SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Mechanics

Journal homepage:
https://semarakilmu.com.my/journals/index.php/appl_mech/index
ISSN: 2289-7895

Modifications of 4253HT Smoother in Extracting Signal from Noise

Adie Safian Ton Mohamed[1], Noor Izyan Mohamad Adnan[1,*], Fadila Amira Razali[1], Sharifah Norhuda Syed Wahid[1], Qasim Nasir Husain[2]

[1] College of Computing, Informatics and Mathematics, Universiti Teknologi MARA Pahang, Jengka Campus, Bandar Tun Abdul Razak, Jengka, Pahang, Malaysia
[2] Department of Mathematics, College of Education for Pure Science, Tikrit University, Iraq

ARTICLE INFO

Article history:
Received 15 September 2023
Received in revised form 17 November 2023
Accepted 5 December 2023
Available online 9 January 2024

Keywords:
4253HT smoother; hanning; non-linear smoother

ABSTRACT

The smoothing technique has been used broadly in data analysis to obtain sharp signals or patterns by eliminating noise or formless patterns from a data set. Various smoothers with great properties were reported in previous studies, mainly non-linear smoothers such as 4253HT smoother. Hence, the 4253HT was selected to be used in this study, with some modifications done on mean 2 with different types of Hanning. The modifications on mean 2 were done by substituting with Geometric, Quadratic, Harmonic, and Contraharmonic means. Meanwhile, three types of Hanning were used in this study, including Tukey, Shitan, and Husain's type. Moreover, through a simulation study, this study aims to determine the best combination of smoother between four signals, which were Doppler, HeaviSine, Bumps, and Block signals. A Residual Mean Square Error (RMSE) was used as an evaluator to determine and assess the performance of modified 4253HT smoother when utilizing each smoother and noise level. As a result, Hanning Husain exhibited the best performance among all. Moreover, Hanning Tukey performed better on signal Block at lower noise levels (10% and 25%), while Hanning Shitan showed the worst performance. Besides, modified 4253HT using contra harmonic mean smoother recorded the most satisfying outcome in smoothing compared to other smoothers. Therefore, the 4253HT smoother will demonstrate the best performance with the combination of S5 smoother and Hanning Husain. Thus, it is suggested to utilize this smoother in further analysis, mainly in forecasting to provide accurate values and pattern for prediction.

1. Introduction

Generally, the purpose of conducting data analysis and nearly all exploratory data analysis is to observe the patterns of data. Smoothing is one of the methods used in data analysis to obtain appropriate signals from the uneven sequence of data values. Moreover, it is necessary to perform smoothing of a time series as it mainly aims to eliminate noises consisting in data. Previous study [1,2] had proven this in their studies by applying a non-linear smoother in forecasting of Malaysian crude palm oil prices. Meanwhile, Azmi [3] also did a modification on smoother before

* *Corresponding author.*
E-mail address: noorizyan@uitm.edu.my

https://doi.org/10.37934/aram.113.1.92107

being applied in forecasting. Following that, an appropriate signal of possible trend and shape of the distribution could be attained by utilizing the smoothed values.

According to Gabbouj [4], the linear smoothing method has been used broadly in studies since the procedure is not complicated to be applied in data analysis. Furthermore, linear smoothers are proven ideal in removing common tracking trends, and Gaussian noise consists of a data series [5]. Despite its simplicity, Bernholt [5] also stated that linear smoothers do not function well in non-linear data series and are highly susceptible to outliers. Besides, it causes loss of crucial information and blurry edges in data series when sudden changes occur.

Nevertheless, a running median smoother was introduced by Tukey [6] to overcome the limitation of linear smoothers, as these smoothers are able to preserve edges caused by sudden changes and are robust to outliers. Moreover, Velleman [7] acknowledged running median as a great tool used in statistical studies to extract signals from possibly spikey noise or long-tailed distribution. Regardless of its advantages, running the median is likely to eliminate Gaussian noise, which would impair important signals in a data series due to over-smoothing. In addition, standard median smoothers have both edge preservation and noise removal properties for heavy-tailed distribution.

Running median has been improved extensively through the years, including windsorized smoother, recursive running median, weighted running median, repeated running median, and compound smoother. Once again, Tukey [6] initiated the concept of compound smoother, which has since been expanded to several intriguing forms such as Hanning, splitting, repeating, and re-smoothing the rough [7]. Amongst all, 4253HT is the most prevalent and well- established compound smoother technique. This technique is very effective in eliminating spikes consisting of a data series before performing further parametric analysis.

Furthermore, 4253HT is an excellent tool for analyzing trends without destroying essential aspects of a data series. Many scholars used 4253HT in preliminary analysis to study trend trajectory in various fields, such as image signal processing [8], medical [9,10], seismology [11], agriculture [12], microbiology [13], climatology [14], and finance [15].

Researchers have been conducting several comparison studies to evaluate the efficacy of smoothers through simulation and practical analysis of real data. Amongst non-linear smoothers that have been explored, the 4253HT indicates the most remarkable performance in smoothing. However, Tothmeresz [16] and Jin [17] identified that 4253HT does not work efficiently in oscillated trend. Moreover, Janosky [18] stated that 4253HT should have not less than seven observations. Otherwise, it will converge to a constant root. Apart from that, modifications on Velleman’s [7] compound smoother have been performed by experimenting with various combinations of multiple stages of running median, Hanning, and rough re-smoothing [19]. Note that the improvements to the current compound smoother have yet to be thoroughly investigated.

For the purpose of enhancing the performance of a compound smoother, Sargent [19] combined the smoothing algorithm made up of the running median of various span sizes, Hanning, and "twice" to fit the performance of Australian football players. The output of smoother was then applied for forecasting using exponential smoothing method. The findings showed that forecasts performed using smoothed data from a compound smoother method were superior to those produced using actual data. In order to reconstruct normalised difference vegetation index (NDVI) time series data, Jin [17] suggested RMMEH, a compound smoother that consists of moving average, median smoother, maximum smoother, and Hanning. Besides, Jin [17] also accepted that 4253HT is a good smoother over others although RMMEH is known as to be better at smoothing the NDVI data based on particular criteria. However, improvements on the existing compound smoother, 4253HT has yet been investigated and explored broadly.

