Environmental and Ecological Statistics (2022) 29:779–799
https://doi.org/10.1007/s10651-022-00547-2

Fast estimation and choice of confidence interval methods for step regression

Shuangcheng Hua[1] · **Youyi Fong**[1] · **Jarrod Kath**[2]

Received: 9 August 2022 / Accepted: 26 August 2022 / Published online: 6 September 2022

Abstract

In this paper we propose a new fast grid search algorithm for finding the least square estimators of a step regression model. This algorithm makes it practical to compute resampling-based confidence intervals for step regression models. We introduce five data generating models, including one where the mean model is a step model (model correctly specified) and four where the mean models are not step models (model misspecified), and use them to study the coverage probabilities of two new types of resampling-based confidence intervals for step regression: symmetric percentile bootstrap confidence intervals and subsampling confidence intervals using a new set of rules-of-thumb to select block size. Our results show that when the model is correctly specified, the symmetric percentile Efron bootstrap confidence intervals provide close-to-nominal coverage and have shorter intervals than the subsampling methods; when the model is misspecified, the subsampling method using the rules-of-thumb provides good coverage and shorter confidence intervals than the symmetric percentile Efron bootstrap method and the subsampling method using a double bootstrap-like procedure for block size selection. Finally, we apply the proposed methods to a real world environmental dataset on the relationship between grassland productivity, soil moisture anomalies and other hydro-climatic and land use variables to provide inference for the threshold in soil moisture anomalies, across which there is a jump in grassland productivity.

Keywords Dynamic programming · Ecology · Subsampling · Symmetric percentile bootstrap confidence intervals

Handling Editor: Luiz Duczmal

✉ Youyi Fong
youyifong@gmail.com

[1] Department of Biostatistics, University of Washington, Seattle, WA, USA

[2] University of Southern Queensland, Darling Heights, QLD, USA

1 Introduction

Step regression, also known as discontinuous threshold regression, is a type of non-regular regression model where the mean of an outcome variable is a step function of the covariate of interest, e.g.

$$Y = \alpha + \alpha_z^T z + \beta I(x > e) + \epsilon,$$

where Y is the response, x is the predictor with threshold effect, z is a covariate vector with $p-1$ dimensions, and ϵ denotes an error term with bounded variance that is independent between observations, e is the threshold parameter, β denotes the size of the jump at the threshold, α is the intercept, and α_z denotes the coefficients corresponding to z. This model is often used in practice because it provides a simple but elegant and interpretable way to model certain kinds of threshold-dependent relationships between an outcome and a predictor. Despite the simplicity of step regression models, the presence of threshold parameters changes both estimation and inference for the models in profound ways. For example, it is well understood that the asymptotic distribution of the threshold parameter is nonstandard: while the least square estimator of the change point or threshold converges at a rate of n^1 when both the data generating model and the working model are step linear regression models, it converges at a rate of $n^{1/3}$ when the true underlying model does not actually follow a step regression model (e.g. Bühlmann and Yu 2002; Pons 2003; Banerjee and McKeague 2007; Kosorok 2008; Song et al. 2016).

The use of step regression in data analysis is currently limited by two factors: (i) Existing methods for step regression model estimation either produce locally optimal instead of globally optimal solutions or are too slow. (ii) The performance of resampling-based confidence interval methods depends critically on block size selection, but there is a dearth of studies on how to select block sizes when the data generating model and the working model do not match. Here we address these two limitations. In Sect. 2 we propose a fast grid search algorithm for step linear regression that seeks globally optimal solutions and apply the method to empirically study the convergence rates under five different simulation scenarios. In Sect. 3, we study symmetric percentile Efron bootstrap confidence intervals and subsampling confidence intervals both when the data generating model and the working model match and when they do not match. We further propose new rules for selecting block size in subsampling that compare favorably with the existing double bootstrap-like procedure. In Sect. 4, we apply the proposed methods to a real data example from a study of hydro-climatic drivers of agricultural grassland productivity under extreme drought and rainfall. We end with a discussion in Sect. 5.

2 Fast grid search-based estimation of step linear regression models

Due to the threshold parameter, the likelihood function of a step threshold model is non-smooth and non-convex. A natural approach to maximize the likelihood is to approximate the step function in the likelihood by a smooth function (e.g. Gallant and

Fuller 1973; Tishler and Zang 1981; Pastor and Guallar 1998; Muggeo 2003; Fong et al. 2017). However, as shown in Fong (2019), such approximation may lead to inaccurate coverage of bootstrap confidence intervals in threshold linear regressions because the criterion functions are still non-convex and the approach finds locally optimal solutions. Conceptually, the grid search approach (Friedman and Silverman 1989) is a two-stage process: in the first stage we optimize a series of submodels conditional on a grid of candidate threshold values (typically the realized covariate values in the dataset); in the second stage we simply find the maximum in the set of likelihoods computed in the first stage and take the corresponding threshold value as the estimated threshold. Since conditioning on the threshold values removes both non-smoothness and non-convexity associated with the presence of the threshold parameter, the grid search approach yields globally optimal solutions. The main disadvantage of the grid search approach is its computational burden.

We propose a fast grid search-based estimation method for step linear regression models. The method uses the same strategy from Elder and Fong (2019) and Son and Fong (2020), which deals with continuous two-phase regression models. We sketch the outline of the algorithm here; further details can be found in Section B of the Supplementary Materials. Consider the model given in the introduction, the least squares estimator for the model parameters minimizes the sum of squares of the residuals, which equals $[(I - H_e)Y]^T[(I - H_e)Y] = Y^TY - Y^TH_eY$, where $\boldsymbol{H_e} \equiv \boldsymbol{X}_e(\boldsymbol{X}_e^T\boldsymbol{X}_e)^{-1}\boldsymbol{X}_e^T$ is the hat matrix, and $\boldsymbol{X}_e \equiv [\mathbf{1}, \boldsymbol{Z}, \boldsymbol{v}_e] \equiv [\boldsymbol{X}, \boldsymbol{v}_e]$ is the design matrix, where $\mathbf{1}$ is a vector of ones, $\boldsymbol{Z}$ is a $n \times (p-1)$ matrix, $\boldsymbol{X}$ is a $n \times p$ matrix, and $\boldsymbol{v}_e \equiv \boldsymbol{I}(\boldsymbol{x} > e)$ is a n-dimensional vector which equals 1 when $x_i > e$ and 0 otherwise. Thus it is equivalent to maximizing $\boldsymbol{Y}^T\boldsymbol{H_eY}$ with respect to e.

The key factor for accelerating computation is to avoid computing $\boldsymbol{Y}^T\boldsymbol{H_eY}$ *de novo* for every candidate e. We achieve this by breaking down $\boldsymbol{Y}^T\boldsymbol{H_eY}$ into components and deriving a recursive relationship for each component between successive e's. We can write

$$\boldsymbol{Y}^T\boldsymbol{H_eY} = \boldsymbol{Y}^T\boldsymbol{HY} + \left(\boldsymbol{v}_e^T\boldsymbol{r}\right)^2/(\boldsymbol{v}_e^T\boldsymbol{v}_e - \boldsymbol{v}_e^T\boldsymbol{Hv}_e), \tag{1}$$

where $\boldsymbol{H} \equiv \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$ and $\boldsymbol{r} \equiv (\boldsymbol{H} - \boldsymbol{I})\boldsymbol{Y}$. By the QR decomposition, we can further write $\boldsymbol{v}_e^T\boldsymbol{Hv}_e = (\boldsymbol{Q}_1^T\boldsymbol{v}_e)^T\boldsymbol{Q}_1^T\boldsymbol{v}_e$, where $\boldsymbol{Q}_1$ is the first p columns of the orthogonal matrix Q from the QR decomposition.

