LOCAL SMOOTHING COMPARISON

OCTOBER 9, 2021

KO TSZ NGA (VALERIE)
valerie.ktn@gmail.com

Introduction

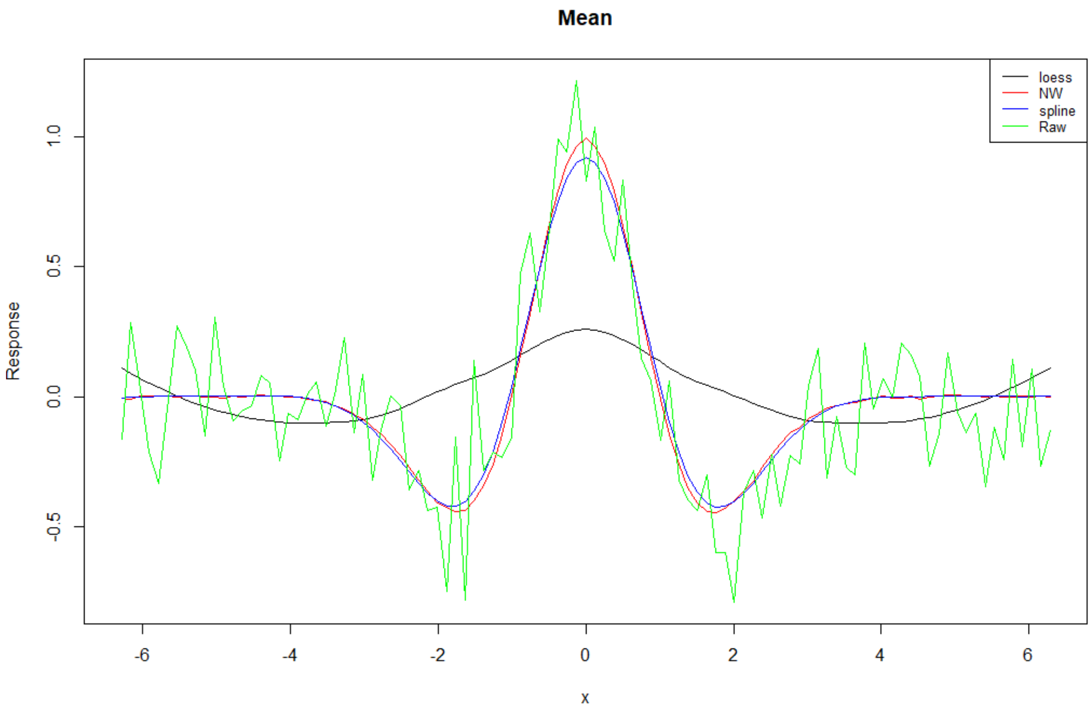
This project focuses on the comparison between 3 smoothers, namely **loess**, **Nadaraya-Watson (NW) kernel smoothing** and **spline smoothing**. m = 1000 Monte Carlo runs will be used, and in each run, we simulate a data set of n = 101 observations with the Mexican hat function. The comparison will be made based on **mean**, **empirical bias**, **empirical variance**, and **empirical mean square error (MSE**). 2 sets of data points based on deterministic fixed design will be generated: (**1) Equi-distant points (2) Non-equidistant points.**