Throughout this study, 4253HT was used to determine the performance of modifications on mean 2 with different types of Hanning. Several modifications were done on 4253HT, where mean 2 was substituted with Contraharmonic, Harmonic, Quadratic, and Geometric means. Besides, 4253HT was operated together with three different types of Hanning, which are Husain, Shitan, and Tukey. In addition, a simulation study was performed to identify the most effective combination of smoother between four signals, either Doppler, HeaviSine, Bumps, and Block signals.

2. Methodology

This section is divided into four parts that will further describe the procedures of each modification process performed in this study. The first part of this section introduces compound smoothers and explains the procedure of operating 4253HT smoother throughout the study. Besides, the modifications of 4253HT smoother are precisely described in the next part. Following that, the third part consists of the details related to the types of Hanning used in the modification process. Moreover, the simulation procedure and noise used are described at the end of this section.

Figure 1 represents a general procedure to determine the best smoother using simulation process. Further details of each process are described in the next section.

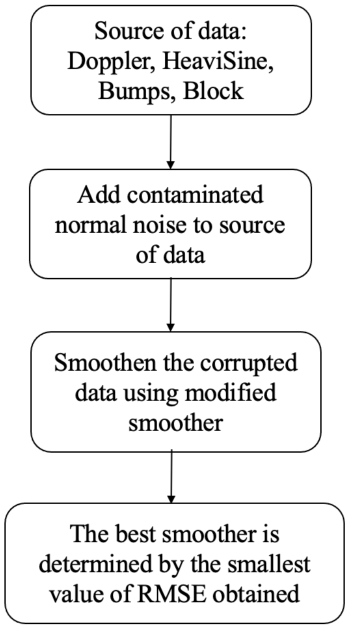

Fig. 1. Flowchart of general smoothing procedures in determination of the best smoother

2.1 Compound Smoother and 4253HT

A compound smoother is a non-linear method that can lessen the heavy noise of a signal while unaffected by sudden shifts and impulses in a data series. 4253HT smoother is one of the compound smoothers proven to deliver a great performance in smoothing. This smoother has been explored extensively by modifying its algorithm to focus on estimating the middle point of running median for even span size, where several types of means were applied, including Quadratic, Geometric, Contraharmonic, and Harmonic.

A 4253HT smoother S is an example of a compound smoother, which is also recognized as a non-linear smoother. Initially, 4253HT was proposed by Tukey [6]. A few years later, Vellemen [7] had done various modifications to 4253HT to enhance its performance. This smoother is made up of odd and even window medians. The function of even window medians is to reduce the issues related to odd window medians. In addition, Shitan [20] and Velleman [7] claimed that 4253HT is the most excellent non-linear smoother among all.

In most cases, the application of single running median smoother (e.g., 2-median, 3-median. . . etc.) in smoothing is inefficient in reducing noises from a signal (data set). In fact, it causes difficulty in observing the actual pattern of a data series. Due to that, the evolution of smoother carried out by Tukey [6] had resulted in a compound smoother, which was formed from the process of repeated running median of different windows, Hanning or running weighted average, splitting, and re-roughing that is called as twice. Besides, compound smoothers have beneficial properties such as edge preservation, variation reduction, idempotency, co-idempotency, etc.

Generally, the smoothing method using 4253HT is performed based on a combination of running weighted averages and median. Tukey [6] pioneered this smoothing method, which Velleman [7] described to a great extent. Moreover, compound smother 4253HT is an algorithm that is applied to Y to produce a new series of smoothed values S(yi), where Y is a double-infinite sequence of real data.

$\ldots, Y_{t-2}, Y_{t-1}, Y_t, Y_{t+1}, Y_{t+2}, \ldots$

Compound smother 4253HT is an algorithm that is applied to Y to produce a new series of smoothed values S(yi). Hence, the steps of applying 4253HT smoother are as below:

Step 1: A running median of window four was performed, then re-centered using the running median of window two:

$$S_4(y_i) = median\,[y_{i-2},\ y_{i-1}, y_i,\ y_{i+1}] \tag{1}$$

$$S_{42}(y_i) = median\,[S_4(y_i),\ \ S_4(y_{i+1})] \tag{2}$$

Step 2: Running median of window size five was applied and followed by running median of window size three:

$$S_{425}(y_i) = median[S_{42}(y_{i-2}),\ S_{42}(y_{i-1}),\ S_{42}(y_i),\ S_{42}(y_{i+1}), S_{42}(y_{i+2})] \tag{3}$$

$$S_{4253}(y_i) = median[S_{425}(y_{i-1}), S_{425}(y_i),\ S_{425}(y_{i+1})] \tag{4}$$

Step 3: Using coefficients, $A = \left\{\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right\}$ as an illustration for algorithm, Hanning, H was performed as below:

$$S_{4253H}(y_i) = \frac{1}{4}S_{4253}(y_{i-1}) + \frac{1}{2}S_{4253}(y_i) + \frac{1}{4}S_{4253}(y_{i+1}) \tag{5}$$

Step 4: The residual or rough, e_i, then was inserted back into the smoothed values, which is called twicing and denoted by "T" in the algorithm.

$$e_i = y_i - S_{4253H}(y_i) \tag{6}$$

$$S_{4253HT}(y_i) = S_{4253H}(y_i) + S_{4253H}(e_i) \tag{7}$$

2.2 Modification of 4253HT

This section further describes the modification procedure involving mean 2 in 4253HT smoother. The average of two middle points of an arranged sequence was calculated using the arithmetic mean to determine the output of a median smoother for an even window size. Furthermore, several modifications were suggested by replacing the average of the middle with various types of means. A simulation process was performed using Eq. (2), where Step 1 was substituted with several modified means. Hence, the algorithms of four different types of mean applied in this study are expressed below:

i. Geometric Mean:
$$\bar{Y}_{Geometric} = \left(\prod_{i=1}^{n} Y_i\right)^{\frac{1}{n}} \tag{8}$$

ii. Quadratic Mean:
$$\bar{Y}_{Quadratic} = \left(\prod_{i=1}^{n} Y_i^2\right)^{\frac{1}{2}} \tag{9}$$

iii. Harmonic Mean:
$$\bar{Y}_{Harmonic} = \frac{n}{\left(\sum_{i=1}^{n} \frac{1}{Y_i}\right)} \tag{10}$$

iv. Contraharmonic Mean:
$$\bar{Y}_{Contraharmonic} = \frac{\sum_{i=1}^{n} Y_i^2}{\sum_{i=1}^{n} Y_i} \tag{11}$$