Consider two successive (different) values of e: e_t and e_{t+1}. Suppose e_t and e_{t+1} correspond to the t^{th} and $t+1^{th}$ ascending ordered values of x. Let $\boldsymbol{v}_{e_{t+1}} = \boldsymbol{v}_{e_t} - \boldsymbol{\delta}_t$. We find that:

$$\boldsymbol{v}_{e_{t+1}}^T\boldsymbol{v}_{e_{t+1}} = \boldsymbol{v}_{e_t}^T\boldsymbol{v}_{e_t} - 1 \tag{2}$$

$$\boldsymbol{v}_{e_{t+1}}^T\boldsymbol{Q}_1 = \boldsymbol{v}_{e_t}^T\boldsymbol{Q}_1 - \boldsymbol{\delta}_t^T\boldsymbol{Q}_1 \tag{3}$$

$$\boldsymbol{v}_{e_{t+1}}^T\boldsymbol{r} = \boldsymbol{v}_{e_t}^T\boldsymbol{r} - \boldsymbol{\delta}_t^T\boldsymbol{r} \tag{4}$$

Here we use the fact that $\boldsymbol{\delta}_t$ is a vector of size n with the $(t+1)^{th}$ entry equal to 1 and 0 everywhere else. $\boldsymbol{\delta}_t^T\boldsymbol{Q}_1$ is the $t+1^{th}$ row vector of $\boldsymbol{Q}_1$, and $\boldsymbol{\delta}_t^T\boldsymbol{r}$ is the $t+1^{th}$

Table 1 Run time (sec) for fitting step linear regression models on a single Intel(R) Xeon(R) CPU E5-2667 v4 @ 3.20GHz, averaged over 20 Monte Carlo datasets

	Grid search	Fast grid search
$n = 10^3$, $p = 2$	2604	0.40
$n = 10^4$, $p = 2$	–	3.80
$n = 10^5$, $p = 2$	–	42.7
$n = 10^6$, $p = 2$	–	465
$n = 10^6$, $p = 10$	–	1223

element of $\boldsymbol{r}$. Hence, these update steps can be done very quickly. The full algorithm is described more formally below. The results of some benchmarking experiments are shown in Table 1.

Algorithm 1 ***Fast grid search algorithm for step linear regression model***

1. *Sort the samples by the ascending order of* x_i
2. *Compute and store the initial values:* $\boldsymbol{v}_{e_1}^T \boldsymbol{v}_{e_1}$, $\boldsymbol{v}_{e_1}^T \boldsymbol{Q}_1$, $\boldsymbol{v}_{e_1}^T \boldsymbol{r}$
3. *Compute and store the initial value of* $\boldsymbol{Y}^T \boldsymbol{H}_{e_1} \boldsymbol{Y}$
4. *For* t *in* 1 *to* $n - 1$*:*
   - *update* $\boldsymbol{v}_{e_{t+1}}^T \boldsymbol{v}_{e_{t+1}}$, $\boldsymbol{v}_{e_{t+1}}^T \boldsymbol{Q}_1$, $\boldsymbol{v}_{e_{t+1}}^T \boldsymbol{r}$ *based on* (2), (3), *and* (4)
   - *update* $\boldsymbol{Y}^T \boldsymbol{H}_{\boldsymbol{e_{t+1}}} \boldsymbol{Y}$ *based on* (1)
   - *update* $\hat{e} = e_{t+1}$ *if* $\boldsymbol{Y}^T \boldsymbol{H}_{\boldsymbol{e_{t+1}}} \boldsymbol{Y} > \boldsymbol{Y}^T \boldsymbol{H}_{\boldsymbol{e_t}} \boldsymbol{Y}$

With the fast grid search algorithm, we can more easily study the convergence rates of the parameter estimates empirically. To compare these converge rates to the asymptotic rates, we use five data generating models: (i) Step: The mean of Y is a step function of x; (ii) Sig_γ for $\gamma \in \{1, 5, 15\}$: The mean of Y is a sigmoid function of x. Sig_15 more closely resembles a step function than Sig_1; (iii) Quad: The mean of Y is a quadratic function of x:

$$E(Y|Z, X) = \alpha + \alpha_z Z + \beta I(X > e) \tag{Step}$$

$$E(Y|Z, X) = \alpha + \alpha_z Z + \beta \frac{\gamma e^{(X-4.7)}}{1 + \gamma e^{(X-4.7)}} \tag{Sig}$$

$$E(Y|Z, X) = \alpha + \alpha_z Z + \gamma_1 X + \gamma_2 X^2 \tag{Quad}$$

The model parameters and the covariate distributions are given in Section A of the Supplementary Materials. The step linear regression model is correctly specified under the Step model and misspecified under the Sig_γ models and the Quad model. Under the Sig_γ models, the step linear regression model can be seen as an approximation of the data generating models because they have the same overall shape; under the Quad model, the interpretation of the step model fit needs to be more carefully considered,

e.g. the limits of the step linear regression model parameters depend heavily on the distribution of X. To obtain the limits of the step linear regression model parameters when the model is misspecified or when the limits cannot be inferred from symmetry, we fit the step linear regression model $Y = \alpha + \alpha_z Z + \beta I(X > e)$ to datasets with sample size $n = 10^6$ and take the average over ten Monte Carlo replicates, the results of which are listed in Tables A.1–A.4 of the Supplementary Materials.

To study the empirical rate of convergence over a spectrum of sample sizes, we consider two sets of sample sizes. The first set has sample sizes of 500, 1000, 1500, and 2000, and the second set has sample sizes of 1024,000, 2048,000, 4096,000, and 8192,000. At each sample size, we estimate the variability of the estimator by conducting 10,000 Monte Carlo runs and computing the standard deviation of the parameter estimates across the Monte Carlo replicates, which we denote by σ_n. To estimate the rate of convergence, we fit the model $\sigma_n = a \times n^b$ by fitting a straight line through $(n, log(\sigma_n))$ so that the slope $\hat{b}$ of the fitted line gives an estimated convergence rate.

The results for the two sets of sample sizes are summarized in Tables 2 and 3, respectively. When the model is correctly specified, as shown in both Tables 2 and 3, the estimated convergence rates are close to the asymptotic rates: n for the threshold parameter e and $n^{1/2}$ for the slope parameters. When models are misspecified, the results are complex. We discuss the results one parameter at a time. In larger datasets, $\hat{e}$ is $n^{1/3}$-convergent. In smaller datasets, the convergence rate is between $n^{1/2}$ and $n^{1/3}$, and among the three sigmoid models, the closer the model is to the step model, the faster it converges.