Results

(1) Deterministic fixed design with equi-distant points

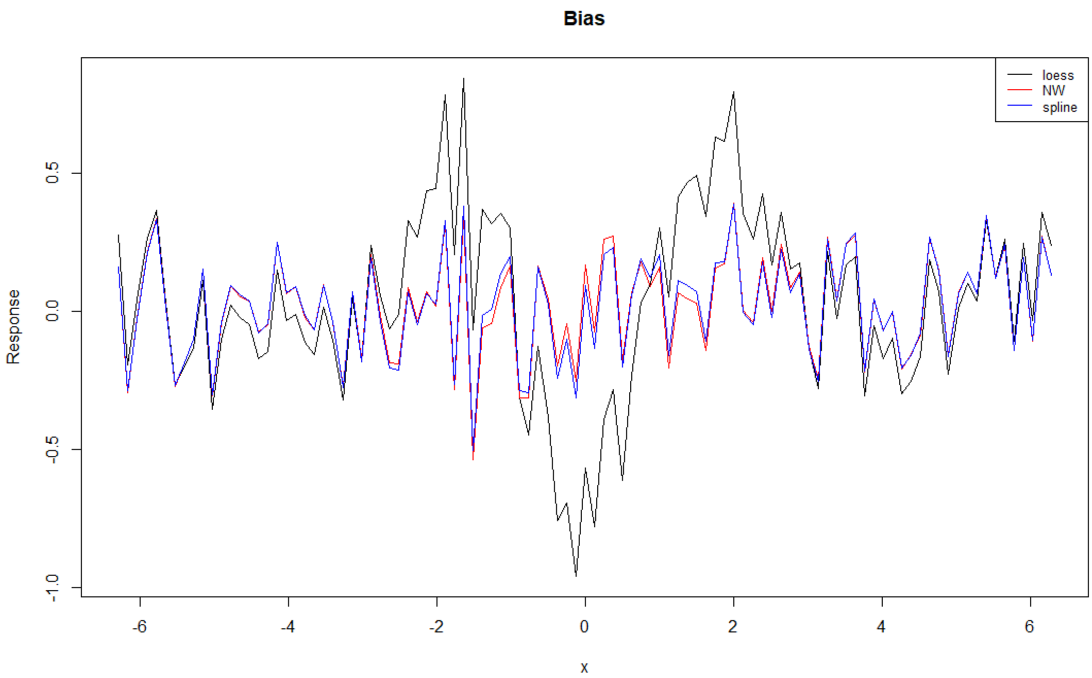
a. Mean

The means obtained are plotted against x_i. The result is compared with the Y_i, noted in green.



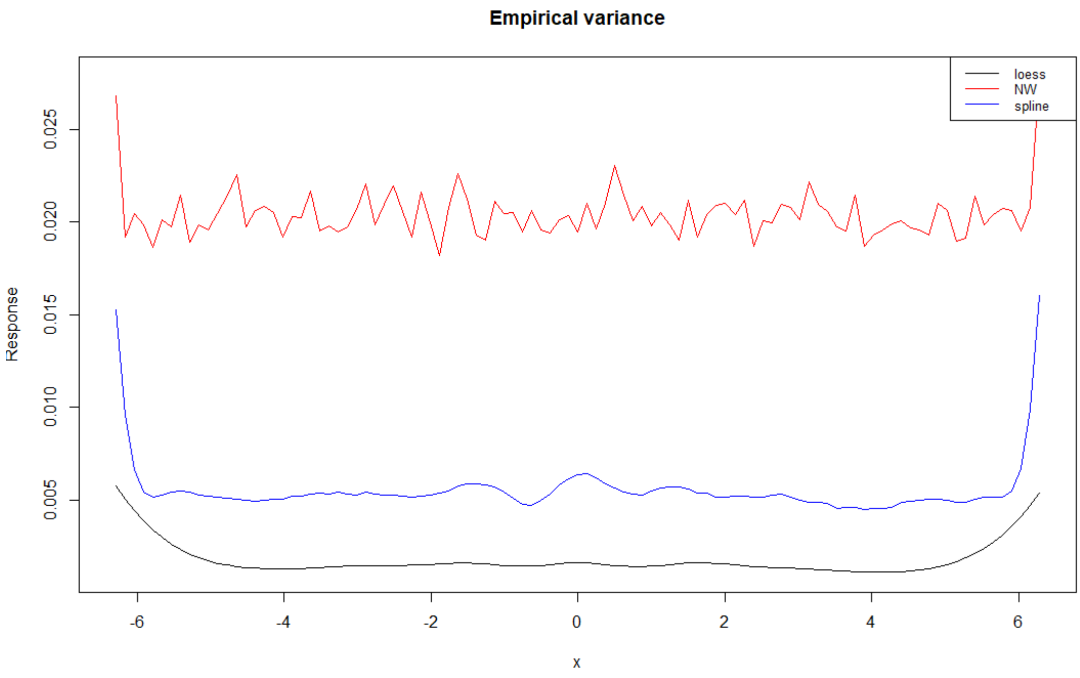
Noted that both **Nadaraya-Watson (NW) kernel smoothing** and **spline smoothing** show a more similar trend compared with the raw Y_i data. Particularly, the **NW kernel** smoothing is **more sensitive** in capturing the trend. For example, it gives a higher response in the peak (ie. when x=0) and lower responses in the dips (ie. when x=-0.2 and x=0.2). In contrast, the smoothing line given by **loess** does not really fit the peaks and dips really well. For instance, its lowest dips were around x=-0.4 and x=0.4, but the lowest dips of the raw Y_i in general were around x=-0.2 and x=0.2. Based on the comparison on the mean, **NW kernel** seems to be better.

