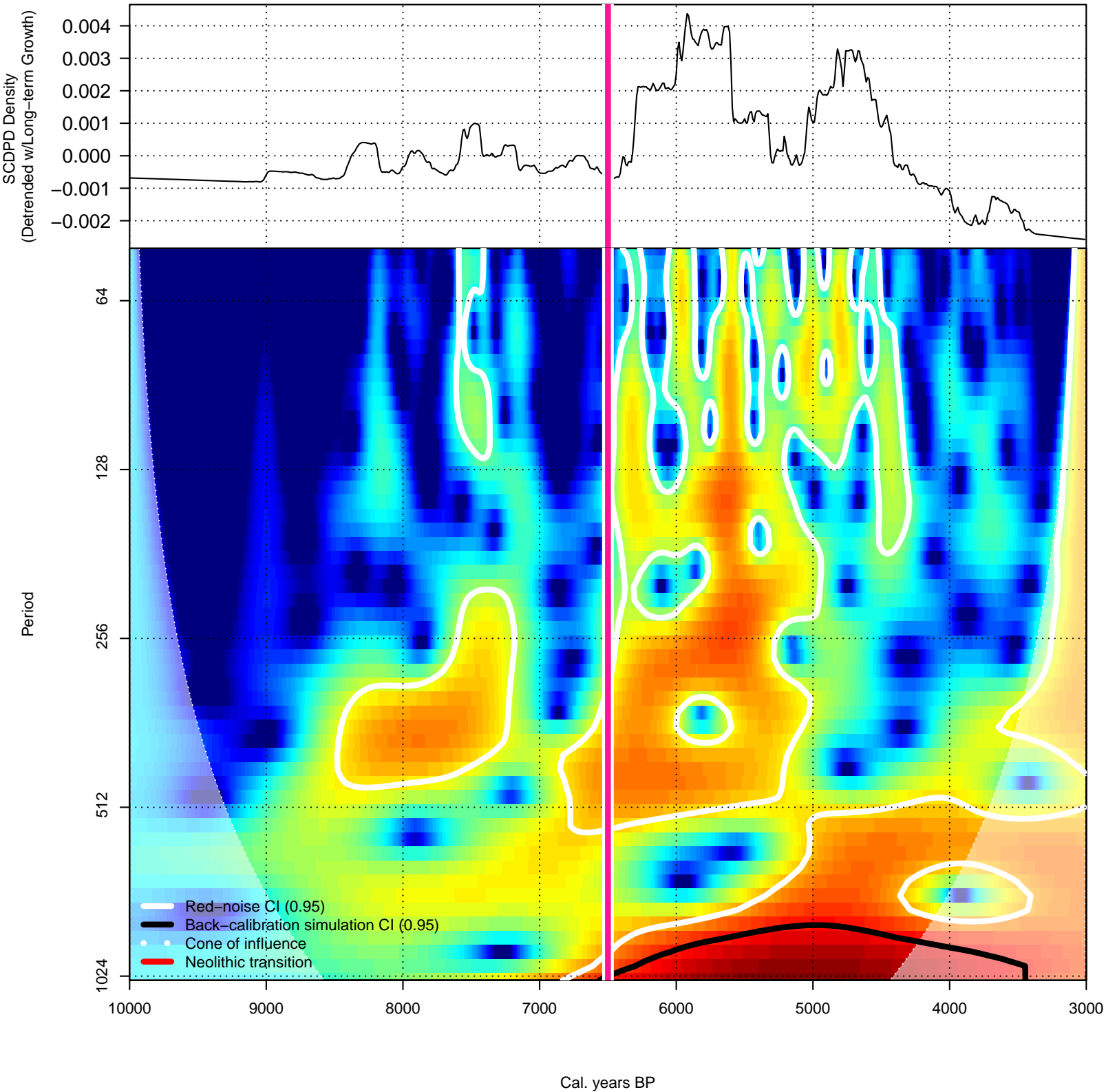


Wavelet analysis for all regions.