The estimates of α and $\alpha + \beta$, which correspond to the contribution of x to the mean when $x \leq e$ and $x > e$, respectively, converge at a rate between $n^{1/2}$ and $n^{1/3}$. The convergence rate is closer to $n^{1/3}$ when the sample size is larger, and in the series of sigmoid models, closer to $n^{1/3}$ when the true model is further away from the step model.

The behaviors of $\hat{\beta}$, the estimated jump at the threshold, are more interesting. For the quadratic model, the convergence rate is close to $n^{1/2}$ in smaller datasets but tends towards the asymptotic rate of $n^{1/3}$ in larger datasets. For all three sigmoid models, however, the convergence rate is $n^{1/2}$ in both smaller and larger datasets, suggesting that the asymptotic rate of convergence is $n^{1/2}$ and not $n^{1/3}$. The reason for this can be seen from Theorem 2.1 of Banerjee and McKeague (2007). The limiting distribution of β, when scaled by $n^{1/3}$, equals to $n^{1/3}(c_1 - c_2) \arg\max_t Q(t)$, but due to the symmetry of the these sigmoid models and the choice of the distribution of X, c_1 equals to c_2 in this special case causing this limiting distribution to be degenerate.

Finally, the estimates α_z, the slope associated with Z, converges at the regular $\sqrt{n}$ rate in both smaller and larger datasets.

3 Empirical studies of resampling-based confidence interval methods

3.1 Symmetric percentile Efron bootstrap confidence intervals

Efron bootstrap confidence intervals (Carpenter and Bithell 2000) are widely used for making inference in $\sqrt{n}$-convergence problems. The percentile Efron bootstrap

Table 2 Estimated convergence rates and their 95% CIs based on $n \in \{500, 1000, 1500, 2000\}$ from 10^4 Monte Carlo replicates

	e	β	α	$\alpha + \beta$	α_z
Step	−1.03 (−1.07, −1.00)	−0.50 (−0.50, −0.49)	−0.51 (−0.52, −0.51)	−0.51 (−0.52, −0.50)	−0.51 (−0.51, −0.50)
Sig_15	−0.45 (−0.53, −0.38)	−0.50 (−0.54, −0.46)	−0.51 (−0.52, −0.50)	−0.50 (−0.52, −0.47)	−0.51 (−0.53, −0.49)
Sig_5	−0.39 (−0.45, −0.34)	−0.50 (−0.54, −0.46)	−0.50 (−0.51, −0.48)	−0.47 (−0.51, −0.44)	−0.51 (−0.53, −0.49)
Sig_1	−0.35 (−0.39, −0.30)	−0.49 (−0.52, −0.46)	−0.44 (−0.48, −0.41)	−0.43 (−0.48, −0.37)	−0.51 (−0.52, −0.50)
Quad	−0.40 (−0.43, −0.38)	−0.49 (−0.50, −0.48)	−0.45 (−0.48, −0.42)	−0.46 (−0.48, −0.44)	−0.54 (−0.55, −0.53)

e: threshold; β: jump; α_z: slope of z; α: "lower step"; $\alpha + \beta$: "upper step"

Table 3 Estimated convergence rates and their 95% CIs based on $n \in \{1{,}024{,}000, 2{,}048{,}000, 4{,}096{,}000, 8{,}192{,}000\}$ from 10^4 Monte Carlo replicates

	e	β	α	$\alpha+\beta$	α_z
Step	−0.99 (−1.01, −0.98)	−0.50 (−0.50, −0.49)	−0.50 (−0.50, −0.49)	−0.50 (−0.50, −0.49)	−0.50 (−0.50, −0.49)
Sig_15	−0.33 (−0.34, −0.32)	−0.50 (−0.52, −0.47)	−0.42 (−0.44, −0.40)	−0.42 (−0.44, −0.41)	−0.50 (−0.52, −0.48)
Sig_5	−0.34 (−0.34, −0.33)	−0.50 (−0.52, −0.48)	−0.37 (−0.38, −0.36)	−0.38 (−0.40, −0.35)	−0.50 (−0.52, −0.48)
Sig_1	−0.33 (−0.35, −0.32)	−0.50 (−0.50, −0.50)	−0.35 (−0.35, −0.34)	−0.35 (−0.35, −0.34)	−0.50 (−0.50, −0.50)
Quad	−0.34 (−0.36, −0.33)	−0.42 (−0.43, −0.40)	−0.36 (−0.38, −0.34)	−0.37 (−0.39, −0.35)	−0.51 (−0.52, −0.49)

e: threshold; β: jump; α_z: slope of z; α: "lower step"; $\alpha+\beta$: "upper step"

confidence interval is well known to be inconsistent for non-regular problems, including step regression (e.g. Bühlmann and Yu 2002; Seijo and Sen 2011; Yu 2014). The theoretical properties of other types of general purpose Efron bootstrap confidence intervals (Carpenter and Bithell 2000), such as inverse percentile and symmetric percentile, are largely unknown. A type of bootstrap method specific for step regression, smoothed percentile bootstrap confidence intervals (Seijo and Sen 2011), has been proposed, but it involves a tuning parameter and can be hard to apply in practice (Yu 2014) and only works when the model is correctly specified. In this section, we compare the coverage of two types of Efron bootstrap confidence intervals under each of the five data generating models introduced in the previous section. In addition to the percentile method, we focus on the symmetric percentile ("symmetric") method (Hansen 2017). The symmetric method can be seen as a compromise between the percentile and inverse percentile methods. For example, for α the 95% confidence interval is defined as $\hat{\alpha} \pm q^*$, where q^* is the $95\%^{th}$ quantile of the bootstrap distribution of $|\hat{\alpha}^* - \hat{\alpha}|$.

As shown in Table 4, when the data is generated from a step model, both percentile and symmetric Efron bootstrap confidence intervals have coverage probabilities close to the nominal level 0.95 for the $n^{1/2}$-convergent slope parameters β, α_z and α. As for the threshold parameter e, the percentile method suffers from undercoverage. Importantly, this does not improve with larger sample sizes, providing empirical evidence that the Efron bootstrap is inconsistent for the n-convergent $\hat{e}$. Interestingly though, the symmetric method produces reasonable coverage with only a small increase in the width of the confidence intervals.

In contrast with the results under correct model specification, under model misspecification both the percentile and symmetric methods show a small amount of over-coverage for the threshold parameter e. For the slope parameter β, even though it is $n^{1/2}$-convergent under the Sig_1 model, the percentile method shows substantial undercoverage when the sample size is small-to-moderate; the symmetric method also under-covers, but to a much smaller extent. For the $n^{1/3}$-convergent α, both methods provide reasonable coverage, but show some over-coverage when the sample size increases. Moreover, the amount of over-coverage also increases from Sig_15 to Sig_1. These results, together with the results in Tables 2 and 3, suggest that the more the "effectual" convergence rates move from $n^{1/2}$ towards $n^{1/3}$, the more over-coverage there will be for both bootstrap confidence interval methods.

To better understand the reasons behind the difference in performance between the percentile and symmetric methods, we divide the 10^4 Monte Carlo replicates into three categories based on the skewness of the bootstrap sampling distributions of $\hat{e}$, as measured by the moment coefficient of skewness (Joanes and Gill 1998). We regard those calculated to be less than -0.5, between -0.5 and 0.5, and greater than 0.5 as left-skewed, not-skewed,and right-skewed, respectively (Bulmer 1966). We find that 40% of the bootstrap distributions are left-skewed, 20% are not-skewed, and 40% are right-skewed. While the percentile method covers the truth 71%, 89% and 89% of the time for left-skewed, not-skewed, and right-skewed bootstrap distributions, respectively, the symmetric method covers 96%, 95% and 95% of the time, respectively.