b. Bias



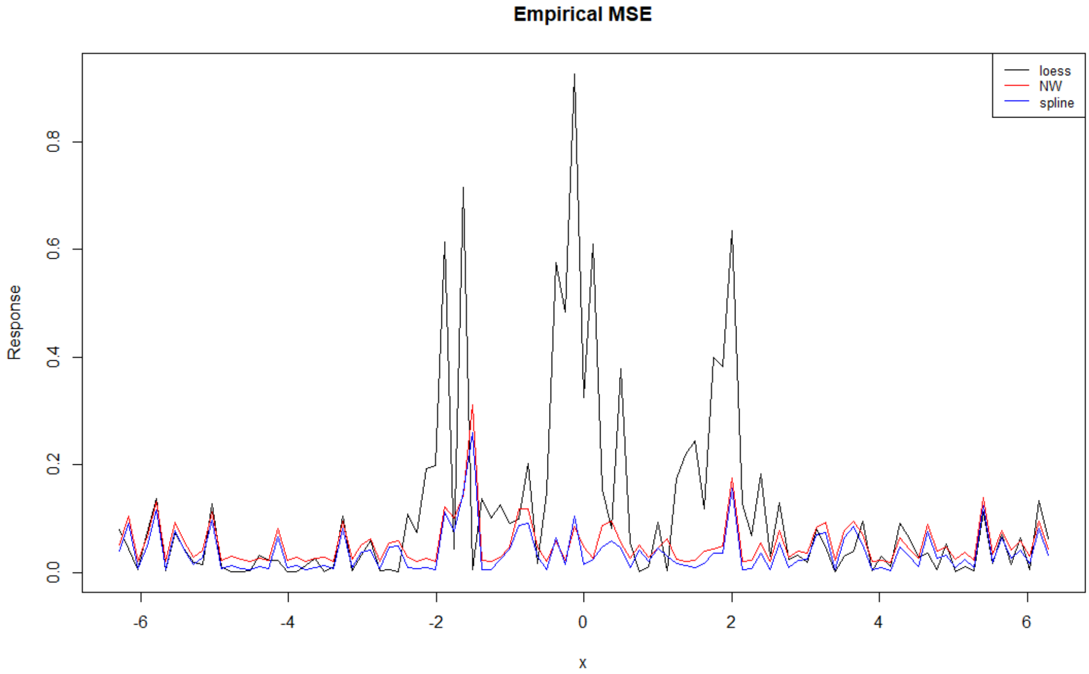
In a general point of view, the biases of **NW kernel** smoother and **spline** smoother are around 0, with fluctuation within -0.5 and 0.5. That implies they have relatively **low or even 0 biases** stably. On the other hand, the bias of loess varies a lot within a range between -1 and 1. To be precise, the bias reaches nearly -1 when x=0 and they reach near 1 when x is around -2 and 2. It shows that **loess** has a strong bias in predicting smaller magnitude of x.

c. Empirical variance



The empirical variance gives an interesting result — No overlap between the lines is found. Also, the ‘smoothness’ between of the trends are quite different, except all of them share a higher variance value when x=-6 and x=6. The variance of **loess** is the lowest and it gives the “smoothest” graph. That of **spline** is the second lowest with slight variance bumps mainly around x=0. The variances of **NW kernel** are the highest, with an observable gap from the other 2 smoothers. It fluctuates a lot mainly around 0.02.