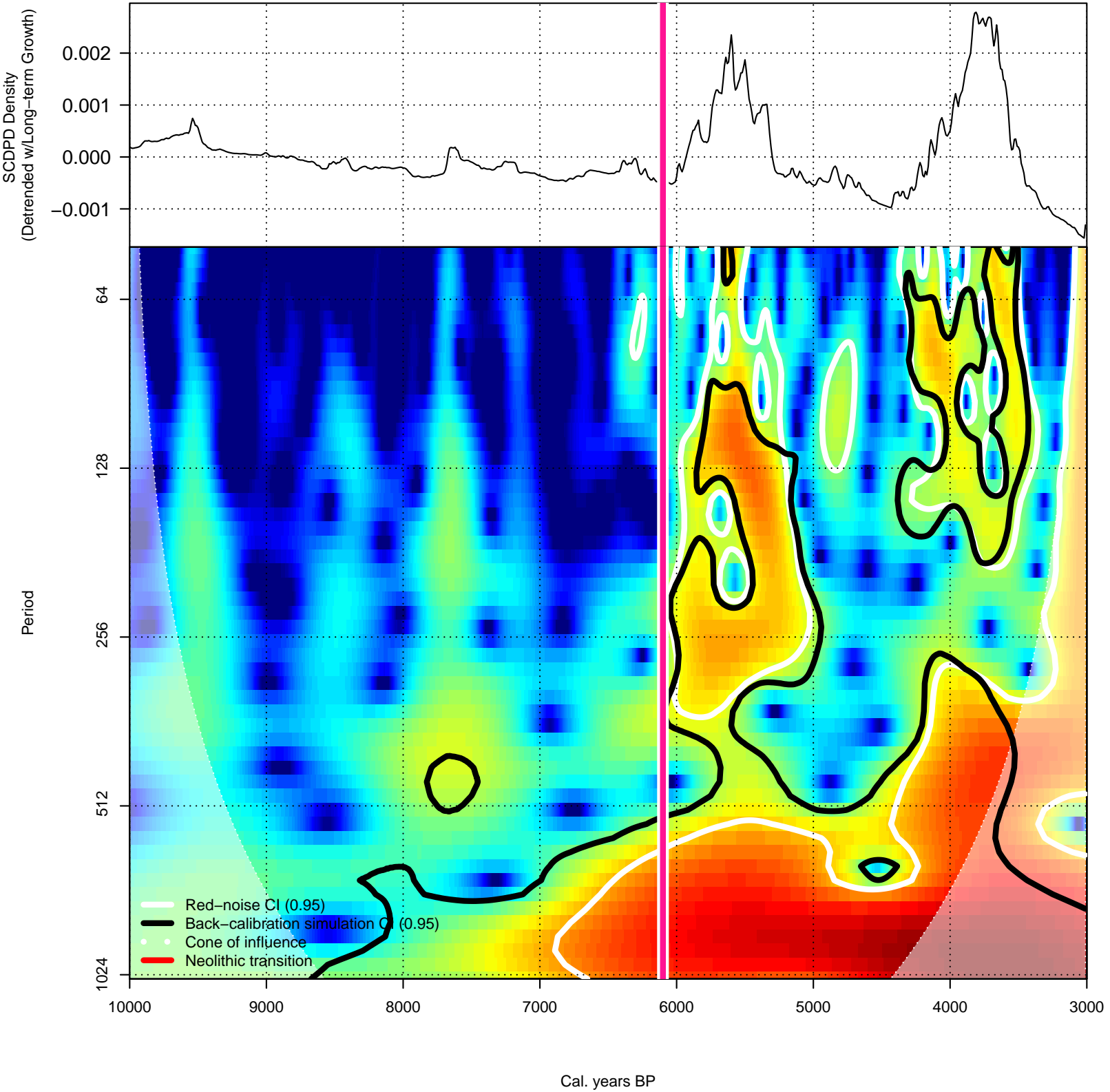
Southern Germany



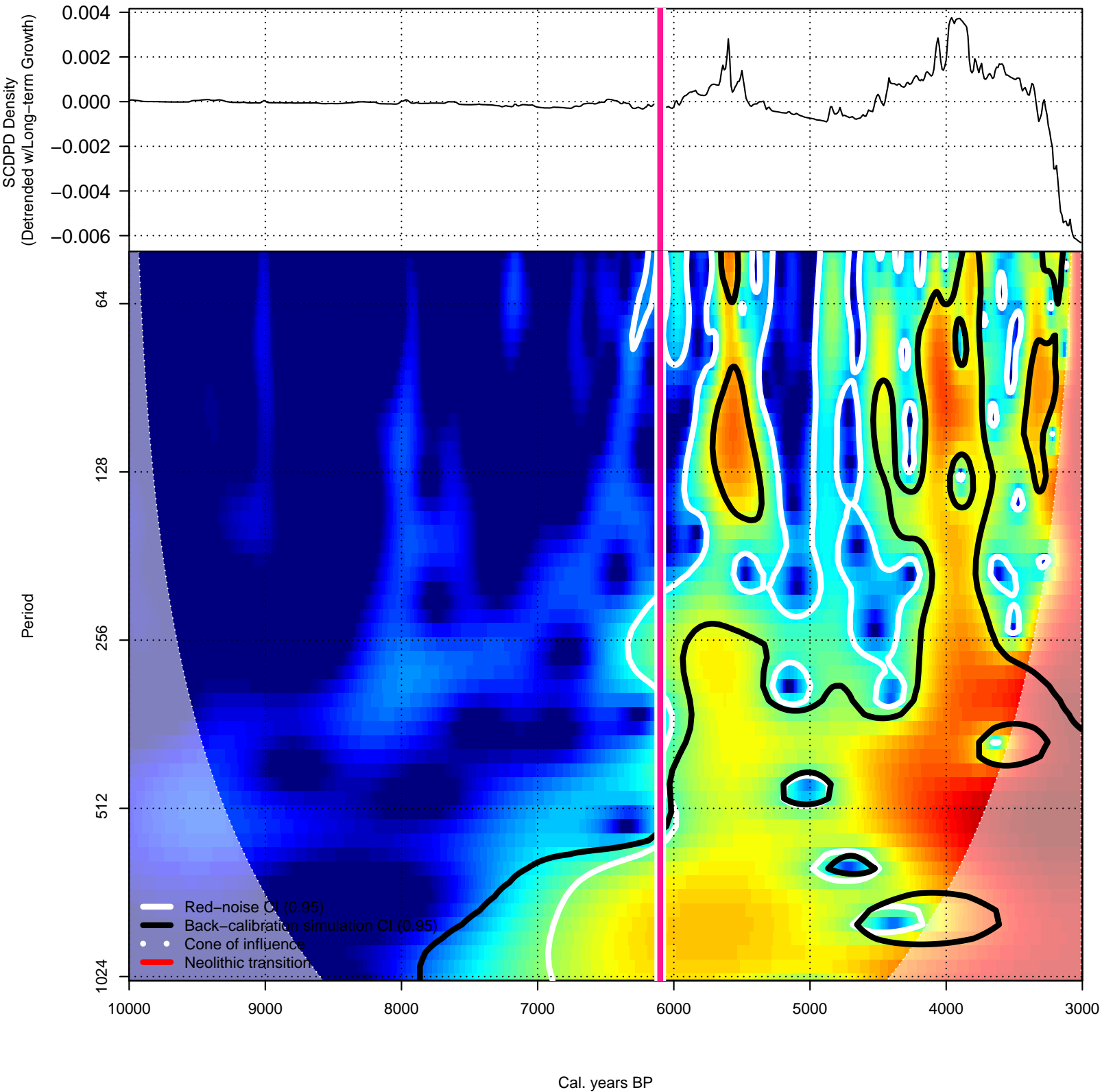
Eastern Switzerland



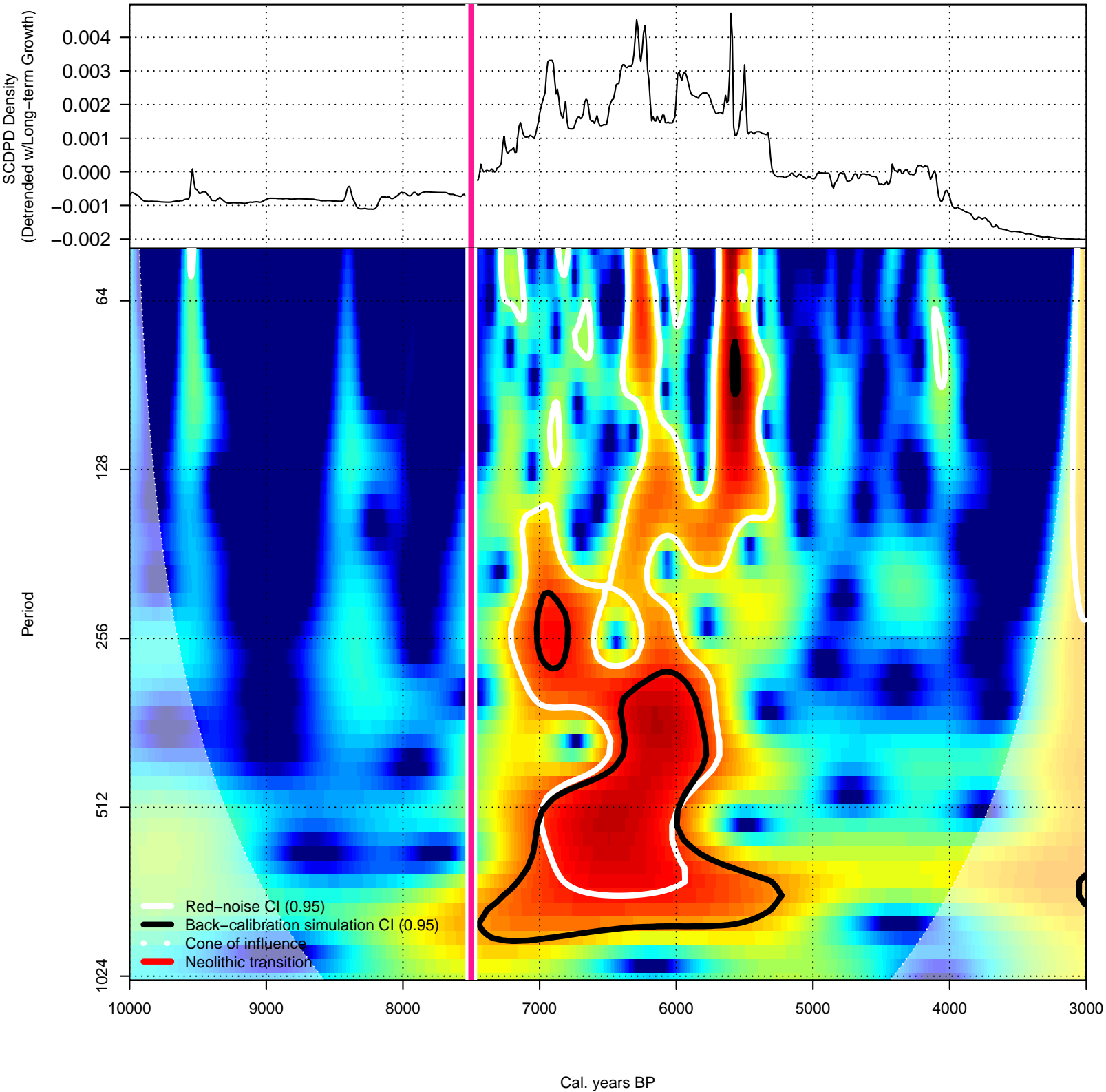
England and Wales (w/o Wessex & Sussex)