2.3 Hanning

Hanning algorithm is applied to the data series to enhance the process of smoothing [21]. There are many types of Hanning coefficient reported in previous studies, but only three types of Hanning were applied in this study. Their formulae are expressed below:

i. Tukey [6]
$$A = \left\{\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right\}$$
$$H_i = \frac{1}{4}Y_{i-1} + \frac{1}{2}Y_i + \frac{1}{4}Y_{i+1}$$

ii. Shitan [20]
$$A = \left\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right\}$$
$$H_i = \frac{1}{3}Y_{i-1} + \frac{1}{3}Y_i + \frac{1}{3}Y_{i+1}$$

iii. Husain [22]
$$A = \left\{\frac{3}{8}, \frac{2}{8}, \frac{3}{8}\right\}$$
$$H_i = \frac{3}{8}Y_{i-1} + \frac{2}{8}Y_i + \frac{3}{8}Y_{i+1}$$

These three types of Hanning were applied in the Hanning algorithm using Eq. (5), which later were recognized as Tukey, Shitan, and Husain Hanning, named after their pioneers, respectively.

2.4 Simulation Procedure

This section emphasizes the analysis of the performance of 4253HT smoother and its modified forms in eliminating noises and unwanted elements to capture the original signal through empirical simulation. Apart from that, this study also investigated the influence of Hanning on 4253HT smoother and its modified forms. Data is generally expressed as:

$$Y_t = W_t + D_t, \tag{12}$$

where Y is data or input, W is signal, and D is noise at $\mathrm{t^{th}}$ time.

The overall idea of the simulation procedure was from Conradie [23], where simulation was done 200 times in this study. Moreover, a signal used for the simulation process has been representing most real lifetime series data and has special functions such as Doppler, HeaviSine, Bumps, and Block, as follows:

i. Doppler:
$$W(i) = \sqrt{i(1-i)}\sin\left(\frac{2\pi(1+\delta)}{i+\delta}\right), where\ \delta = 0.05$$

ii. HeaviSine:
$$W(i) = 4\sin(4\pi i) - sgn(i-0.3) - sgn(0.72-i)$$

iii. Bumps:
$$W(i) = \sum h_m k\left(\frac{i-i_m}{w_m}\right)$$
where $k(i) = (1+|i|^4)^{-1}$
$$i_m = (0.1,\ 0.13,\ 0.15,\ 0.23,\ 0.25,\ 0.40,\ 0.44,\ 0.65,\ 0.76,\ 0.78,\ 0.81)$$
$$w_m = (0.005,\ 0.005,\ 0.006,\ 0.01,\ 0.01,\ 0.03,\ 0.01,\ 0.01,\ 0.005,\ 0.008,\ 0.005)$$

iv. Block:
$$W(i) = \sum h_m k(i-i_m)$$
where $i_m = (0.1,\ 0.13,\ 0.15,\ 0.23,\ 0.25,\ 0.40,\ 0.44,\ 0.65,\ 0.76,\ 0.78,\ 0.81)$
$$h_m = (4,\ -5,\ 3,\ -4,\ 5, -4.2,\ 2.1,\ 4.3,\ -3.1,\ 5.1,\ -4.2)$$

The formula of this signal function originated from Donoho [24], where their contributions were reported by Y_i [25] as their plots are shown in Figure 2. Later, all four signals were corrupted with five levels of noise.

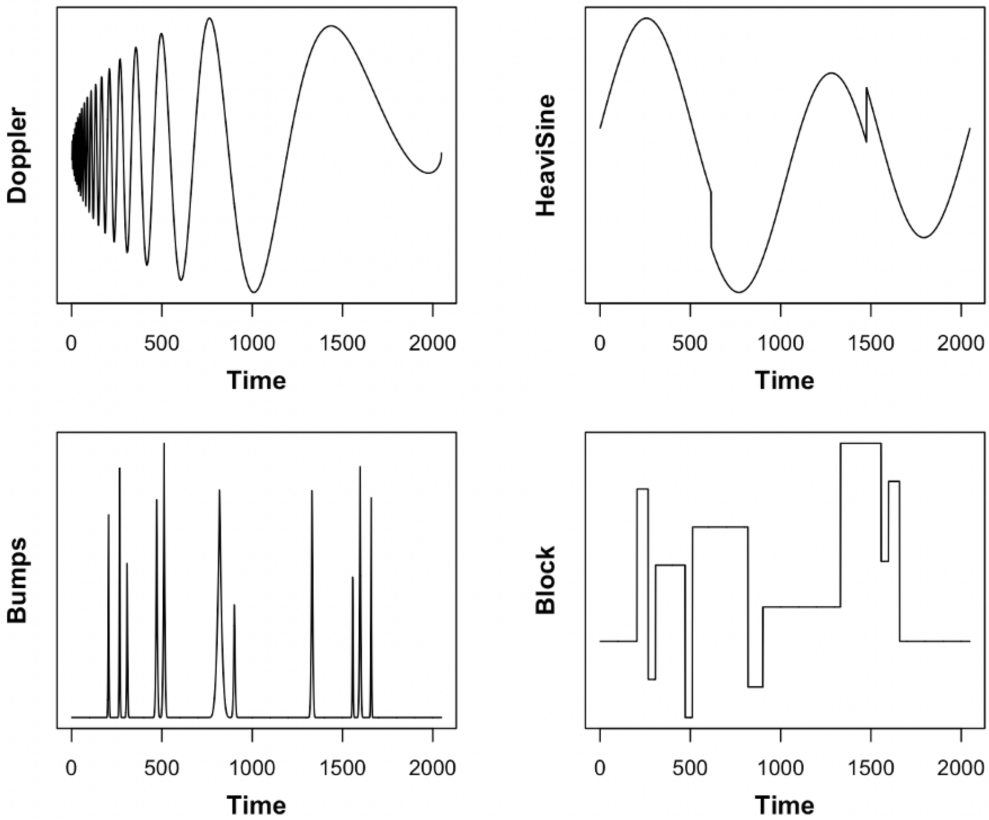

Fig. 2. Signal of doppler, heaviSine, bumps and block

Noise, D_i was created as independent, which is equally distributed random variables from two normal distributions, $D_{1i} \sim N(0, 1^2)$ and $D_{2i} \sim N(0, 5.06^2)$. Later, it was recognized as a contaminated normal noise due to the combination of both normal distributions. Furthermore, Wicklin [26] mentioned that the combination of noise from normally distributed and similar mean will create outliers in the data series.