d. Empirical MSE

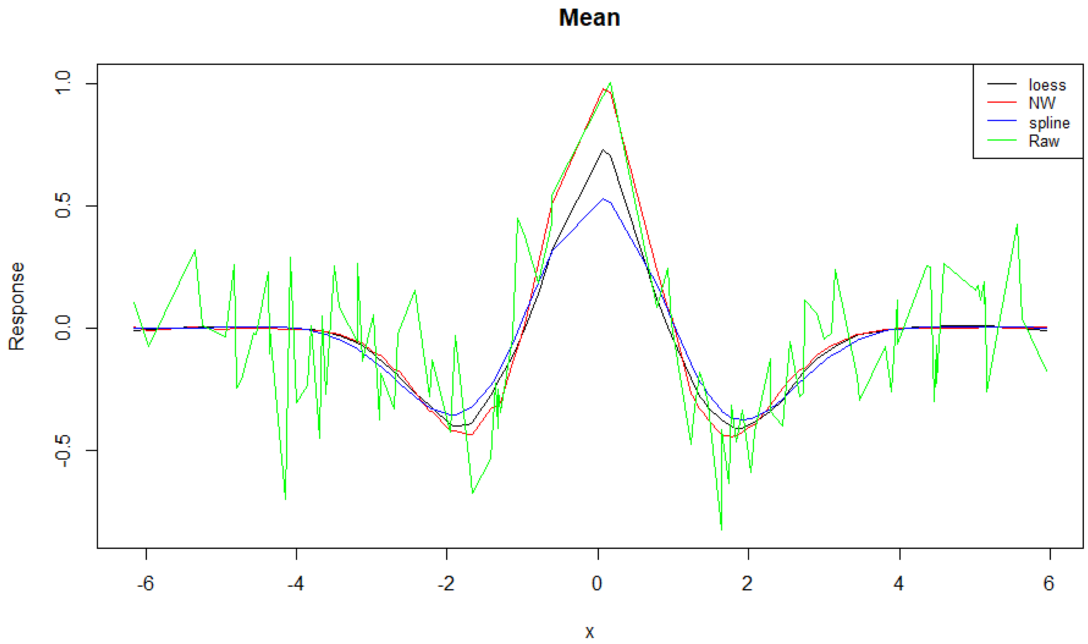


The empirical MSE of **NW kernel** and **spline smoother** are quite similar. The slight variation between these two is just the NW kernel one gives a slightly higher MSE. In general, they both has a low MSE mainly below 0.2 except for x near -2. However, the graph for **loess** shows a stark contrast with the other two. It has 3 high ‘peaks’ when x is around -2, 0 and 2 respectively. The peaks notify that **loess** has a worse predicting performance around these values.

(2) Deterministic fixed design with non-equidistant points

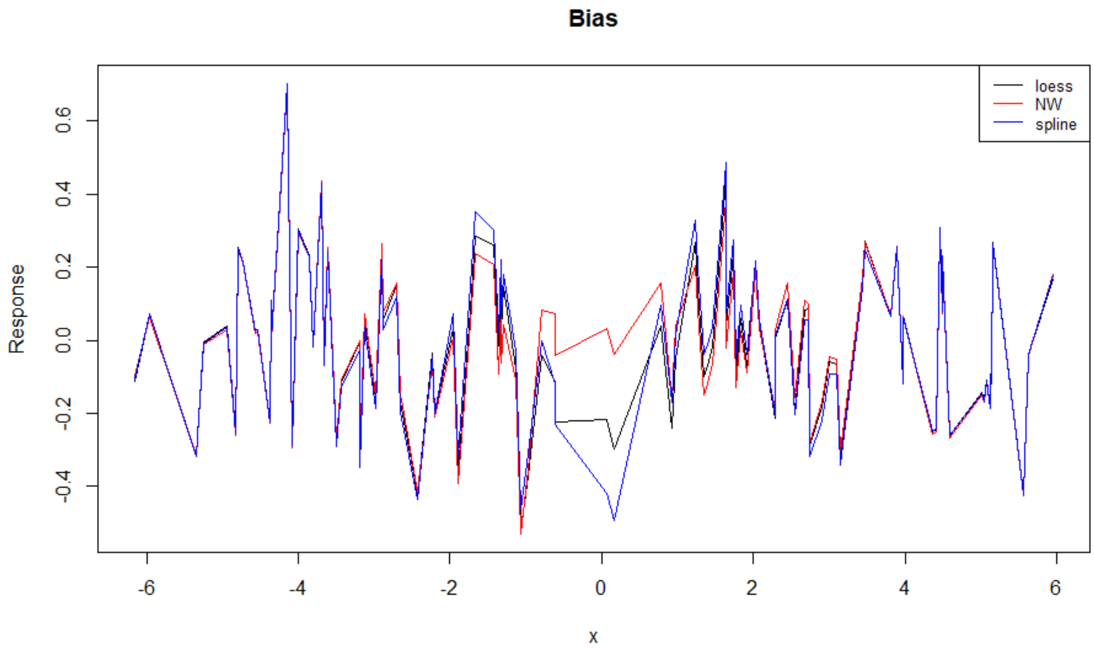
a. Mean

Again, the means obtained are plotted against x_i. The result is compared with the Y_i, noted in green.