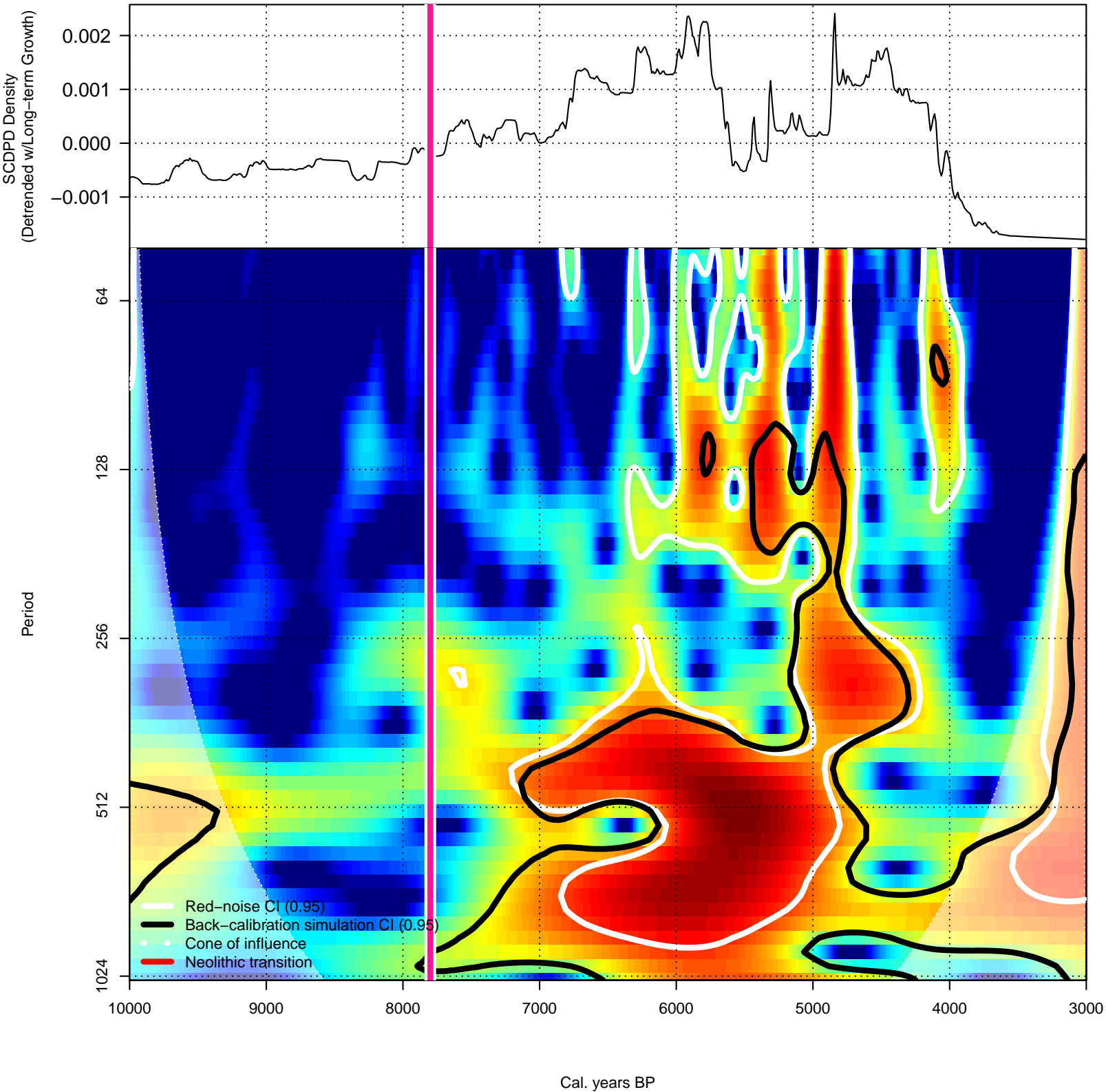
Ireland



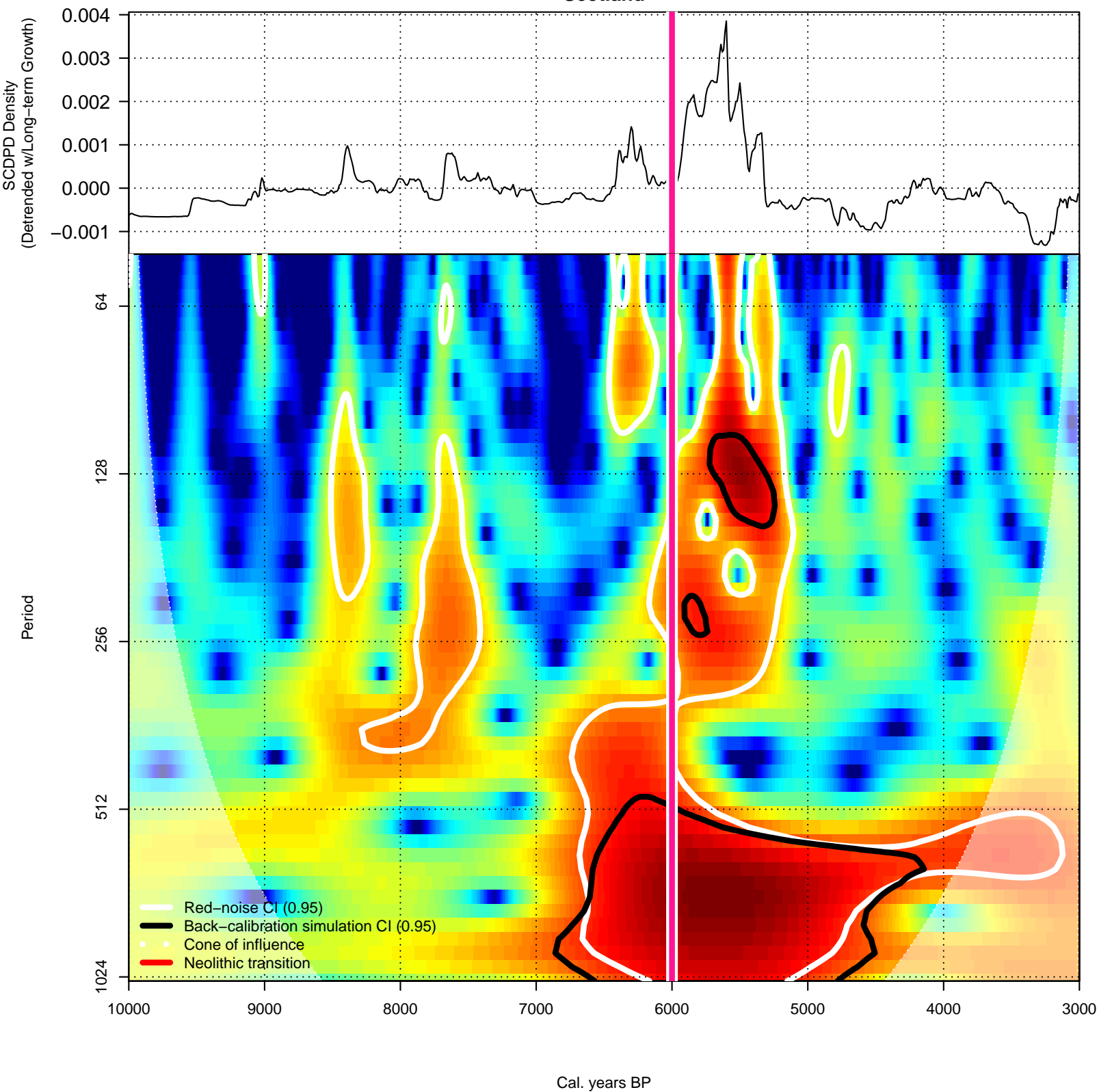
Paris Basin



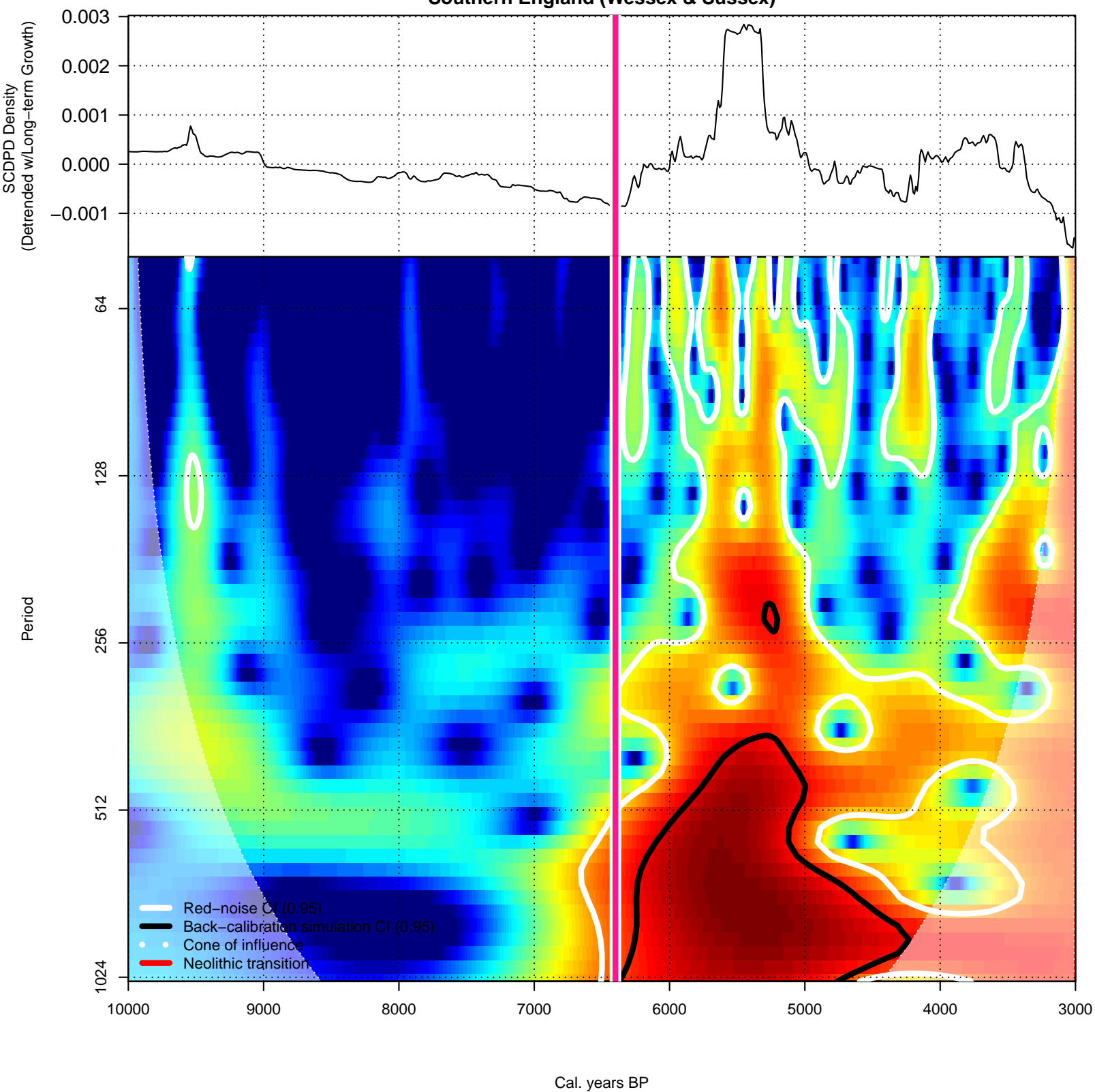
Rhone-Languedoc



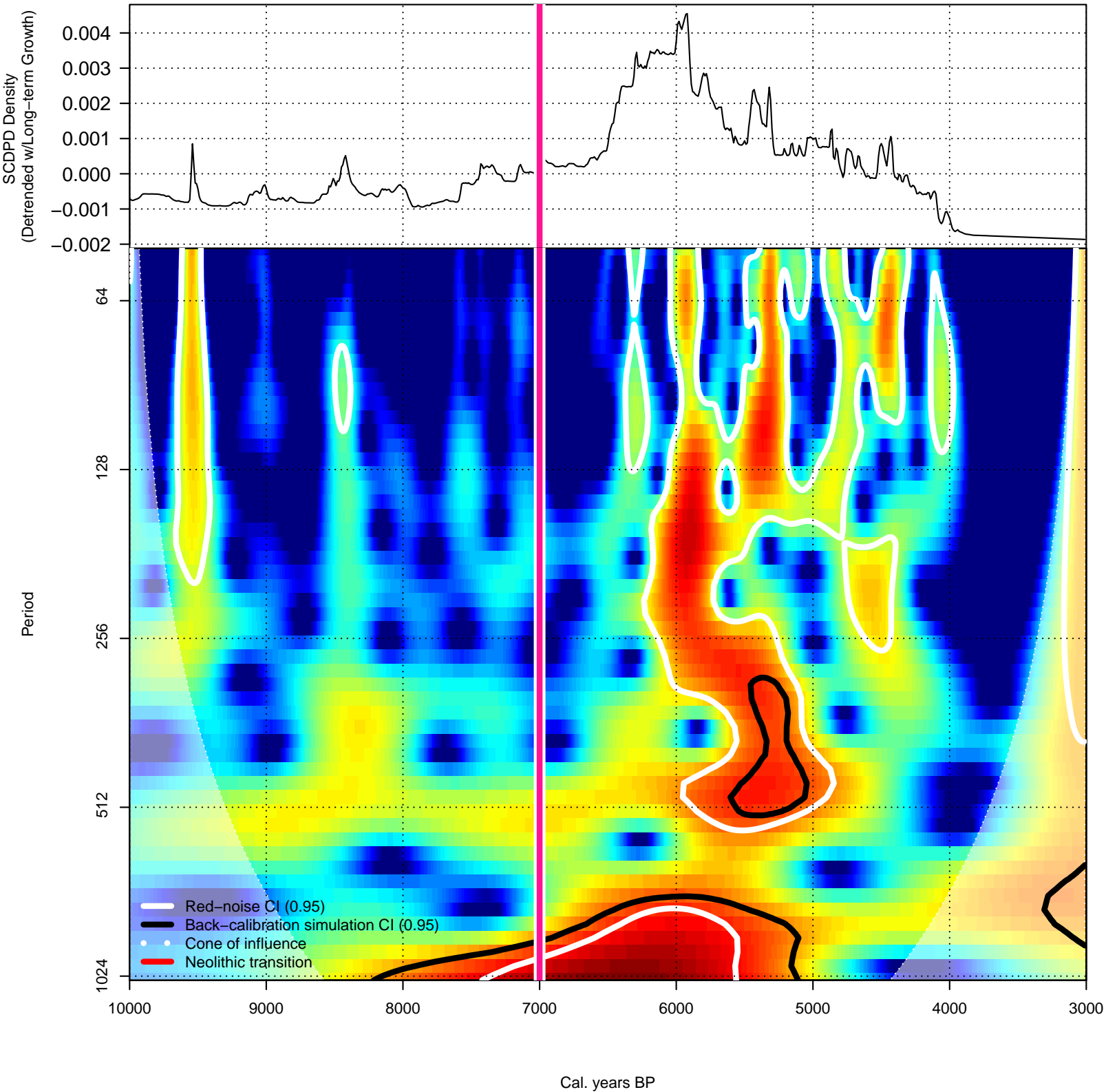
Scotland



Southern England (Wessex & Sussex)

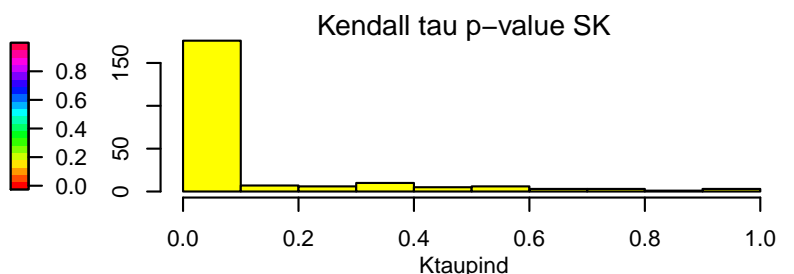
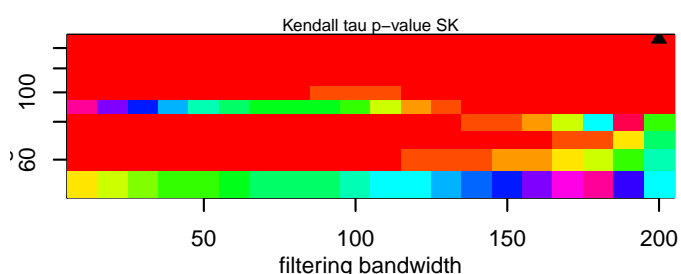
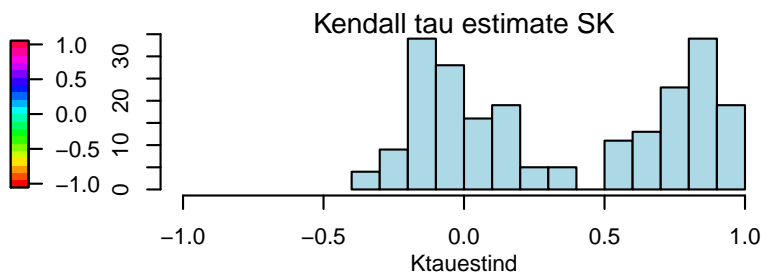