To summarize, these Monte Carlo experiment results show that the symmetric Efron bootstrap confidence interval method may correct for the under-coverage of the

Table 4 Simulation results from 10,000 Monte Carlo runs

	e		β		α		α_z	
n	Percentile	Symmetric	Percentile	Symmetric	Percentile	Symmetric	Percentile	Symmetric
Step								
250	83.0 (0.50)	96.7 (0.61)	94.9 (0.15)	95.1 (0.15)	95.1 (0.11)	95.3 (0.11)	94.6 (0.08)	94.8 (0.08)
500	83.1 (0.24)	96.8 (0.30)	94.9 (0.10)	95.1 (0.10)	94.7 (0.08)	94.9 (0.08)	94.6 (0.05)	94.8 (0.05)
1000	82.7 (0.12)	96.3 (0.14)	94.3 (0.07)	94.5 (0.07)	94.7 (0.05)	94.9 (0.05)	95.1 (0.04)	95.0 (0.04)
2000	82.0 (0.06)	96.3 (0.07)	94.8 (0.05)	94.9 (0.05)	94.9 (0.04)	95.1 (0.04)	94.7 (0.03)	94.9 (0.03)
8000	81.6 (0.01)	96.2 (0.02)	94.7 (0.03)	94.8 (0.02)	95.0 (0.02)	95.2 (0.02)	95.1 (0.01)	95.4 (0.01)
32,000	81.5 (0.00)	96.8 (0.00)	94.7 (0.01)	94.7 (0.01)	95.0 (0.01)	95.2 (0.01)	94.9 (0.01)	95.0 (0.01)
Sig_15								
250	91.3 (0.60)	96.4 (0.74)	93.9 (0.15)	94.7 (0.15)	94.8 (0.11)	95.3 (0.11)	94.6 (0.08)	94.7 (0.08)
500	93.3 (0.36)	96.2 (0.44)	94.2 (0.10)	94.8 (0.10)	94.6 (0.08)	95.0 (0.08)	94.6 (0.05)	94.7 (0.05)
1000	95.0 (0.23)	96.5 (0.29)	93.7 (0.07)	94.2 (0.07)	95.0 (0.06)	95.2 (0.06)	95.0 (0.04)	95.0 (0.04)
2000	95.6 (0.17)	96.4 (0.21)	94.5 (0.05)	94.9 (0.05)	95.2 (0.04)	95.6 (0.04)	94.8 (0.03)	94.9 (0.03)
8000	96.4 (0.10)	96.6 (0.13)	94.6 (0.03)	94.8 (0.03)	95.5 (0.02)	95.6 (0.02)	95.0 (0.01)	95.0 (0.01)
32,000	96.9 (0.06)	96.9 (0.08)	94.5 (0.01)	94.6 (0.1)	95.6 (0.01)	95.5 (0.01)	94.1 (0.01)	94.2 (0.01)

Table 4 continued

	e		β		α		α_z	
n	Percentile	Symmetric	Percentile	Symmetric	Percentile	Symmetric	Percentile	Symmetric
Sig_5								
250	94.3 (0.91)	96.2 (1.11)	92.2 (0.15)	94.2 (0.15)	95.1 (0.12)	95.5 (0.13)	94.6 (0.08)	94.7 (0.08)
500	95.5 (0.63)	96.7 (0.78)	92.9 (0.10)	94.4 (0.11)	95.1 (0.08)	95.3 (0.09)	94.6 (0.05)	94.8 (0.05)
1000	96.0 (0.46)	96.6 (0.57)	92.7 (0.07)	94.1 (0.08)	95.6 (0.06)	95.5 (0.06)	95.1 (0.04)	95.1 (0.04)
2000	96.5 (0.35)	96.9 (0.44)	93.5 (0.05)	94.8 (0.05)	96.2 (0.04)	95.9 (0.05)	94.9 (0.03)	94.9 (0.03)
8000	96.9 (0.22)	96.9 (0.27)	94.0 (0.03)	94.8 (0.03)	97.0 (0.02)	96.6 (0.03)	95.0 (0.01)	95.2 (0.01)
32,000	96.8 (0.13)	96.8 (0.17)	94.4 (0.01)	94.9 (0.1)	97.4 (0.01)	96.8 (0.01)	94.1 (0.01)	94.4 (0.01)
Sig_1								
250	96.6 (2.59)	96.9 (3.27)	77.9 (0.15)	90.6 (0.17)	95.3 (0.17)	96.2 (0.19)	95.0 (0.08)	94.9 (0.08)
500	97.1 (2.02)	96.9 (2.54)	81.4 (0.11)	91.5 (0.12)	95.8 (0.12)	96.3 (0.13)	94.8 (0.05)	94.9 (0.05)
1000	97.2 (1.57)	97.4 (1.97)	84.4 (0.07)	92.1 (0.08)	96.5 (0.09)	96.9 (0.10)	95.1 (0.04)	95.2 (0.04)
2000	97.3 (1.23)	96.9 (1.54)	86.5 (0.05)	92.7 (0.06)	97.3 (0.06)	97.1 (0.07)	95.1 (0.03)	95.1 (0.03)
8000	97.2 (0.77)	97.0 (0.96)	89.3 (0.03)	93.5 (0.03)	98.2 (0.04)	97.5 (0.04)	95.3 (0.01)	95.4 (0.01)
32,000	97.0 (0.49)	97.2 (0.61)	91.4 (0.01)	94.2 (0.01)	98.0 (0.02)	97.5 (0.03)	95.0 (0.01)	95.0 (0.01)
Quad								
250	94.8 (0.47)	96.4 (0.57)	95.4 (0.96)	95.0 (0.97)	95.8 (0.61)	94.7 (0.65)	96.4 (0.45)	95.9 (0.47)
500	95.1 (0.35)	96.2 (0.42)	95.3 (0.68)	95.0 (0.69)	96.2 (0.45)	95.4 (0.48)	96.7 (0.31)	96.3 (0.32)
1000	95.5 (0.26)	96.5 (0.32)	95.6 (0.49)	95.5 (0.49)	96.2 (0.33)	95.7 (0.35)	96.2 (0.21)	96.4 (0.21)
2000	95.6 (0.20)	96.3 (0.24)	95.8 (0.34)	95.6 (0.35)	96.3 (0.25)	95.8 (0.26)	96.2 (0.14)	96.1 (0.14)
8000	96.4 (0.12)	96.7 (0.15)	96.2 (0.18)	96.0 (0.18)	96.7 (0.13)	96.3 (0.15)	95.5 (0.07)	95.5 (0.07)
32,000	96.5 (0.07)	96.6 (0.09)	95.4 (0.09)	94.9 (0.09)	97.0 (0.07)	96.5 (0.08)	95.3 (0.03)	95.2 (0.03)

Estimated coverage and mean width of percentile and symmetric Efron bootstrap CIs

n-convergent threshold parameter exhibited by the percentile Efron bootstrap method when the step regression model is correctly specified. Moreover, the symmetric Efron bootstrap method provides reasonable coverage for the $n^{1/3}$-convergent model parameters when the step regression model is misspecified and the sample size is small to moderate.