A variance of D_{2i} was taken as 5.06^2 due to the interest particularly on data with high kurtosis. There is diverse noise density used in previous studies, such as Ahmed [27], who used noise density ranging from 20% to 70%. In this study, an increasing noise density of 10%, 25%, 50%, 75% and 90% were utilized, where the simulation of 10% contaminated normal distribution refers to 10% of values from $N(0, 5.06^2)$ distribution and approximately 90% from $N(0, 1^2)$ distribution, Jankowitz [28]. The increase of contaminated normal noise percentage improves the effectiveness of measuring 4253HT smoother performance in signal extraction from heavy noises.

A residual mean square error (RMSE) was used to determine the performance of modified 4253HT smoother, which is expressed as:

$$RMSE = \sqrt{\frac{1}{k}\sum_{j=1}^{k}\frac{1}{t}\sum_{i=1}^{t}\left(S_{ij} - Z_j\right)^2} \quad (13)$$

where S_{ij} represents an original noise-free signal, Z_j is the 4253HT smoother, and constants t and k denote as data length and number of simulations respectively. A low RMSE value signifies a good performance by the smoother in eliminating normal noise.

3. Results and Discussion

This section elaborates on the outcomes obtained by the simulation study performed for the modifications of 4253HT. For 4253HT modifications, the mean 2 was substituted with Geometric, Quadractic, Harmonic, and Contrahamonic means. Besides, 4253HT was operated together with three different types of Hanning, which are Tukey, Shitan, and Husain. Figure 3 and 4 illustrate about 10% and 90% of series for all signals were corrupted by $N(0, 5.06^2)$, while the rest was corrupted by N $(0, 1^2)$ in a random process respectively.

Based on the plots obtained, a noise that is low volatile and first noise distribution produces a sharp spike which acts as a high variation. Besides, the signal is visible and noticeable by human naked eyes. Figure 4 depicts 90% contaminated normal noise added into the signal. The signal becomes less recognizable. This is due to the extremely high density of noise distribution. Moreover, it causes the original pattern of each signal to become blurry and unclear.