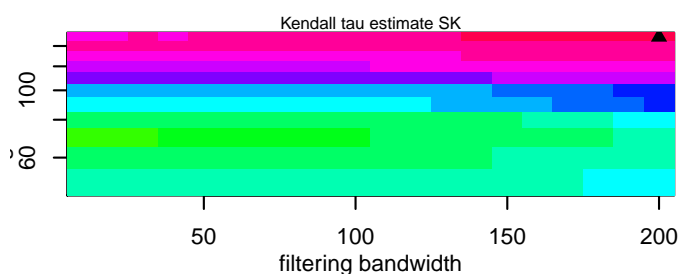
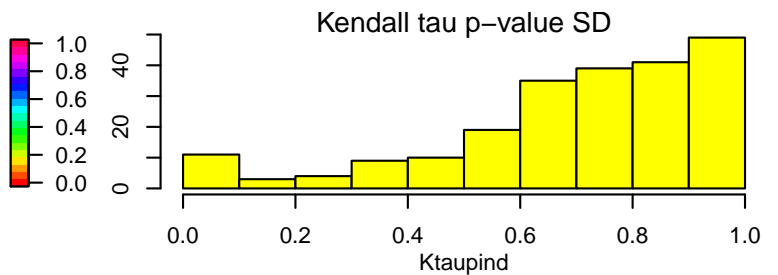
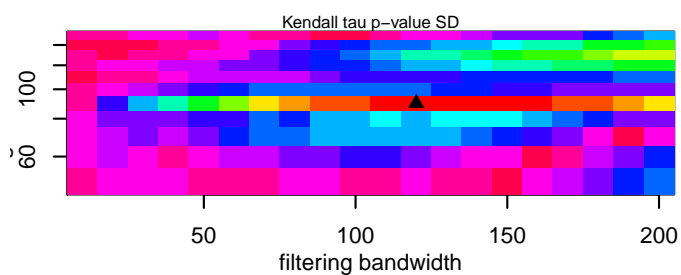
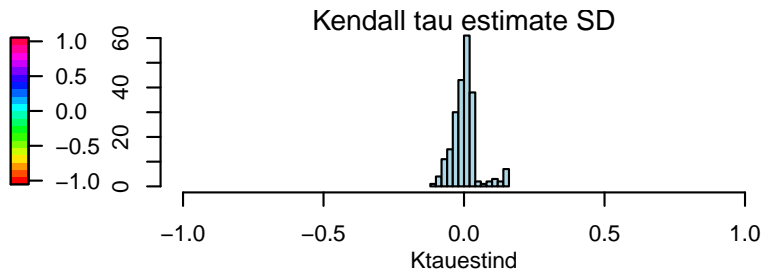
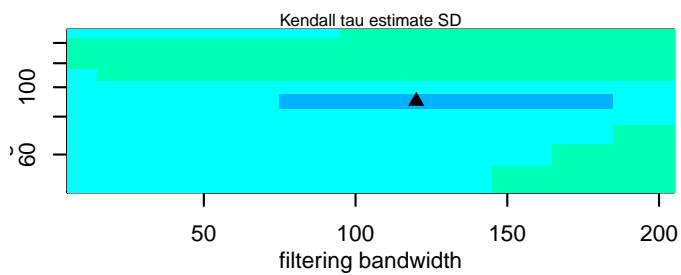
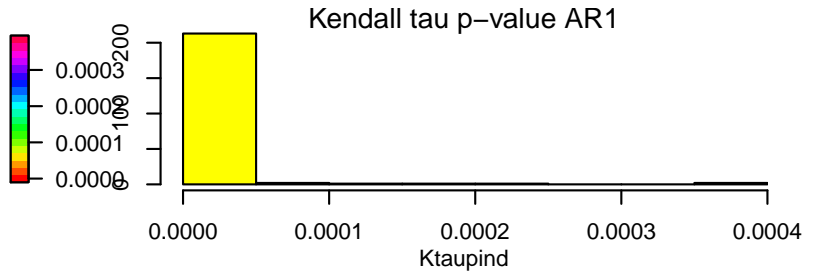
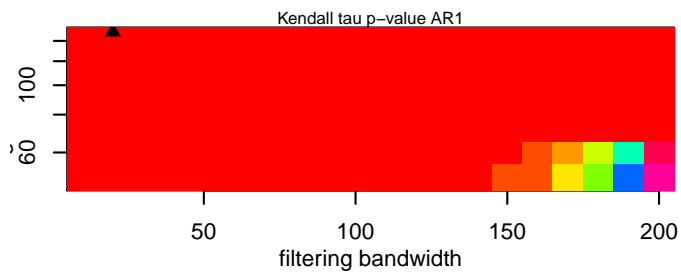
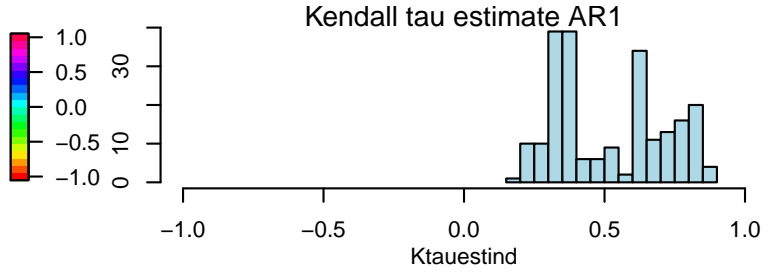
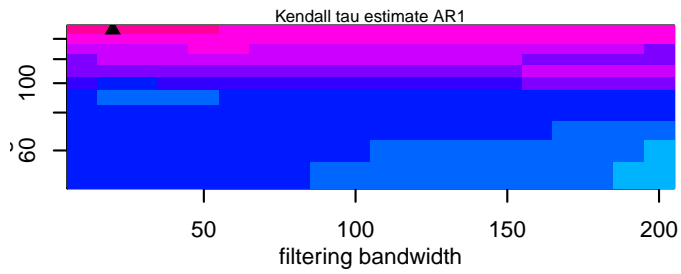


