Supplementary materials for *Breakpoints in the global temperature*

Junbo Duan, Lanling Zhao, Qing Wang, and Pei Li

I. Simulation data

We simulated a piecewise linear (polynomial order $r = 1$) signal as the ground truth to match the CP model, and the time interval ranging from 1880 to 2018 to correspond with the GISTEMP and NOAA datasets. Two breakpoints were set at years 1940 and 1970, yielding three segments. The middle segment was horizontal with a value of 0, and the left and right segments had a slope parameter α°C per year. Since the estimated α was $0.014-0.018$°C per year [1], α was set to 0.005, 0.01, 0.015, and 0.02, which indicated there were inconsequential to significant trends. The two breakpoints have amplitudes of β°C. Since the estimated β from the five datasets was 0.20 ± 0.07°C (Sec. III-C), β was set as 0 and 0.2, which corresponded to the continuous points and breakpoint statuses.

To mimic real data, white noise with a standard deviation σ was added. Since the estimated σ from the real data (last 50 data points of NOAA) is 0.086, σ was set as 0.1 and 0.05 in our simulation, corresponding to a real noise scenario and a less heavy noise scenario. The simulated signal was processed by both the proposed and the CP model. For the proposed model, the polynomial order r was set to 1, and the Bayesian/Schwarz information criterion (BIC/SIC) was used. For the CP model, a configuration with two change points was used; Fig. 1 shows an illustrative example of the simulation data and the processing results with the following configuration: $\alpha = 0.02, \beta = 0.2, \sigma = 0.1$.

The detection results were analysed in terms of sensitivity (or statistical power, recall) and specificity, which characterizes the statistical performance of a detector by the tradeoff between the true positive rate and the true negative rate [2]. The value of the sensitivity and the specificity ranged between 0 and 1, with 1 indicating an ideal detector. In the simulation study, if the bias of a detected breakpoint (or CP) with respect to the ground truth (1920 or 1970) is less than 5 years, the detection is deemed to be a true positive; otherwise, a false positive.

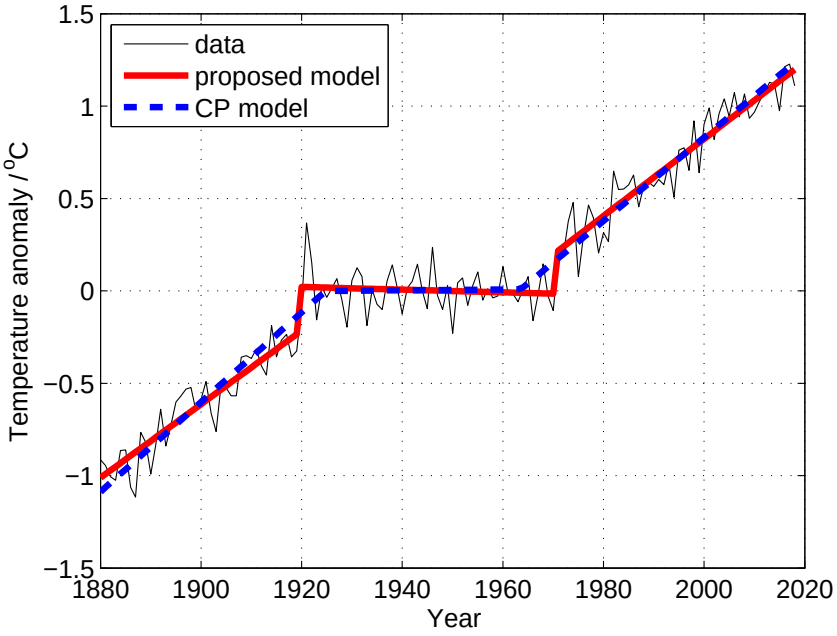


Fig. 1. An example simulation with $\alpha = 0.02, \beta = 0.2, \sigma = 0.1$.

Tab. I lists the detection results of both the proposed model and the CP model in terms of sensitivity and specificity with respect to the slope parameter α, break amplitude parameter β, and noise level parameter σ. Each value is the average of 100 Monte-Carlo replicates with different noise values.