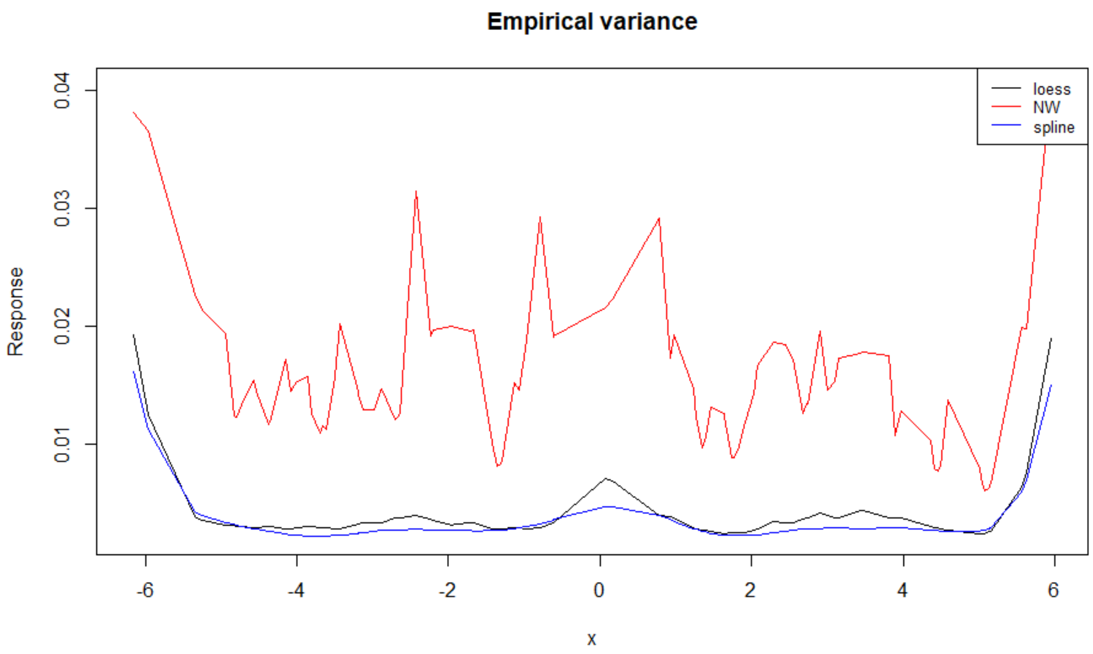
All 3 smoothers follow the general trend of the raw Y data points. Their performance varies only in the dips (ie. around x=-2 and x=2) and the peak at x=0. Specifically, the **NW kernel smoother** is the most sensitive one, it gives the highest peaks and lowest dips. For example, it nearly matches the exact mean of the response Y when x=0. From this point of view, **NW kernel** smoother seems to be the best among the 3 in terms of prediction.