Western France

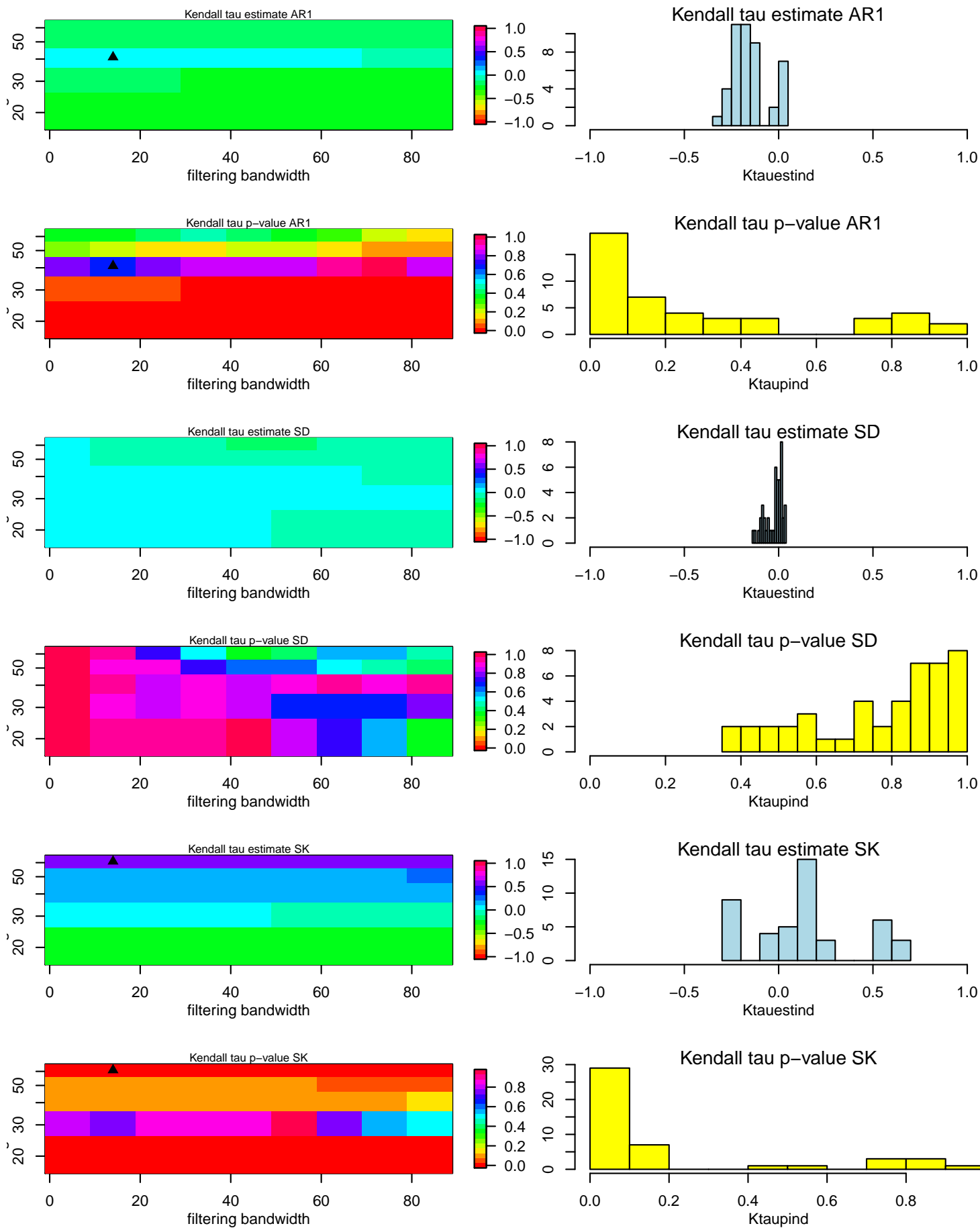


EWS sensitivity analysis for all regions.