The results show that the sensitivity of both models increases as the slope parameter α increases or as the noise level parameter σ decreases. As the break amplitude parameter β increases, the sensitivity of the proposed model increases, and that of the CP model decreases. In all the simulations, the specificity remains at a high level (greater than 0.84). In the end, when the configuration is close to the real data (($\alpha = 0.015$, $\beta = 0.2$ and $\sigma = 0.1$)), the proposed model achieved higher sensitivity (0.91) and specificity (0.98) compared with the CP model (0.20 and 0.86).

TABLE I
SUMMARY OF SIMULATION RESULTS.

α	β	σ	Proposed model		CP model	
			Sensitivity	Specificity	Sensitivity	Specificity
0.005	0	0.05	0.19	0.85	0.29	0.89
		0.1	0.15	0.85	0.13	0.88
	0.2	0.05	0.98	1.00	0.08	0.88
		0.1	0.89	0.98	0.04	0.88
0.01	0	0.05	0.42	0.89	0.55	0.92
		0.1	0.28	0.86	0.48	0.91
	0.2	0.05	0.98	1.00	0.12	0.84
		0.1	0.92	0.98	0.17	0.86
0.015	0	0.05	0.52	0.91	0.77	0.96
		0.1	0.42	0.89	0.69	0.94
	0.2	0.05	0.97	0.99	0.20	0.86
		0.1	0.91	0.98	0.20	0.86
0.02	0	0.05	0.67	0.94	0.79	0.96
		0.1	0.53	0.91	0.76	0.96
	0.2	0.05	0.97	0.99	0.30	0.88
		0.1	0.91	0.98	0.33	0.88

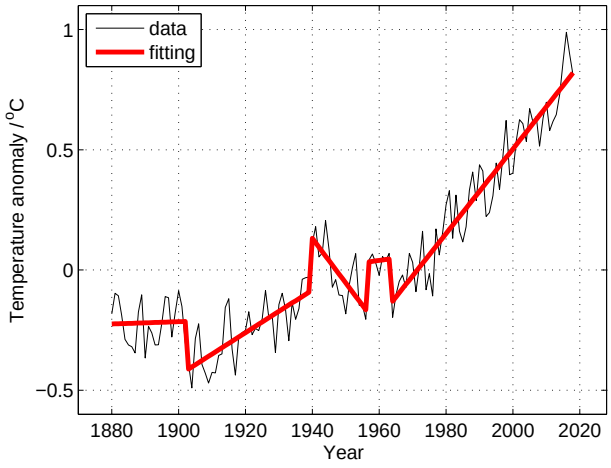
II. REAL DATA

Full results of five datasets are shown if Fig. 2.

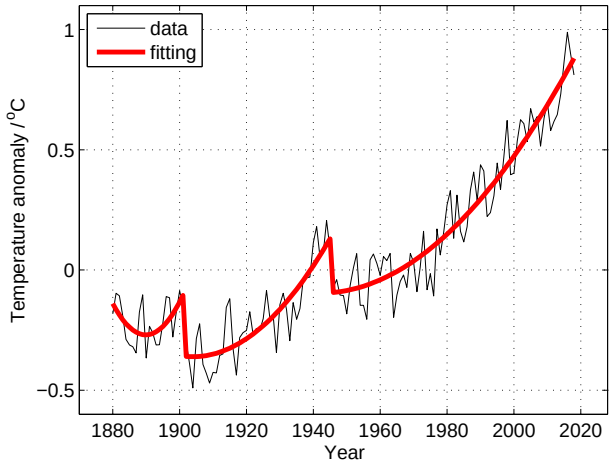
REFERENCES

[1] G. Foster and S. Rahmstorf, "Global temperature evolution 1979–2010", *Environmental Research Letters*, vol. 6, no. 4, pp. 044022, 2011.