3.2 Subsampling with a double bootstrap-like procedure for block size selection

Two types of resampling-based methods have been proposed for making inference for non-regular problems: m-out-of-n bootstrap, which resamples with replacement fewer than n observations (e.g. Bickel et al. 1997), and subsampling, which resamples without replacement fewer than n observations (e.g. Politis and Romano 1994; Bertail et al. 1999). Seijo and Sen (2011) showed that m-out-of-n bootstrap is valid in a step linear regression model when the mean model is correctly specified. However, it is not clear whether the method is still consistent when the mean model is misspecified. In our empirical studies (results not shown) the m-out-of-n bootstrap confidence intervals provided close-to-nominal coverage with appropriately chosen block sizes when the mean model is correctly specified, but over-covered for all block sizes when the mean model is misspecified. We thus focus our attention on subsampling.

An important consideration for using the subsampling method is the choice of the block size m_n, or m for short (Politis et al. 1999). A common approach in all blocking methods (e.g. Delgado et al. 2001; Gonzalo and Wolf 2005; Chakraborty et al. 2013) is nested resampling. This leads to a double bootstrap-like (DBL) procedure (Delgado et al. 2001), in which the original dataset is bootstrapped and subsampling is performed on each bootstrap dataset at a grid of candidate block sizes to look for the block size that provides a close to nominal estimated coverage. The procedure can be described more formally as follows:

1. Draw B_1 first-level samples by Efron bootstrap from the data, and calculate the first-level estimates of the parameter of interest $\hat{e}^{b_1}$, $b_1 = 1, \ldots, B_1$.
2. For each first-level bootstrap sample ($b_1 = 1, \ldots, B_1$):

(a) Draw B_2 second-level samples by sampling m subjects without replacement and calculate the second-level parameter estimates $\hat{e}^{b_1,b_2}$, $b_2 = 1, \ldots, B_2$ for the smallest m in a grid of 25 values evenly spaced between 0.05 and $0.8n$.

(b) Denote the $(\alpha/2) \times 100$ and the $(1 - \alpha/2) \times 100$ percentiles of $\hat{e}^{b_1,b_2}(b_2 = 1, \ldots, B_2)$ by $\hat{e}^{b_1}_{\left(\frac{\alpha}{2}\right)}$ and $\hat{e}^{b_1}_{\left(1-\frac{\alpha}{2}\right)}$, respectively. Construct the subsampling confidence interval for a first-level sample as $(\hat{e}^{b_1}_{\left(\frac{\alpha}{2}\right)}, \hat{e}^{b_1}_{\left(1-\frac{\alpha}{2}\right)})$.

3. Estimate the coverage probabilities at each m by $\frac{1}{B_1}\sum_{i=1}^{B_1} \mathbb{I}\{\hat{e}^{b_1}_{\left(\frac{\alpha}{2}\right)} \leq \hat{e}_n \leq \hat{e}^{b_1}_{\left(1-\frac{\alpha}{2}\right)}\}$, where $\hat{e}_n$ is the estimate of the parameter of interest from the original dataset.
4. Increment the block size m from the smallest to the next higher value in the grid and repeat steps 2 and 3 until the estimated coverage probability goes above the nominal level.

5. Find m that corresponds to the nominal coverage level by linear interpolation between the two block sizes whose estimated coverage probabilities bracket the nominal level.

For the choice of B_1, which affects how well we can estimate the coverage probabilities under different block sizes, and B_2, which affects how well we can estimate the subsampling confidence intervals, we experiment with different combinations of B_1 and B_2 with a fixed $B_1 \times B_2$. The results, shown in Table C.3 of the Supplementary Materials, suggest that the performance of the procedure is not overly sensitive to the choice of B_1, but when B_2 is too small (50), it leads to small selected block sizes, wide confidence intervals, and over-coverage. For the remainder of the paper, we let $B_1 = 200$ and $B_2 = 200$.

The block size m selected by the double bootstrap-like procedure and the corresponding coverage probabilities and widths of subsampling confidence intervals for the threshold parameter e are shown in Table 5. When the model is correctly specified, the DBL confidence intervals appear to over-cover; comparing with Table 4, we see that the DBL confidence intervals are on average longer than the symmetric Efron bootstrap confidence intervals for the threshold parameter, e.g. at $n = 1000$, the mean confidence interval width is 0.14 and 0.20 for symmetric Efron bootstrap and DBL, respectively. When the model is misspecified, the DBL confidence intervals also over-cover, but the degree of over-coverage decreases as the sample size increases. Comparing with Table 4, we see that the DBL confidence intervals are on average shorter than the symmetric Efron bootstrap confidence intervals for the threshold parameter, e.g. at $n = 1000$, the mean confidence interval width is 0.32 and 0.28 for symmetric Efron bootstrap and DBL, respectively, when the data is simulated from the quadratic model.

3.3 Subsampling with simple rules-of-thumb for block size selection

The heavy computational burden of the double bootstrap-like procedure motivates us to develop alternative methods for block size selection. We start by determining the optimal block sizes for each of the five data generating models we have studied. To do this, we estimate the coverage probabilities of subsampling confidence intervals using 10^4 Monte Carlo replicates for each m in a grid of block sizes. We then find the m that corresponds to the nominal coverage level by linear interpolation between the two block sizes whose estimated coverage probabilities bracket the nominal level.

The selected block sizes are listed in Table C.1 in the Supplementary Materials. We use a linear model to model the relationship between the Monte Carlo-selected block sizes and the sample sizes, both on the log scale, under correct model specification, and we use a linear mixed effects model to model their relationship under the four misspecified models (Fig. 1. We obtain the following relationships:

$$log(m) = -0.9207 + 0.9804 \times log(n), \text{ under correct model specification} \quad (5)$$

$$log(m) = -0.5565 + 0.9961 \times log(n), \text{ under model misspecification} \quad (6)$$

Table 5 The performance of the double bootstrap-like procedure with $B_1 = 200$ and $B_2 = 200$

	Step		Sig_15		Sig_5		Sig_1		Quad	
n	m	cvg (width)	m	cvg (width)	m	cvg (width)	m	cvg (width)	m	cvg (width)
250	109	98.0 (0.77)	121	97.4 (0.76)	129	96.6 (0.99)	134	95.3 (2.62)	126	96.8 (0.56)
500	212	98.2 (0.39)	248	97.6 (0.43)	255	96.8 (0.67)	267	95.3 (1.99)	259	96.4 (0.38)
1000	404	98.2 (0.20)	502	97.1 (0.27)	518	96.1 (0.47)	530	96.2 (1.54)	511	96.1 (0.28)
1500	605	98.2 (0.13)	759	96.9 (0.21)	778	96.2 (0.40)	793	95.5 (1.34)	791	96.1 (0.23)
2000	787	98.4 (0.10)	1037	96.7 (0.18)	1047	96.1 (0.36)	1055	95.9 (1.21)	1037	95.6 (0.20)

m average selected block size, *cvg* coverage probabilities of *e* estimated from 10^4 Monte Carlo replicates, *width* average width of subsampling confidence intervals