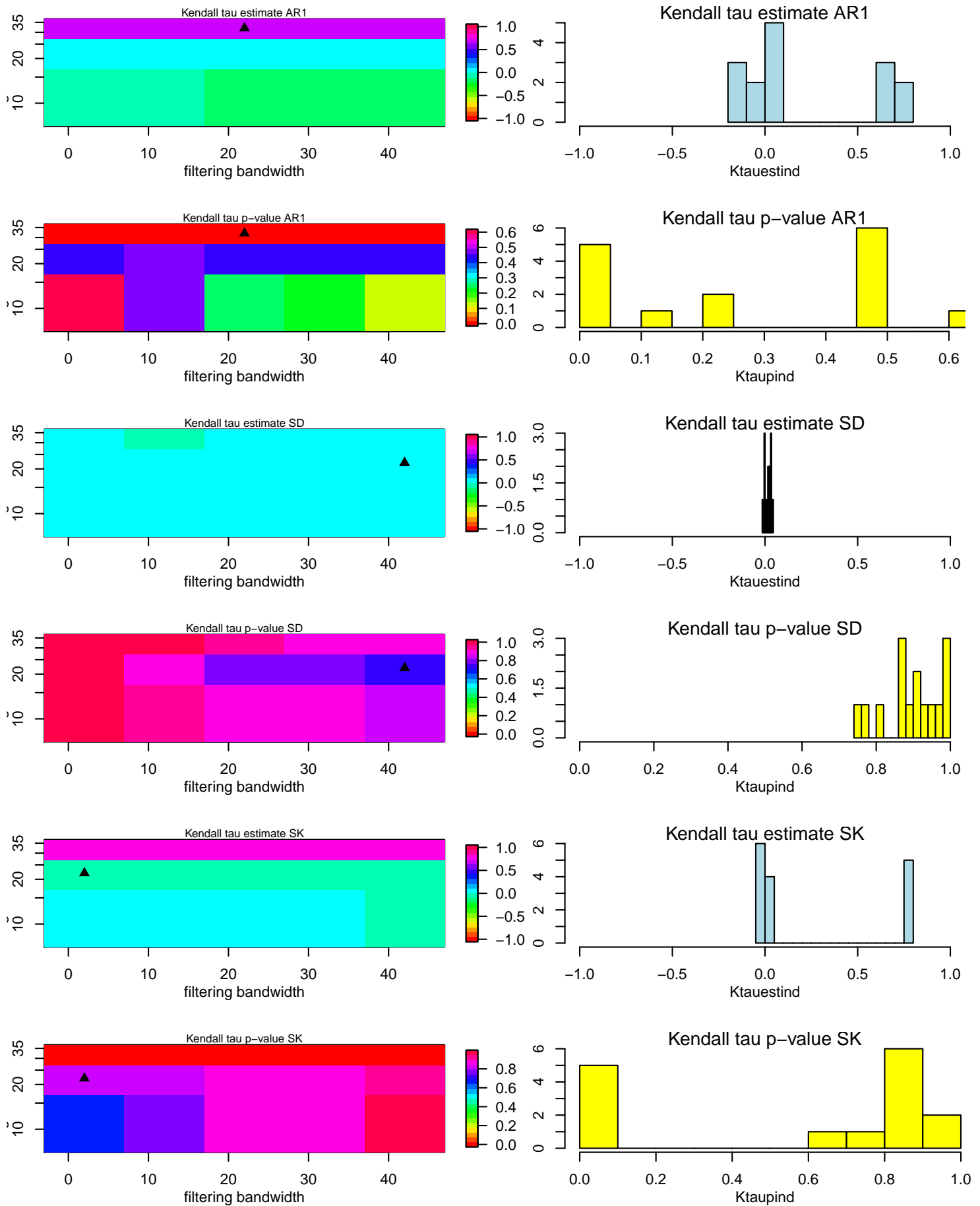
Southern Germany



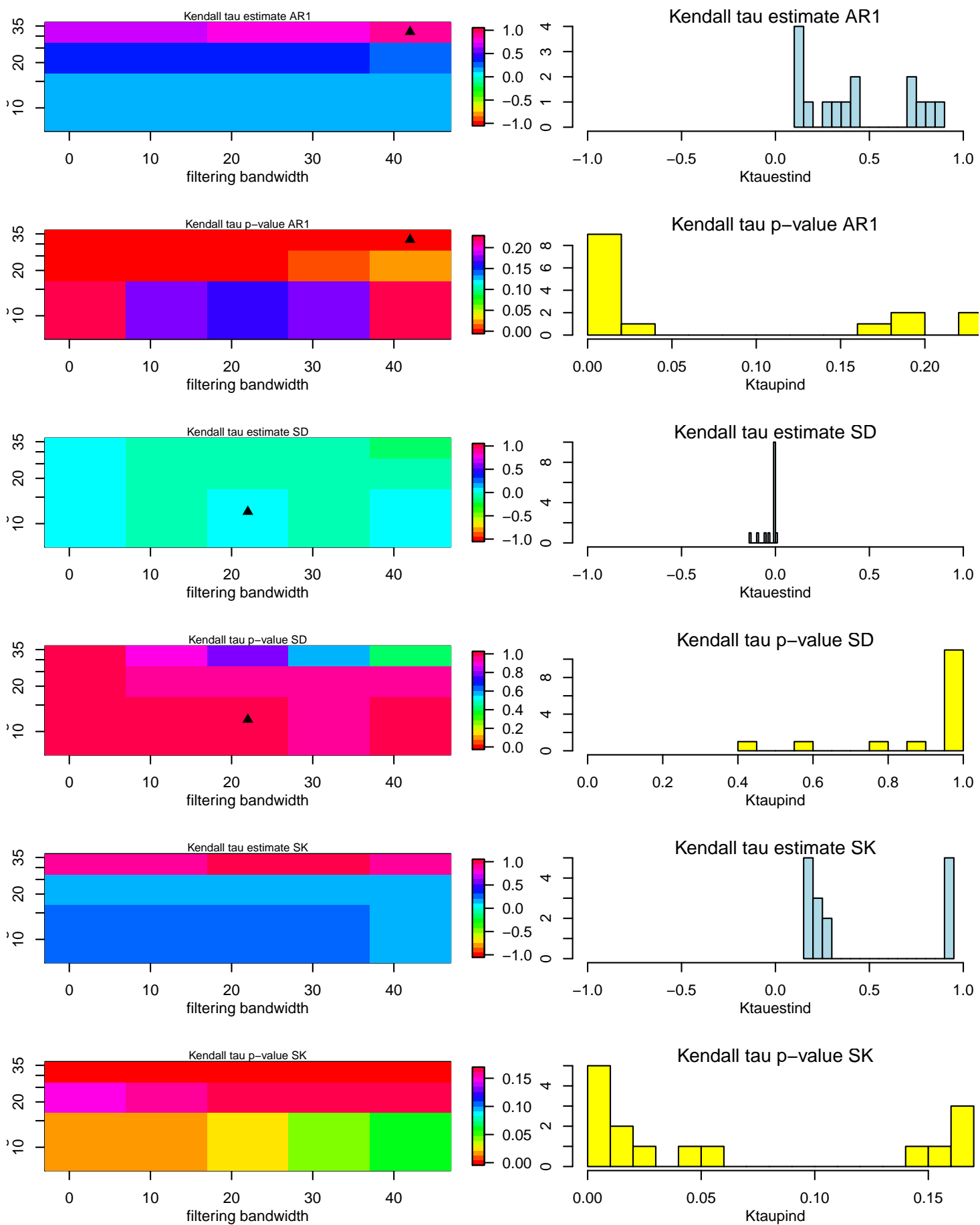
Eastern Switzerland



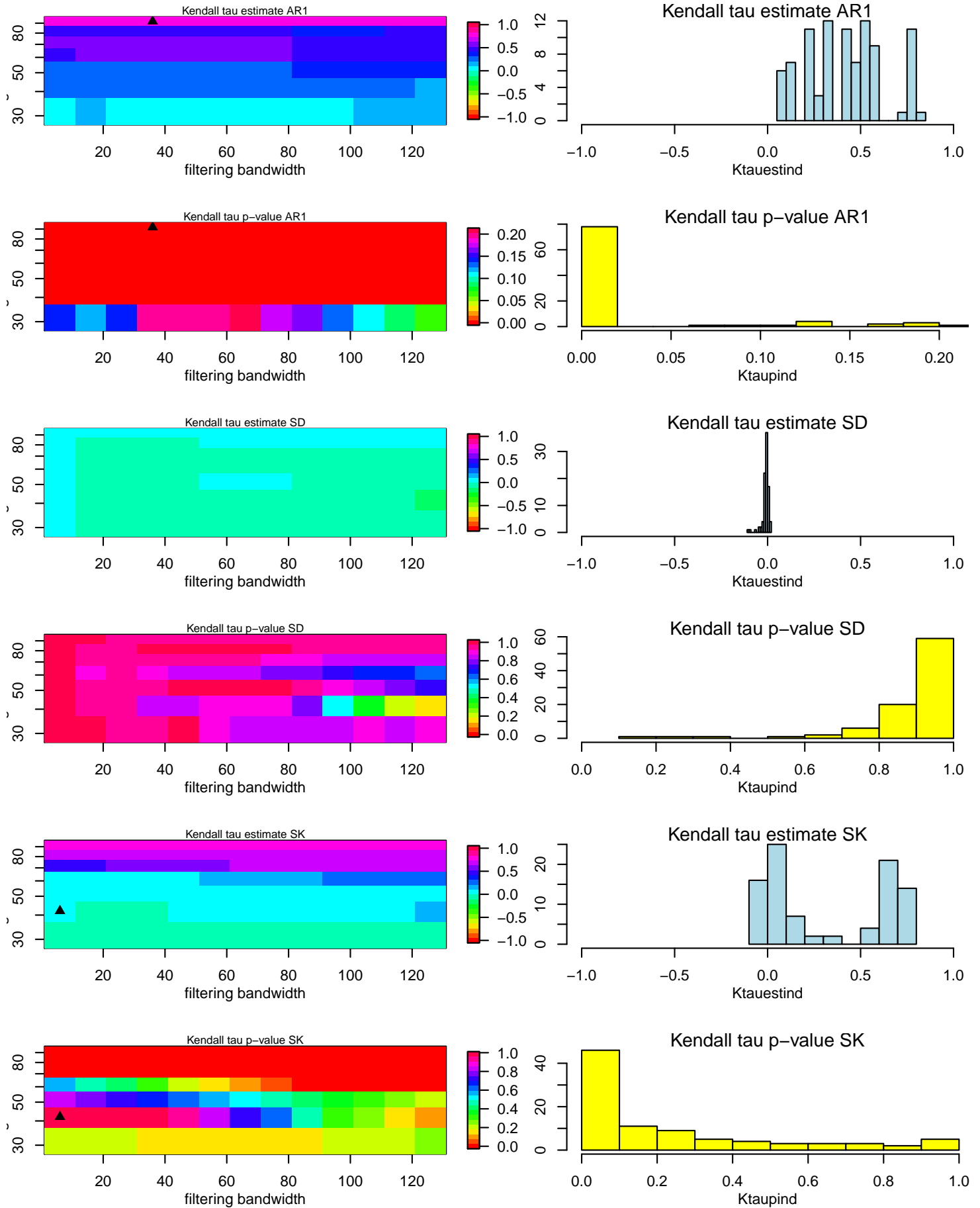
England and Wales (w/o Wessex & Sussex)