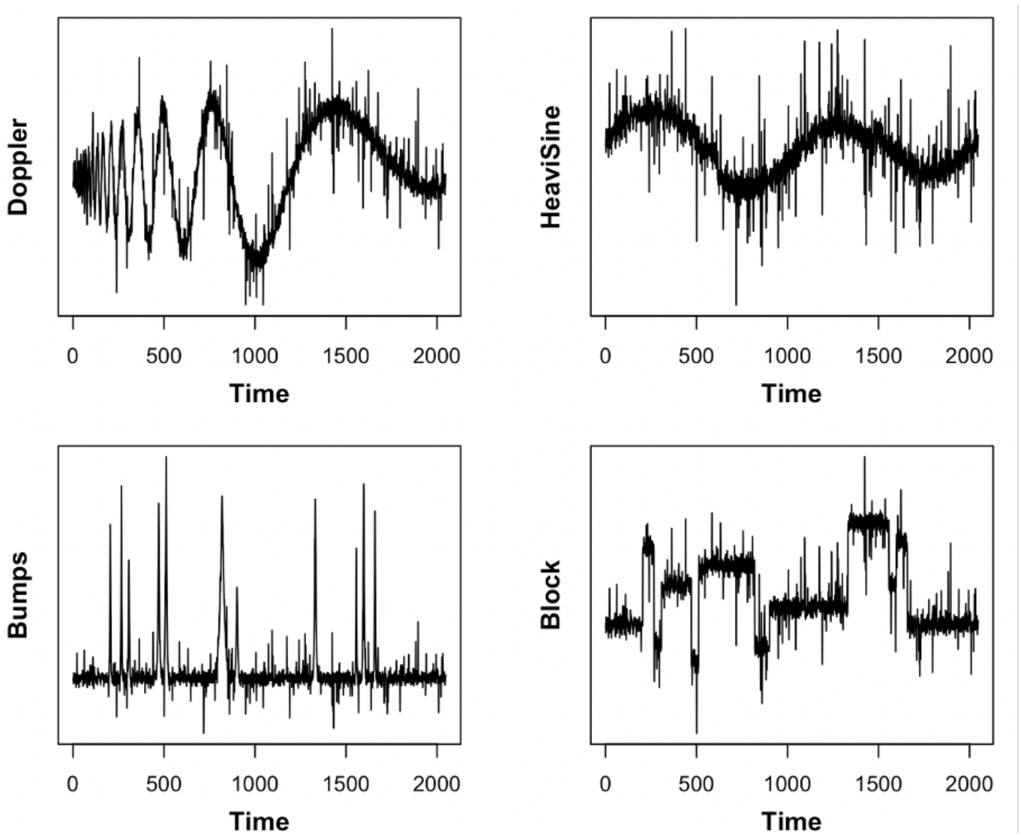

Fig. 3. Signals that were corrupted by 10% contaminated normal noise

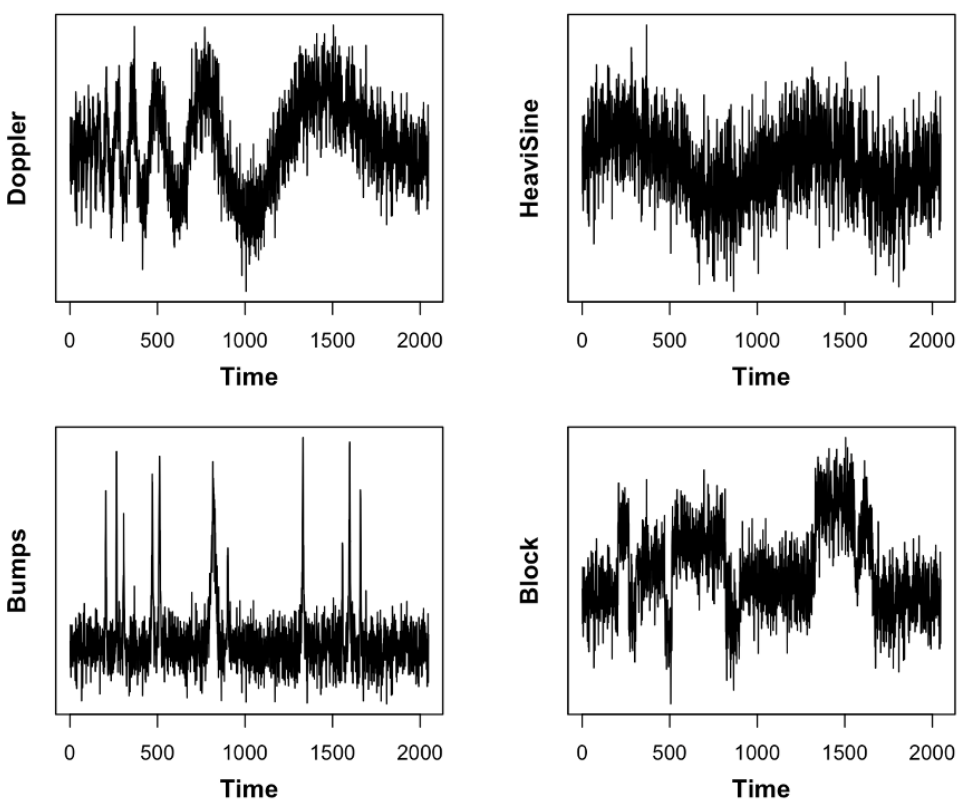

Fig. 4. Signals that were corrupted by 90% contaminated normal noise

Figure 5 and 6 portray the performance of 4253HT in capturing signals with 10% and 90% of contaminated normal noise, respectively. The red line in both figures represents the signal. The black line is a corrupted signal, while the blue line is 4253HT smoother. A lower level of noise enables us to observe the performance of 4253HT in capturing the original trail. Moreover, 4253HT was proven to be able to preserve the original pattern of signal and eliminate the spikes at once. Even with 10% noise, the original signal is still noticeable by the naked eye, as it interrupts only a negligible part of the original signal.

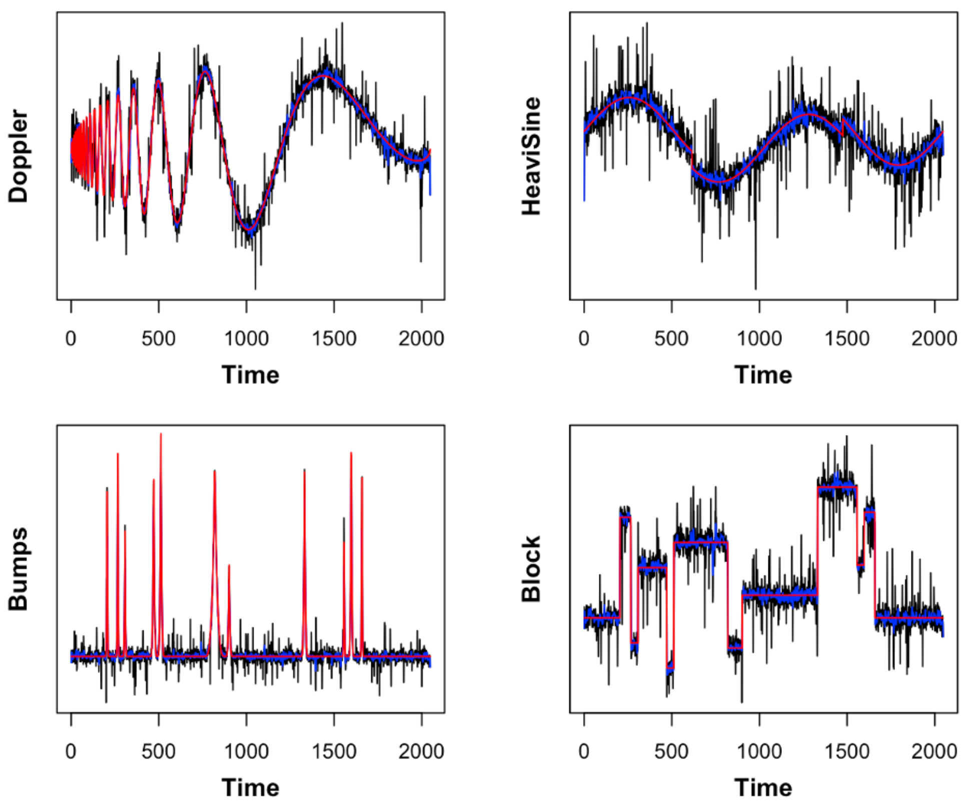

Fig. 5. Performance of 4253HT in capturing signal with 10% of contaminated normal noise