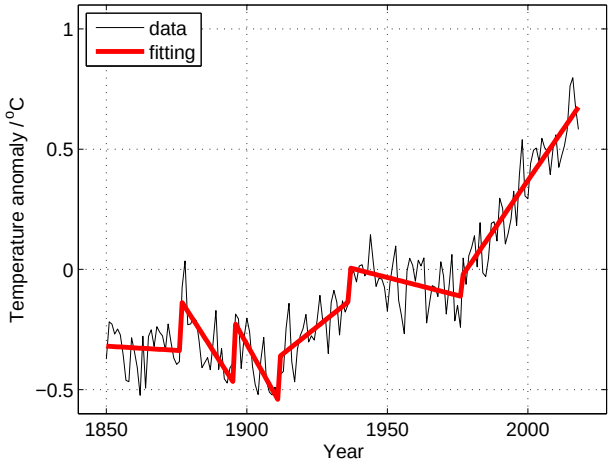
[2] D. M. Powers, "Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation", *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.


data
fitting
Temperature anomaly /°C
Year
1
0.5
0
−0.5
1880
1900
1920
1940
1960
1980
2000
2020

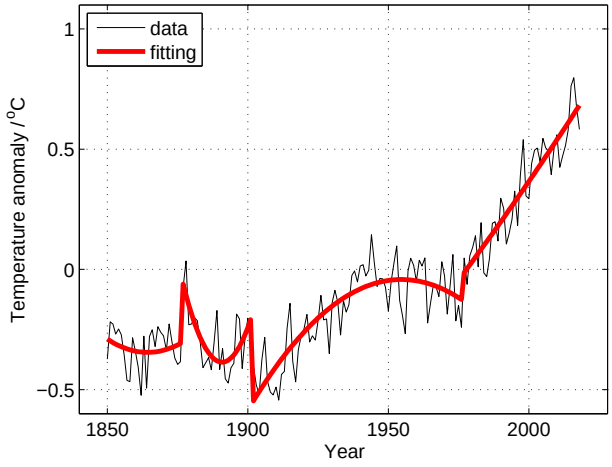
(a)


data
fitting
Temperature anomaly /°C
Year
1
0.5
0
−0.5
1880
1900
1920
1940
1960
1980
2000
2020

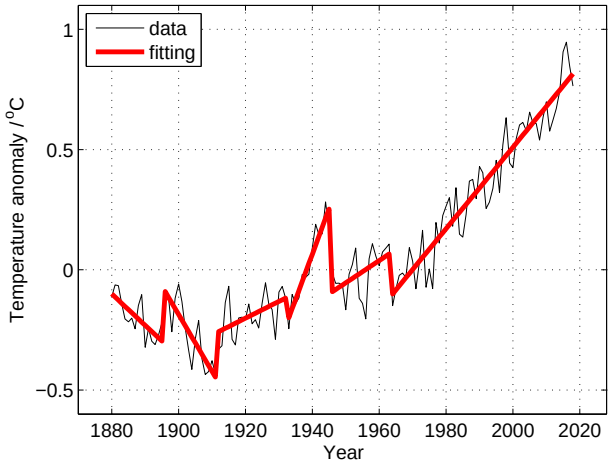
(b)


data
fitting
Temperature anomaly /°C
Year
1
0.5
0
−0.5
1850
1900
1950
2000

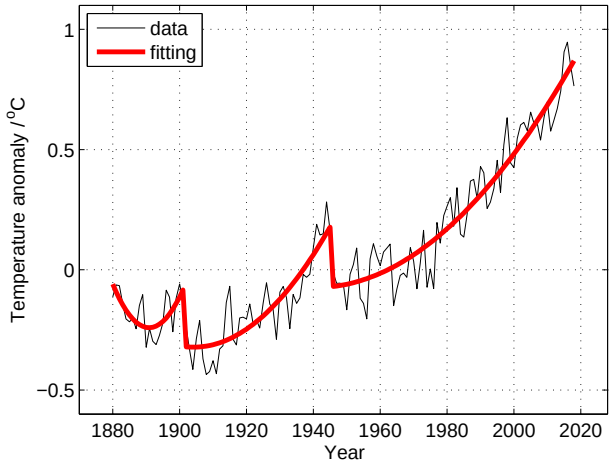
(c)