b. Bias



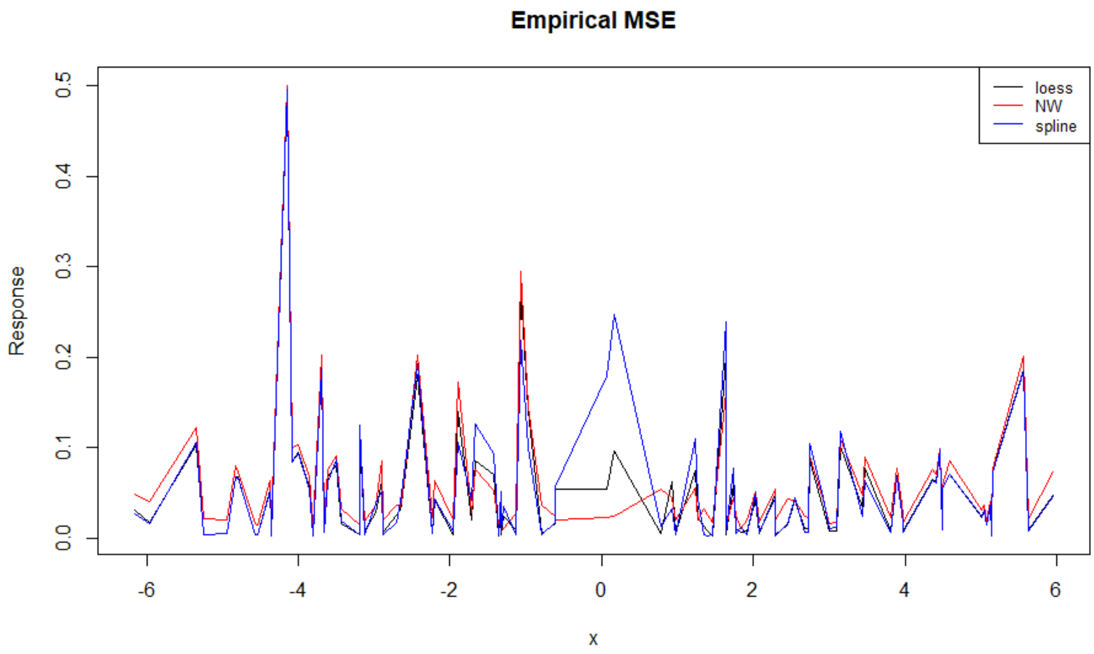
For bias, the 3 smoothers vary a lot when x is between -2 and 2. When x is around -2, **spline** gives the highest bias. Around x=0, spline bias reaches -0.5, loess reaches around -0.2 and NW kernel is around 0. In short, **NW kernel** gives a relatively stable bias and that of the **spline** fluctuates the most.

c. Empirical variance



The graph of the variance of the **NW kernel** smoother is very different from the other 2, both in magnitude and trend. That implies it gives the highest variance in general, especially when x is around -2.5, -0.5 and 0.5. **Spline** and **loess** have a stably lower variance, and both have a variance surge around x =0.

d. Empirical MSE



The first thing that captures my attention is the high MSE around x=-4 for all 3 smoothers. It reaches 0.5 which is abnormal when comparing to the general trend that varies from 0 to 0.3. The other place that I want to stress on is around x=0, it is the only place where the 3 smoothers give a more different performance. The **spline** has the highest MSE at this point versus **NW kernel** smoother stays around 0.

Findings

For better comparison, I summarise the smoothers performance with (1) equi-distant points and (2) non-equidistant points as follows:

	Smoothers that perform well based on the respective criteria	
	(1) Equi-distant points	**(2) Non-equidistant points**
Mean (Criteria: Follow the graph of mean of raw Y)	NW kernel (slightly better) Spline	NW kernel
Bias (Criteria: Have bias around 0 in general)	NW kernel Spline	NW kernel
Empirical variance (Criteria: Have variance close to 0 in general)	loess	Spline (slightly better) loess
Empirical MSE (Criteria: Have MSE close to 0 in general)	NW kernel Spline (slightly better)	Inconclusive

Conclusions

To conclude, I would say **NW kernel** gives a stably well performance when assessing with **mean and bias** regardless of the data points used. In terms of **variance, loess** performs well regardless of the data points. Yet, for **MSE**, I have to conclude that the result is quite different between the 2 sets of data points. None of the smoothers give a generally stable trend around 0 when non-equidistant points are used. In regards of that, I don't think it is fair to assess their performance with **MSE**. Also, **further investigation** should be carried out. I would recommend focussing more on the sudden **MSE** surge when x=-4 and the stark **MSE** variation when x=0 in the **non-equidistant points** case.

In short, **NW kernel** smoother is better in **minimising bias** while **loess** is better in **minimising variance**. It reminds me of the **bias-variance trade-off** concept. The **low bias** of **NW** means it is capable in capturing the trend of the true y, which agrees with the performance in mean graph. However, it may result in overfitting the model. In contrast, the **low variance** in **loess** means it can potentially give consistent prediction but maybe oversimplify the model. Without the conclusive finding in **MSE**, I don't think it is fair to compare them.