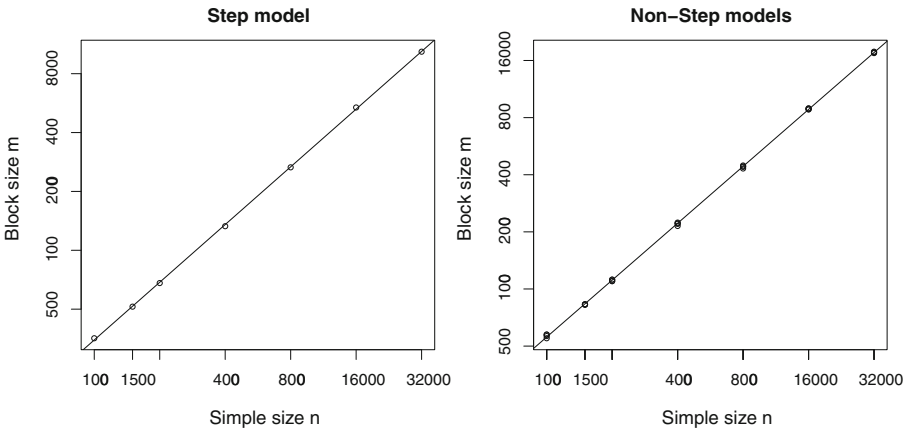


Fig. 1 Fitted linear regression lines (linear model for step model and linear mixed effects model with random intercept for non-step models) of the relationship between the sample size n and the block size m. The block size m is derived by letting the coverage probabilities be closest to 0.95 for threshold parameter e by subsampling bootstrap

Table 6 shows the performance of the subsampling confidence interval method when these rules-of-thumb are used to select block sizes. When the model is correctly specified, the block sizes selected by the rule-of-thumb are on average smaller than those selected by the DBL procedure, which leads the rule-of-thumb confidence intervals to be longer; however, the coverages provided by the rule-of-thumb confidence intervals are actually smaller than those provided by the DBL procedure. If we restrict to the Monte Carlo replicates for which the DBL method covers and the rule-of-thumb method does not cover, the rule-of-thumb confidence intervals are shorter, but among the Monte Carlo replicates for which both methods cover, the rule-of-thumb confidence intervals are longer. When the model is misspecified, the block sizes selected by the rule-of-thumb are on average bigger. This leads both the confidence intervals to be shorter and the coverage probabilities to be smaller, as we would expect.

3.4 Additional Monte Carlo studies

These simulation results suggest that the symmetric Efron bootstrap confidence interval and the two subsampling confidence intervals using either a double bootstrap-like procedure or a rule-based procedure to select block size all provide reasonable coverage. To investigate the robustness of their performance under a wider variety of scenarios, we conduct two additional Monte Carlo studies. First, we shift the distribution of X so that the threshold e_0 is not always at the center of the distribution. The results, summarized in Supplementary Material Section E.1, show that all three methods perform similarly as in the original Monte Carlo study. Second, we let the noise term ϵ be distributed as a Student's t distribution with four degrees of freedom instead of a normal distribution. The results, summarized in Supplementary Material Section E.2, show that while the symmetric Efron bootstrap method and the subsampling method using the double bootstrap-like procedure to select block size perform

similarly as in the original Monte Carlo study, the subsampling method with the rule-based method to select block size under-covers when the data is generated from a step model.

Taken together, these results suggest that both the symmetric Efron bootstrap method and the subsampling method with a double bootstrap-like procedure to select block size provide good coverage and can be recommended. These two methods perform similarly when the model is misspecified, but the symmetric bootstrap method performs better, with narrower confidence intervals and closer-to-nominal coverage, when the model is correctly specified. The symmetric bootstrap method is also computationally more efficient than the subsampling method; the subsampling method, on the other hand, is better understood theoretically. The subsampling method with a rule-based procedure to select block size provides a useful alternative when a faster subsampling method is desired, but its coverage may be insufficient when the model is correctly specified and the noise distribution is heavy-tailed.

4 Hydro-climatic drivers of agricultural grassland productivity under extreme drought and rainfall

Efficient methods for estimating thresholds are important in environmental sciences, where complex and large datasets are frequent and abrupt non-linear threshold responses are common. Eutrophication of lake ecosystems (Carpenter and Lathrop 2008), fire mediated vegetation transitions in forests and woodlands (Adams 2013) and algal responses to light in polar ecosystems (Clark et al. 2013) all show some evidence for step threshold responses. However, despite threshold responses being a common feature of natural systems, accurate and efficient techniques for the estimation of thresholds are lacking.

To demonstrate the utility of the fast grid search algorithm developed here for step regression on a real world dataset, we used a remotely sensed derived dataset ($n = 2549$) on grassland productivity responses to soil moisture changes during drought on the Darling Downs, eastern Australia (Plant et al. 2021; Kath et al. 2019). Grasslands and forests are good model systems for investigating thresholds because the roots of vegetation have a discrete physiological limit to the amount of soil moisture they can access. Once moisture levels decline below a critical level, plants can no longer access water and so a rapid decline in plant biomass (possibly leading to plant death) occurs. Given the rapid nature of this change, it is likely to occur as a step threshold (Elmore et al. 2006; Kath et al. 2014). In line with these expectations, Plant et al. (2021) applied Bayesian additive regression trees (BART) to model grassland productivity responses, using the Darling Downs dataset, to soil moisture. Their results suggested a clear step-like response of the rate of grassland productivity change with soil moisture anomalies. The BART approach, while allowing the subjective visualization of a threshold, does not provide an estimate of where that step threshold occurs, nor its uncertainty.

To provide an estimate and quantify the uncertainty of a potential threshold response of grassland productivity change to soil moisture anomalies, we fit a step threshold using the fast grid approach developed here. We use the rule-of-thumb assuming correct model specification (5) because both the regression tree modelling result (Plant et al.

Table 6 Performance of the prediction rules

	Step		Sig_15		Sig_5		Sig_1		Quad	
n	m	cvg (width)	m	cvg (width)	m	cvg (width)	m	cvg (width)	m	cvg (width)
250	89	95.4 (0.89)	140	91.3 (0.52)	140	93.7 (0.80)	140	95.0 (2.40)	140	93.1 (0.45)
500	176	94.9 (0.43)	279	93.3 (0.32)	279	94.8 (0.57)	279	95.2 (1.83)	279	93.8 (0.33)
1000	347	95.3 (0.21)	557	94.7 (0.22)	557	95.0 (0.42)	557	95.6 (1.43)	557	94.8 (0.25)
1500	517	95.0 (0.14)	835	94.8 (0.18)	835	94.9 (0.36)	835	94.9 (1.24)	835	94.4 (0.21)
2000	686	94.6 (0.10)	1112	95.0 (0.16)	1112	95.3 (0.32)	1112	95.4 (1.11)	1112	94.3 (0.19)

m block size selected by the proposed rules-of-thumb, *cvg* coverage probabilities of *e* estimated from 10^4 Monte Carlo replicates, *width* average width of subsampling confidence intervals

Table 7 For the grassland example, point estimates and confidence intervals for the threshold e and jump β

	With covariates adjustment		Without covariates adjustment	
	e	β	e	β
Point estimate	-0.26	0.00025	-0.29	0.00065
symmetric bootstrap	$(-0.29, -0.23)$	(0.00019, 0.00032)	$(-0.32, -0.26)$	(0.00058, 0.00072)
Subsampling-dbl	$(-0.32, -0.25)$		$(-0.31, -0.26)$	
Subsampling-rule	$(-0.32, -0.24)$		$(-0.30, -0.26)$	

Subsampling-d and subsampling-r use the double bootstrap-like procedure and the proposed rules of thumb, respectively, to select block size

2021, Fig. 8) and a three-phase segmented model fit (Fig. 2a left panel) suggest a very sharp transition.