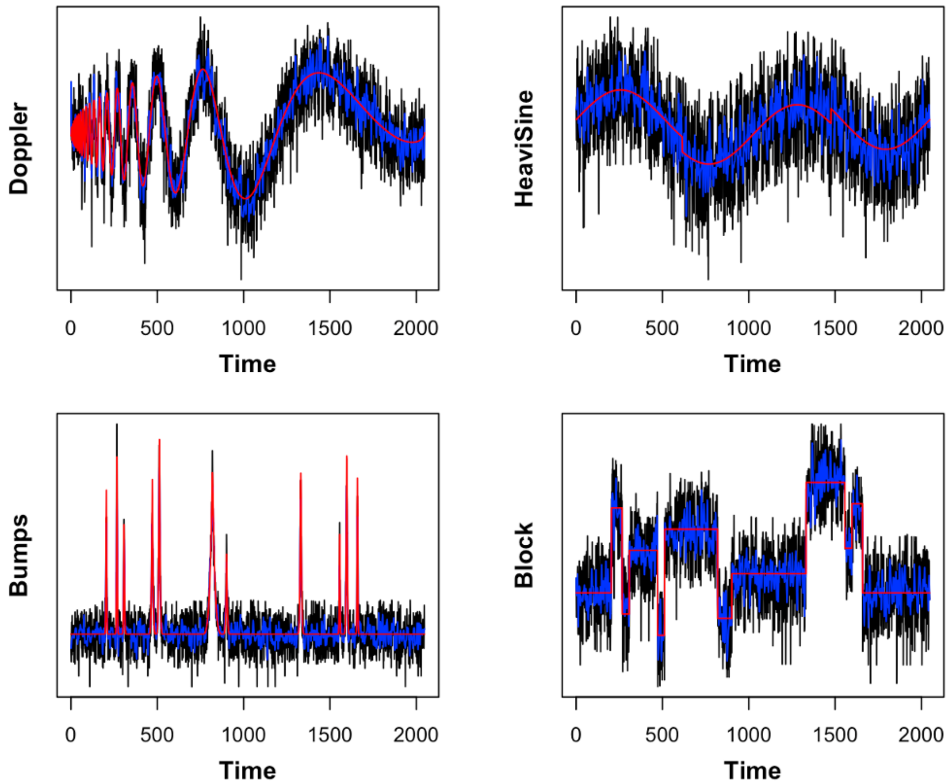

Fig. 6. Performance of 4253HT in capturing signal with 90% of contaminated normal noise

Usually, a higher level of noise causes blurry of the original signal due to the higher density of noise that corrupts the series. However, the original signal is able to be captured by 4253HT, yet the variation of smoothed line produced is slightly larger than applying 10% noise. Despite 90% noise corruption, 4253HT is assumed to be robust to all spikes and capable of preserving edges in extreme conditions.

The values of RSME for all smoother combinations involving three different types of Hanning are tabulated in Table 1 to 4. Let S1 represents the original 4253HT smoother while S2, S3, S4, and S5 represent the modified 4253HT smoother using geometric, quadratic, harmonic, and contra-harmonic means, respectively. Overall, the value of RMSE increases when the percentage of noise increases, as presented in Table 1 to 4. The best performance at each level of contaminated noise is represented by the values in bold in Table 1 to 4. In Table 1, it is obvious that Husain consistently generates the lowest RMSE values for each smoother noise level compared to other Hanning types. Moreover, the outcomes indicate that S4 (Husain) is the best smoother combination for 10% noise, while S5 (Husain) is the best for 25% to 90% noise.

Table 1
The RMSE values of each combination of smoother and hanning in doppler signal

Smoother	Hanning	Contaminated normal noise				
		10%	25%	50%	75%	90%
S1	Tukey	0.6204359	0.9352268	1.453138	1.971939	2.301408
	Shitan	0.6120274	0.9198958	1.430217	1.945379	2.275574
	Husain	0.6088039	0.9134439	1.420419	1.933964	2.264216
S2	Tukey	0.6186428	0.9420793	1.457505	1.981720	2.309059
	Shitan	0.6102903	0.9265098	1.434873	1.955539	2.282733
	Husain	0.6070596	0.9199763	1.425231	1.944197	2.271497
S3	Tukey	0.6223687	0.9293097	1.449392	1.964299	2.295039
	Shitan	0.6138921	0.9141724	1.426213	1.937334	2.269644
	Husain	0.6106554	0.9077370	1.416275	1.925662	2.258720
S4	Tukey	0.6170231	0.9491013	1.461995	1.992141	2.318247
	Shitan	0.6086879	0.9333714	1.439902	1.966516	2.291526
	Husain	**0.6054751**	0.9268252	1.430424	1.955391	2.280120
S5	Tukey	0.6244138	0.9238407	1.445609	1.957758	2.291549
	Shitan	0.6159283	0.9088271	1.422068	1.930794	2.266375
	Husain	0.6126972	**0.9024267**	**1.411942**	**1.919297**	**2.255590**

Table 2 presents the outcomes using HeaviSine Signal. Overall, the RMSE value increases as the noise percentage increases. We concluded that Husain consistently generates the lowest RMSE values for each smoother and noise level. Furthermore, S4 (Husain) performed the best for noise 10% and 25%, while S5 (Husain) indicated the greatest performance for noise 50% and 75%. In addition, S1 (Husain) worked well when HeaviSine was corrupted by 90% contaminated noise.

Table 2
The RMSE values of each combination of smoother and hanning in heaviSine signal

Smoother	Hanning	Contaminated normal noise				
		10%	25%	50%	75%	90%
S1	Tukey	0.5833515	0.9592304	1.519432	2.204500	2.662109
	Shitan	0.5755973	0.9460522	1.498678	2.174885	2.630857
	Husain	0.5722021	0.940596	1.489448	2.162291	2.616957
S2	Tukey	0.5827582	0.9579069	1.524801	2.210877	2.663042
	Shitan	0.5750035	0.9447272	1.504502	2.180087	2.631644
	Husain	0.5716048	0.9393561	1.495289	2.167393	2.617778
S3	Tukey	0.5839236	0.9609322	1.518176	2.200810	2.662310
	Shitan	0.5762366	0.9477070	1.496633	2.171756	2.631218
	Husain	0.5728489	0.9422131	1.487279	2.159134	2.617472
S4	Tukey	0.5822110	0.9575566	1.529953	2.218481	2.665192
	Shitan	0.5744697	0.9445592	1.509629	2.187088	2.633527
	Husain	**0.5710969**	**0.9391946**	1.500124	2.173827	2.619600
S5	Tukey	0.5844434	0.9634025	1.518646	2.198983	2.665201
	Shitan	0.5769406	0.9503861	1.496647	2.170617	2.634102
	Husain	0.5735609	0.9448943	1.487053	2.158468	2.620384

The results using the Bumps signal are presented in Table 3. Generally, the value of RMSE increases as the noise percentage increases. Based on the table, it is assumed that Husain consistently provides the lowest RMSE values for each smoother and noise level. Overall, S5 (Husain) is proven as the best smoother for each noise level. Table 4 displays the outcomes using a Block signal. Similar to other signals, the RMSE value increases when the noise percentage increases. Moreover, S5 (Tukey) recorded the best performance for noise 10% and 25%, while S5 (Husain) presented the greatest work on noise 50% to 90%.