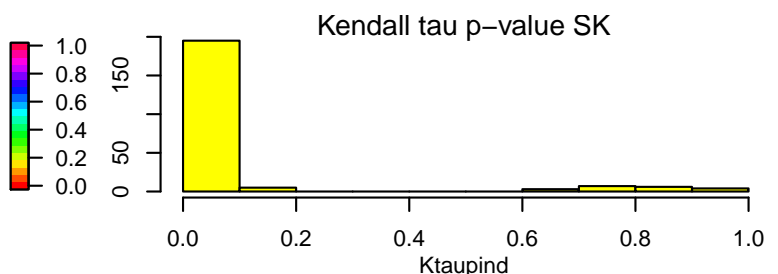
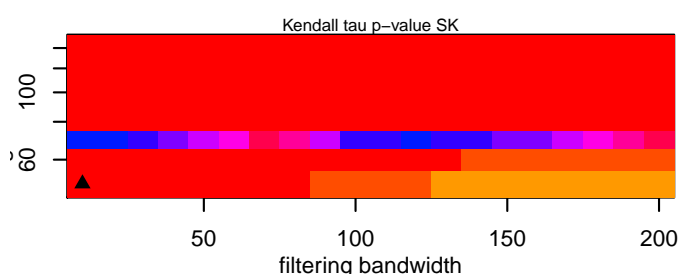
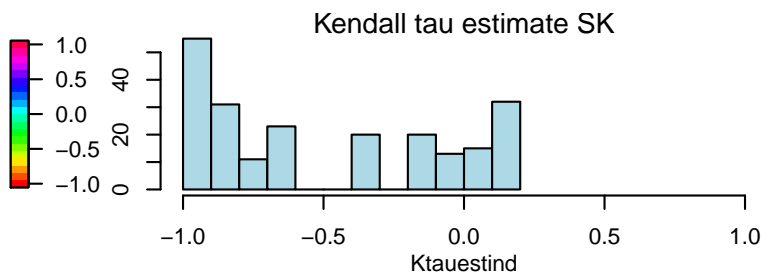
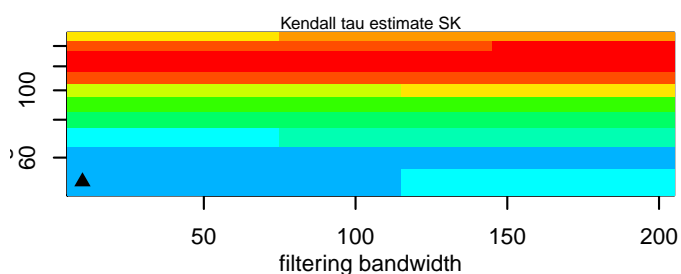
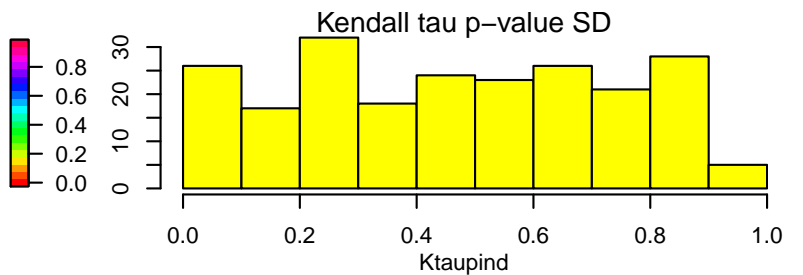
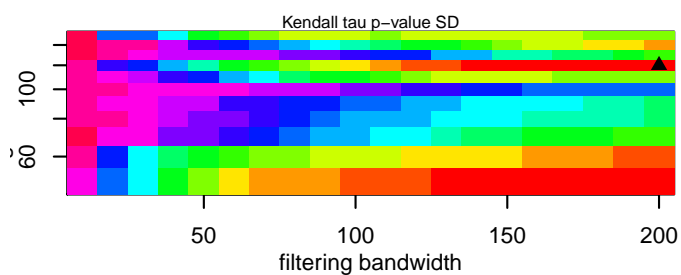
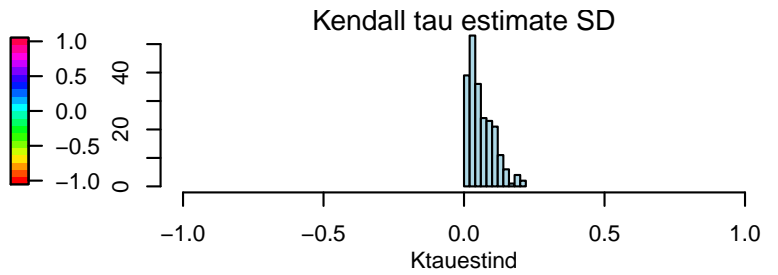
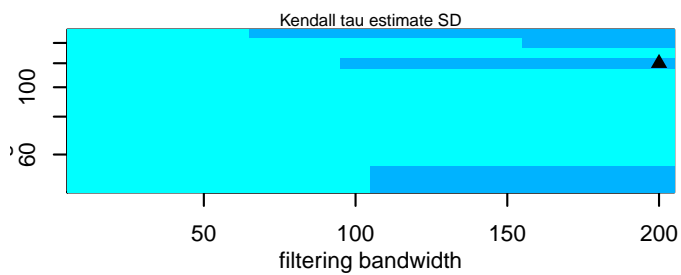
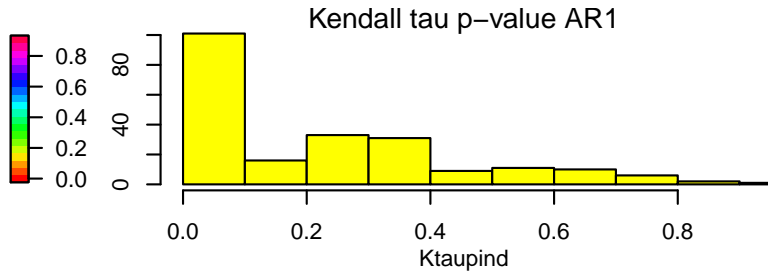
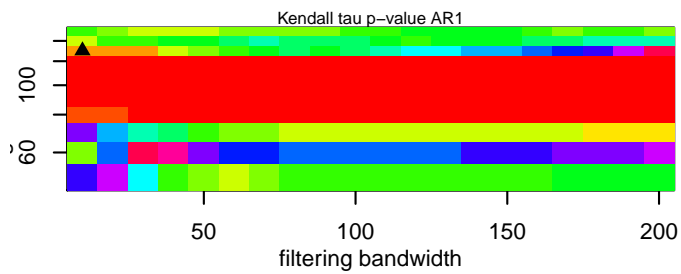
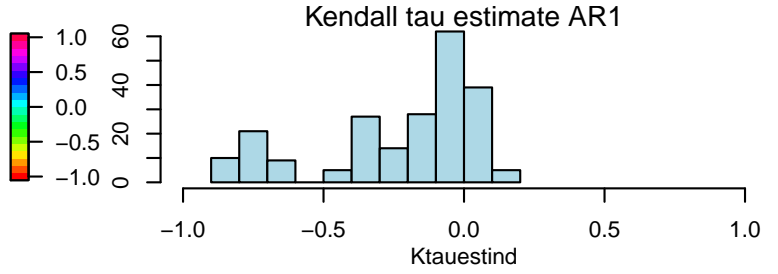
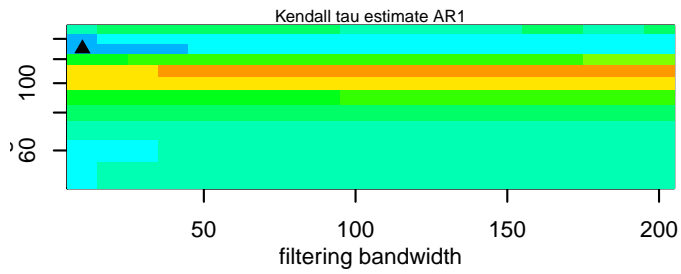
Ireland