data
fitting
Temperature anomaly /°C
Year
1
0.5
0
−0.5
1850
1900
1950
2000

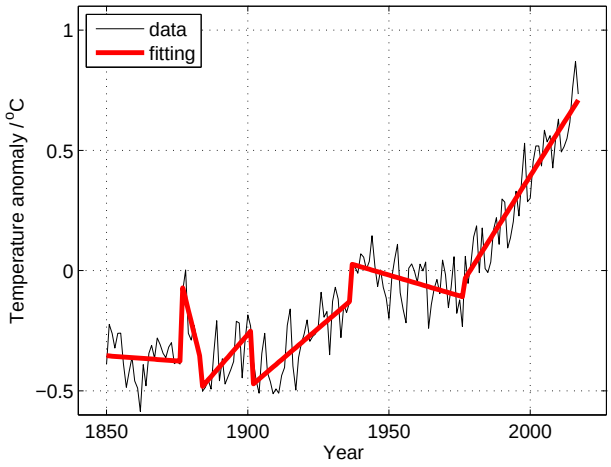
(d)


data
fitting
Temperature anomaly /°C
Year
1
0.5
0
−0.5
1880
1900
1920
1940
1960
1980
2000
2020

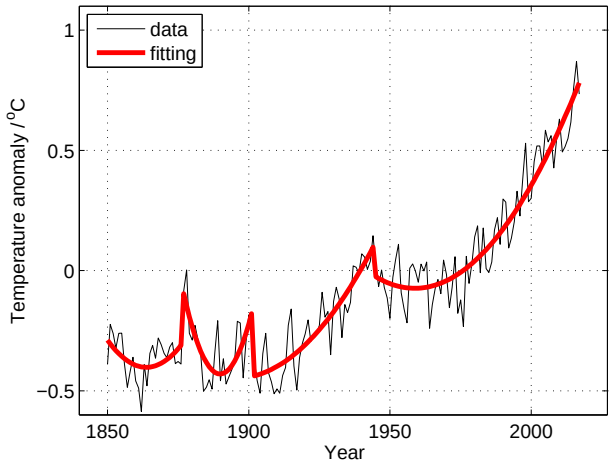
(e)


data
fitting
Temperature anomaly /°C
Year
1
0.5
0
−0.5
1880
1900
1920
1940
1960
1980
2000
2020

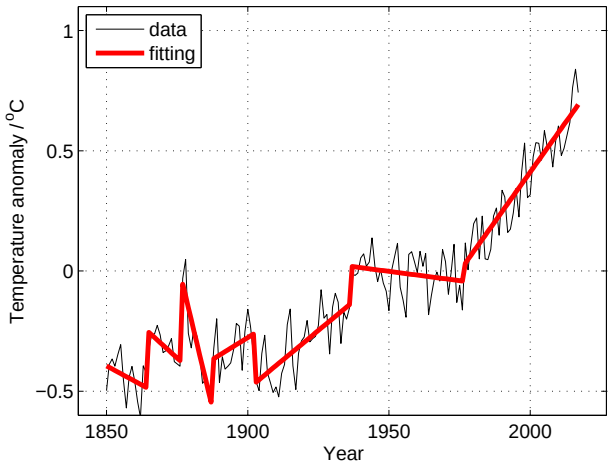
(f)



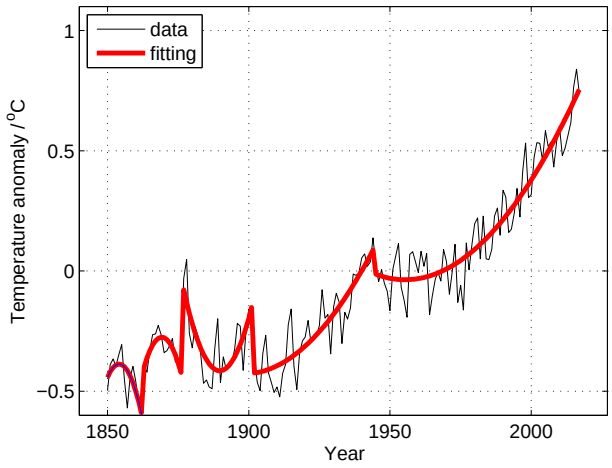
(g)



(h)



(i)



(j)

Fig. 2. Full results of five datasets. Left and right panels show results with polynomial order $r = 1$ and $r = 2$. From the top to bottom rows are the datasets: GISTEMP, HadCRUT4, NOAA, CW, and Berkeley, respectively.