Table 3
The RMSE values of each combination of smoother and hanning in bumps signal

Smoother	Hanning	Contaminated normal noise				
		10%	25%	50%	75%	90%
S1	Tukey	1.027762	1.258685	1.626650	2.202891	2.519408
	Shitan	1.010343	1.239167	1.601718	2.173686	2.491604
	Husain	1.003978	1.232009	1.591992	2.161615	2.480320
S2	Tukey	1.037794	1.268001	1.635582	2.210438	2.533070
	Shitan	1.021396	1.248935	1.611192	2.181480	2.504987
	Husain	1.015518	1.242033	1.601757	2.169425	2.493858
S3	Tukey	1.018370	1.250100	1.618514	2.195970	2.508221
	Shitan	1.000196	1.230165	1.593378	2.166317	2.480122
	Husain	0.993556	1.222782	1.583496	2.154111	2.468429
S4	Tukey	1.048089	1.278987	1.645110	2.219341	2.547386
	Shitan	1.032470	1.259147	1.621252	2.190862	2.519350
	Husain	1.027139	1.252517	1.612056	2.179041	2.508213
S5	Tukey	1.009000	1.242006	1.609943	2.189079	2.498889
	Shitan	0.990365	1.221775	1.584594	2.159382	2.470656
	Husain	**0.983520**	**1.214276**	**1.574365**	**2.147341**	**2.459071**

Table 4
The RMSE values of each combination of smoother and hanning in block signal

Smoother	Hanning	Contaminated normal noise				
		10%	25%	50%	75%	90%
S1	Tukey	0.7264022	0.9726809	1.505257	2.083662	2.399998
	Shitan	0.7383431	0.9735388	1.495668	2.065140	2.374892
	Husain	0.7457611	0.9757886	1.492805	2.057935	2.364063
S2	Tukey	0.7276354	0.9748223	1.510327	2.092598	2.412773
	Shitan	0.7395542	0.9758353	1.501288	2.073560	2.388241
	Husain	0.7470062	0.9783119	1.498820	2.066240	2.377607
S3	Tukey	0.725336	0.9707454	1.501039	2.076991	2.389487
	Shitan	0.7372968	0.9713156	1.490766	2.058543	2.364175
	Husain	0.7446959	0.9733298	1.487658	2.051368	2.353156
S4	Tukey	0.7290017	0.9774172	1.516015	2.103619	2.426299
	Shitan	0.7409003	0.9785719	1.507356	2.083712	2.402237
	Husain	0.7484076	0.9811662	1.505381	2.076046	2.391850
S5	Tukey	**0.7244260**	**0.9692336**	1.497225	2.072684	2.382939
	Shitan	0.7363407	0.9695773	1.486503	2.054226	2.357564
	Husain	0.7437176	0.9713874	**1.483018**	**2.047055**	**2.346350**

Furthermore, Figure 7 depicts residual plots of all signals using only S5 (Husain) smoother with 90% contaminated noise. This smoother was selected as it works the best on all signals. A residual plot is important to determine the stability of smoother through the observation of residual tabulation pattern. A smoother is determined as stable when no pattern is observed (random), while a smoother is not stable if a pattern is found. Overall, the residual plots depicted in Figure 7 indicate no pattern. Besides, there are some spikes occurred due to the randomness of noise.