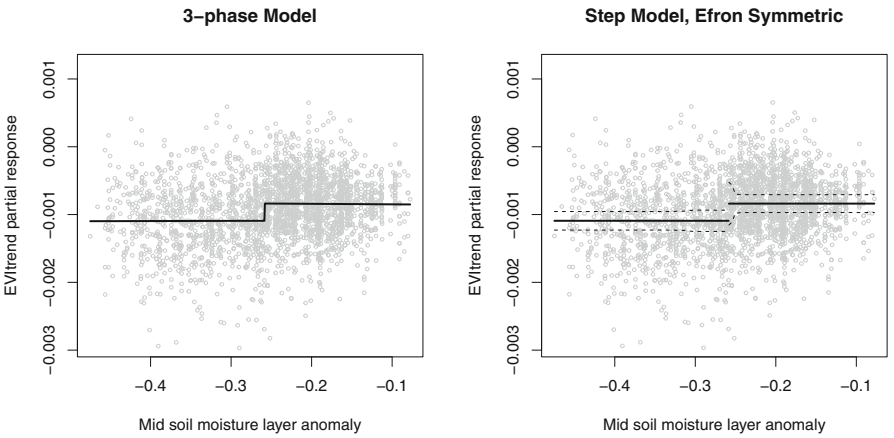
We account for the influence of 11 other hydro-climatic (e.g. soil moisture and evaporation) and land use (e.g. proportion of woody vegetation and agriculture in the landscape) predictors that may influence the relationship between soil moisture and grassland productivity (Table D.1 in the Supplementary Materials). To account for the nonlinear associations between grassland productivity and these variables, we allow each variable 1–9 degrees of freedom as selected by cross-validated generalized additive models (Wood 2017).

The right panel of Fig. 2a shows the step model fit. The change point estimate by the step model is -0.26 with an estimated jump of 2.5×10^{-4} (95% symmetric Efron bootstrap CI 1.9×10^{-4}, 3.2×10^{-4}) (Table 7):

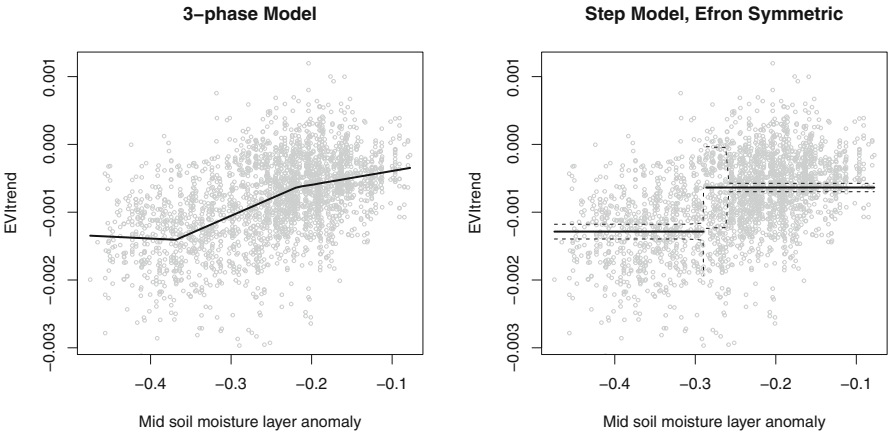
$$\mathbb{E}(\text{EVI trend}) = \alpha + \alpha_z^T z + 0.00025 \times (\text{mid soil moisture layer anomaly} + 0.26)$$

The step model therefore quantifies the step threshold shift to a greater rate of grassland productivity decline under drought once soil moisture anomalies exceed a threshold value of a -0.26 (or a -26% anomaly in soil moisture). Consistent with the simulation study results under Sig_15, symmetric Efron bootstrap produces slightly shorter 95% confidence interval than subsampling with a DBL procedure for block size selection.

For illustration, we also fit a step model between grassland productivity change to soil moisture anomalies without adjusting for any other covariates. We use the rule-of-thumb assuming model misspecification (6) because a three-phase segmented model fit (Fig. 2b left panel) clearly shows that the true model is unlikely to be a step model. The change point estimate by the step model is -0.29 with an estimated jump of 6.5×10^{-4} (95% symmetric Efron bootstrap CI: 5.8×10^{-4}, 7.2×10^{-4}) (Table 7). Consistent with the simulation study results under Sig_1, symmetric Efron bootstrap produces slightly longer 95% confidence interval than subsampling with a DBL procedure for block size selection.



(a) With covariate adjustment.



(b) Without covariate adjustment.

Fig. 2 The grassland example. The solid lines are the fitted curves. The dashed lines are pointwise 95% confidence bands. Subsampling-d and subsampling-r use the double bootstrap-like procedure and the proposed rule-of-thumb, respectively, to select block sizes. In panel **a**, "EVI trend partial response" is defined as the observed EVI trend minus the predicted value based on the adjusted covariates

5 Discussion

Our proposed fast grid search algorithm finds globally optimal solutions to a non-convex, non-smooth problem. It is several orders of magnitude faster than the brute-force grid search algorithm and makes it feasible to analyze large datasets. We illustrated the use of step regression with a dataset on grassland productivity. Since thresholds are often used to inform targets around which to base environmental

management decisions (Simmonds et al. 2019), step regression could be applied in a range of natural settings to inform environmental guidelines and regulations (e.g. safe water quality thresholds that are set in freshwater systems, forest restoration targets, etc.). The proposed methods are implemented in the R package *chngpt*, which is hosted on the Comprehensive R Archive Network. The R scripts for simulation studies and real data analysis are available on the Github code repository youyifong/StepModelSearchBootstrap.

We studied three resampling-based confidence interval methods under a variety of data generating models. The results support recommendation of the symmetric Efron bootstrap method (symmetric bootstrap) and the subsampling method with a double bootstrap-like procedure (subsampling-dbl) for selecting block size. Both methods have nominal or conservative coverage. The symmetric bootstrap is faster, and produces substantially narrower confidence intervals than subsampling-dbl when the model is correctly specified. We designed our simulation studies to include three models, Sig_1, Sig_5, and Sig_15, which increasingly resemble a step model. Interestingly, when the sample size is small ($n = 250$), symmetric bootstrap confidence intervals are also narrower under Sig_15; this difference decreases under Sig_5 and disappears under Sig_1. When the sample size is large enough ($n = 2000$), the two methods produce confidence intervals of similar width under all three sigmoid models. On the other hand, the symmetric bootstrap is not as well understood theoretically as subsampling-dbl and may not be as general as subsampling-dbl.