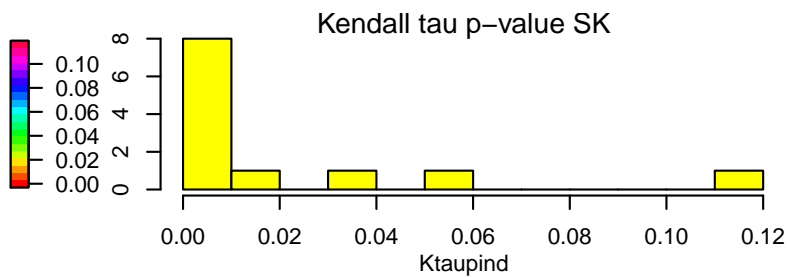
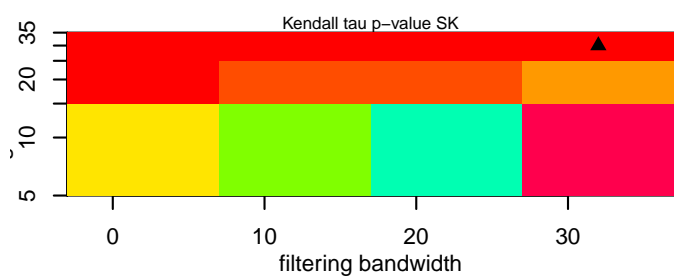
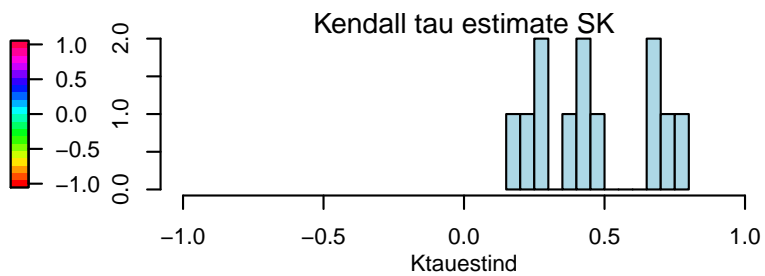
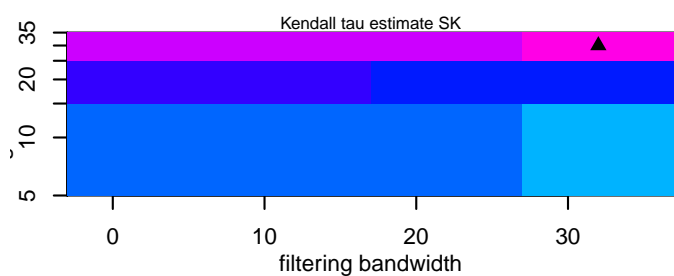
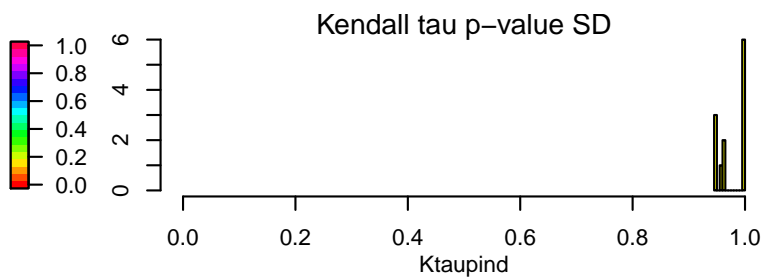
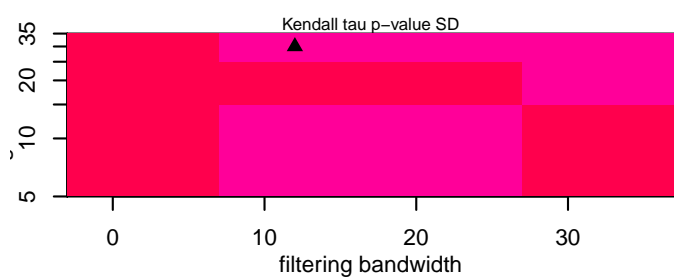
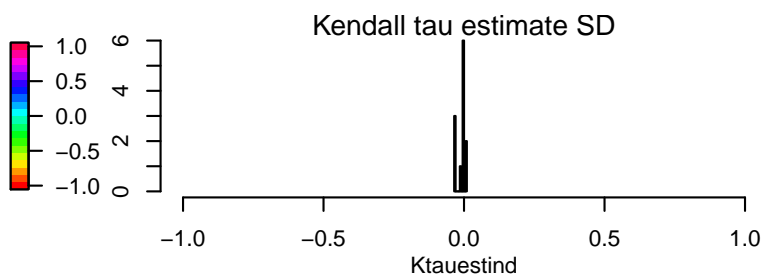
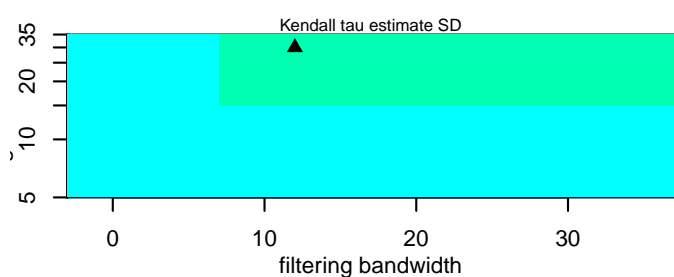
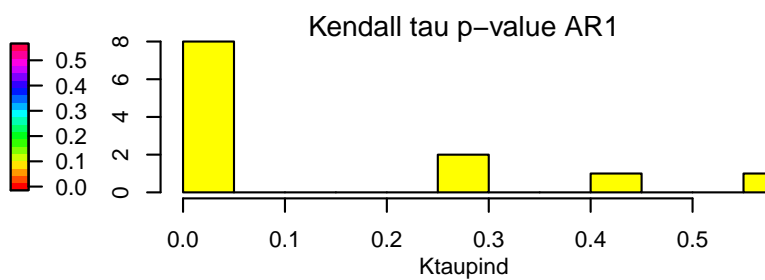
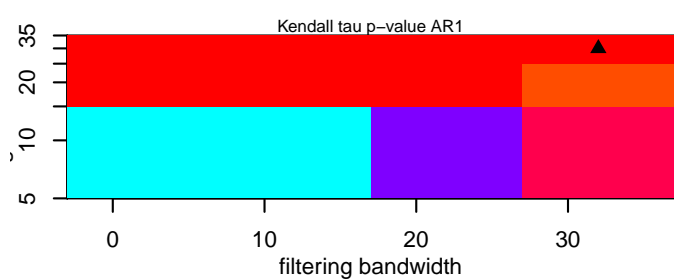
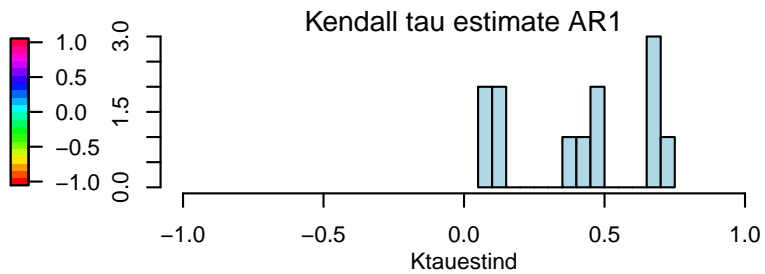
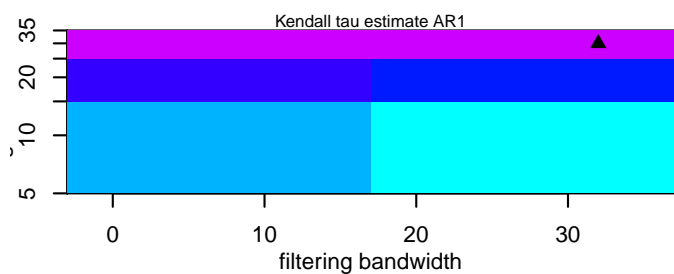
Paris Basin



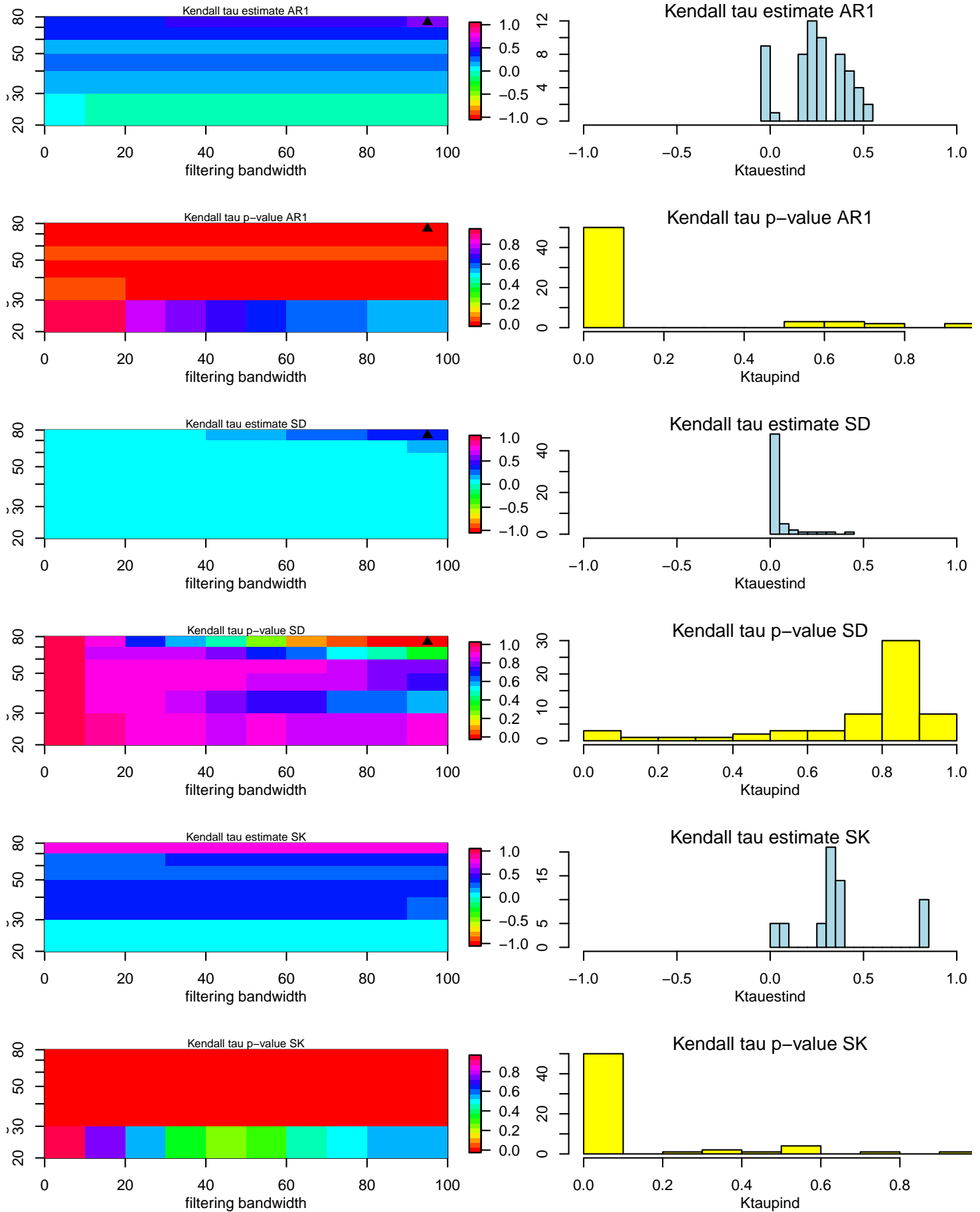
Rhone-Languedoc



Scotland



Southern England (Wessex & Sussex)



Western France