We focused on the threshold parameter in the development of subsampling-based confidence interval methods. Using the block size selected to provide good coverage for the threshold parameter e is not guaranteed to work for other parameters. This is because different parameters converge at different rates and even parameters with the same convergence rates may require different sample sizes for the asymptotics to kick in. Table C.2 of the Supplementary Materials shows the coverage of all model parameters when the subsampling block size is chosen to optimize the coverage of e. The results show that the coverage for β in the Sig_1 model is well below the nominal level. To achieve good coverage for parameters other than e, we recommend symmetric Efron bootstrap, which is fast and provides a coverage between 90-95% depending on the sample size (Table 4). Alternatively, for a specific parameter, e.g., β, subsampling with a modified double bootstrap-like procedure for targets β may provide improved coverage at the cost of a higher computational burden.

When there are multiple covariates with threshold effects in a step regression model, the fast grid algorithm can be generalized by searching through a multi-dimensional grid of candidate thresholds. A detailed description of the algorithm for two covariates with threshold effects is shown in Section F of the Supplementary Materials. It is worth noting that as the number of candidate thresholds increases, exhaustive search using even the type of fast grid search algorithm proposed here quickly becomes impractical due to the exponentially increasing computational burden, which makes heuristic search a necessity. However, the fast grid search algorithm developed here could become part of the heuristic search algorithm as an efficient way for providing high quality solutions to sub-problems.

In this work we have operated under a step regression model that assumes the observations are independent, the error term and the predictors are independent of

each other, and the predictors are without measurement errors. The simplicity of this model allows the least squares estimator for the submodel with a fixed threshold to have a closed-form solution, which provides the target for our optimization. Whether the proposed approach to accelerating computation can be extended to more complex regression models warrants future research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10651-022-00547-2.

Acknowledgements We are grateful to the Editor and four anonymous referees for their constructive comments. We are grateful to Lindsay N. Carpp for her editorial support. We thank Professor Jon A. Wellner for discussion on the convergence behavior of $\hat{\beta}$ in the simulation studies. This work was supported by the National Institutes of Health (R01-AI122991; UM1-AI068635; S10OD028685).

References

Adams MA (2013) Mega-fires, tipping points and ecosystem services: managing forests and woodlands in an uncertain future. For Ecol Manag 294:250–261

Banerjee M, McKeague IW (2007) Confidence sets for split points in decision trees. Ann Stat 35(2):543–574

Bertail P, Politis DN, Romano JP (1999) On subsampling estimators with unknown rate of convergence. J Am Stat Assoc 94(446):569–579

Bickel PJ, Götze F, van Zwet WR (1997) Resampling fewer than n observations: gains, losses, and remedies for losses. Stat Sin 7(1):1–31

Bühlmann P, Yu B (2002) Analyzing bagging. Ann Stat 30(4):927–961

Bulmer MG (1966) Principles of statistics. T. Press, Cambridge

Carpenter J, Bithell J (2000) Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. Stat Med 19(9):1141–1164

Carpenter SR, Lathrop RC (2008) Probabilistic estimate of a threshold for eutrophication. Ecosystems 11(4):601–613

Chakraborty B, Laber EB, Zhao Y (2013) Inference for optimal dynamic treatment regimes using an adaptive m-out-of-n bootstrap scheme. Biometrics 69(3):714–723

Clark GF, Stark JS, Johnston EL, Runcie JW, Goldsworthy PM, Raymond B, Riddle MJ (2013) Light-driven tipping points in polar ecosystems. Glob Change Biol 19(12):3749–3761

Delgado MA, Rodrıguez-Poo JM, Wolf M (2001) Subsampling inference in cube root asymptotics with an application to Manskis maximum score estimator. Econ Lett 73(2):241–250

Elder A, Fong Y (2019) Estimation and inference for upper hinge regression models. Environ Ecol Stat 26(4):287–302

Elmore AJ, Manning SJ, Mustard JF, Craine JM (2006) Decline in alkali meadow vegetation cover in California: the effects of groundwater extraction and drought. J Appl Ecol 43(4):770–779

Fong Y (2019) Fast bootstrap confidence intervals for continuous threshold linear regression. J Comput Graph Stat 28(2):466–470

Fong Y, Huang Y, Gilbert P, Permar S (2017) chngpt: threshold regression model estimation and inference. BMC Bioinform 18(1):454–460

Friedman JH, Silverman BW (1989) Flexible parsimonious smoothing and additive modeling. Technometrics 31(1):3–21

Gallant AR, Fuller WA (1973) Fitting segmented polynomial regression models whose join points have to be estimated. J Am Stat Assoc 68(341):144–147

Gonzalo J, Wolf M (2005) Subsampling inference in threshold autoregressive models. J Econometrics 127(2):201–224

Hansen BE (2017) Regression kink with an unknown threshold. J Bus Econ Stat 35(2):228–240

Joanes DN, Gill CA (1998) Comparing measures of sample skewness and kurtosis. J R Stat Soc Ser D 47(1):183–189

Kath J, Reardon-Smith K, Le Brocque A, Dyer F, Dafny E, Fritz L, Batterham M (2014) Groundwater decline and tree change in floodplain landscapes: identifying non-linear threshold responses in canopy condition. Glob Ecol Conserv 2:148–160

Kath J, Le Brocque AF, Reardon-Smith K, Apan A (2019) Remotely sensed agricultural grassland productivity responses to land use and hydro-climatic drivers under extreme drought and rainfall. Agric For Meteorol 268:11–22

Kosorok MR (2008) Introduction to empirical processes and semiparametric inference. Springer, New York

Muggeo V (2003) Estimating regression models with unknown break-points. Stat Med 22(19):3055–3071

Pastor R, Guallar E (1998) Use of two-segmented logistic regression to estimate change-points in epidemiologic studies. Am J Epidemiol 148(7):631–642

Plant E, King R, Kath J (2021) Statistical comparison of additive regression tree methods on ecological grassland data. Ecol Inform 61:101198

Politis DN, Romano JP (1994) Large sample confidence regions based on subsamples under minimal assumptions. Ann Stat 22(4):2031–2050

Politis DN, Romano JP, Wolf M (1999) Subsampling. Springer, New York

Pons O (2003) Estimation in a cox regression model with a change-point according to a threshold in a covariate. Ann Stat 31(2):442–463

Seijo E, Sen B et al (2011) Change-point in stochastic design regression and the bootstrap. Ann Stat 39(3):1580–1607

Simmonds JS, van Rensburg BJ, Tulloch AI, Maron M (2019) Landscape-specific thresholds in the relationship between species richness and natural land cover. J Appl Ecol 56(4):1019–1029

Son H, Fong Y (2020) Fast grid search and bootstrap-based inference for continuous two-phase polynomial regression models. Environmetrics 1–20 (in press)

Song R, Banerjee M, Kosorok MR (2016) Asymptotics for change-point models under varying degrees of mis-specification. Ann Stat 44(1):153

Tishler A, Zang I (1981) A new maximum likelihood algorithm for piecewise regression. J Am Stat Assoc 76(376):980–987

Wood S (2017) Generalized Additive Models: An Introduction with R, 2nd edn. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, Boca Raton

Yu P (2014) The bootstrap in threshold regression. Econometric Theory 30(3):676–714