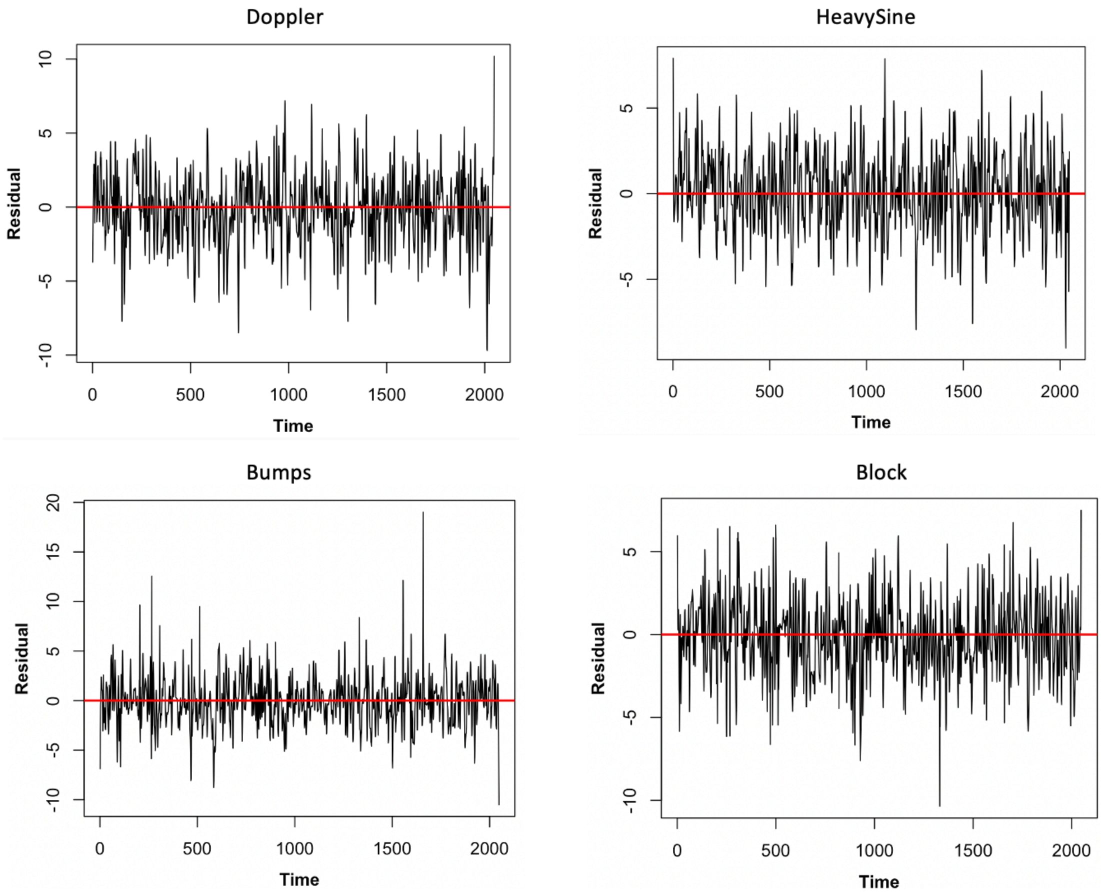

Fig. 7. Residual plots for doppler, heavySine, bumps, and block signals

Table 5 summarizes the performance of Doppler, HeaviSine, Bumps, and Block signals. Generally, the maximum and minimum values of all signals are fairly close, except for the Block signal. Besides, all signals obtain approximately near values of standard deviation. Hence, all these measurements proved that 4253HT smoother and all its modified forms are able to provide stable smoothing values for all four signals. In addition, there are few spikes found, which are just a nature of noise randomness that is added to signals that affect the closeness of smoother with the original signal.

Table 5
The performance of doppler, heavySine, bumps, and block signals

Signal	Residual Summary Statistics		
	Minimum	Maximum	Standard Deviation
Doppler	-9.69600	10.19700	2.247440
HeaviSine	-9.02320	7.90400	2.208919
Bumps	-10.50925	14.25854	2.414172
Block	-10.35540	19.01608	2.134744

4. Conclusions

The main purpose of this study is to identify the most relevant combination of smoother with different signals, including Doppler, HeaviSine, Bumps, and Block signals, by performing a simulation study. Besides, it also aims to determine the performance of modified 4253HT smoother operating with other smoothers and different noise levels using an RMSE as an evaluator.

An overall conclusion is depicted in Table 6. According to the outcomes of this study, Hanning Husain is proven as the best Hanning overall, while Hanning Tukey works the best on Block signals at lower noise levels (10% and 25%). On the other hand, Hanning Shitan records the worst performance on all signals. Besides, the best smoother amongst all is S5 Husain, which is the combination of Contraharmonic Mean with Hanning Husain. Thus, it is suggested to utilize this smoother in further analysis, mainly in forecasting, to provide accurate values and patterns for prediction.

Table 6
Summary of performance of 4253HT smoother and its modified forms in signal recovery for doppler, heavySine, bumps, and block signals

Signal	Noise				
	10%	25%	50%	75%	90%
Doppler	S4 (Husain)	S5 (Husain)	S5 (Husain)	S5 (Husain)	S5 (Husain)
HeaviSine	S4 (Husain)	S4 (Husain)	S5 (Husain)	S5 (Husain)	S1 (Husain)
Bumps	S5 (Husain)	S5 (Husain)	S5 (Husain)	S5 (Husain)	S5 (Husain)
Block	S5 (Tukey)	S5 (Tukey)	S5 (Husain)	S5 (Husain)	S5 (Husain)

Acknowledgement
This research was funded by Universiti Teknologi MARA, Pahang Branch, Malaysia (Grant number: 600-TNCPI 5/3/DDN (06) (014/2022)).

References
[1] Nik Badrul Alam, Nik Muhammad Farhan Hakim, Nazirah Ramli, Adie Safian Ton Mohamed, and Noor Izyan Mohamad Adnan. "Integration of 4253HT Smoother with Intuitionistic Fuzzy Time Series Forecasting Model." *Pakistan Journal of Statistics & Operation Research* 18, no. 4 (2022). https://doi.org/10.18187/pjsor.v18i4.4212

[2] Nik Muhammad Farhan Hakim Nik Badrul Alam, Nazirah Ramli, Adie Safian Ton Mohamed, and Noor Izyan Mohamad Adnan. "A Hybrid Fuzzy Time Series Forecasting Model with 4253HT Smoother." *Applied Mathematics and Computational Intelligence*, November, (2022):325–335.

[3] AZMI, NURUL NISA'BINTI KHAIROL. "EFFICIENCY OF 4253HT SMOOTHERS IN EXTRACTING SIGNAL FROM NOISE AND THEIR APPLICATIONS IN FORECASTING." (2019).

[4] Gabbouj, Moncef, Edward J. Coyle, and Neal C. Gallagher. "An overview of median and stack filtering." *Circuits, Systems and Signal Processing* 11 (1992): 7-45. https://doi.org/10.1007/BF01189220

[5] Bernholt, Thorsten, Roland Fried, Ursula Gather, and Ingo Wegener. "Modified repeated median filters." *Statistics and Computing* 16 (2006): 177-192. https://doi.org/10.1007/s11222-006-8449-1

[6] Hans-Georg, Beyer. "Tukey, John W.: Exploratory Data Analysis. Addison-Wesley Publishing Company Reading, Mass. — Menlo Park, Cal., London, Amsterdam, Don Mills, Ontario, Sydney 1977, XVI, 688 s." *Biometrical Journal* 23 (4), (1981): 413–14. https://doi.org/10.1002/bimj.4710230408

[7] Velleman, Paul F. "Definition and comparison of robust nonlinear data smoothing algorithms." *Journal of the American Statistical Association* 75, no. 371 (1980): 609-615. https://doi.org/10.2307/2287657

[8] Dibeklioğlu, Hamdi, Zakia Hammal, and Jeffrey F. Cohn. "Dynamic multimodal measurement of depression severity using deep autoencoding." *IEEE journal of biomedical and health informatics* 22, no. 2 (2017): 525-536. https://doi.org/10.1109/jbhi.2017.2676878

[9] Stern, L. and Lightfoot, D., Automated outbreak detection: a quantitative retrospective analysis. *Epidemiology & Infection*, *122*(1), (1999):103-110. https://doi.org/10.1017/s0950268898001939

[10] Verma, Kesari, Bikesh Kumar Singh, and A. S. Thoke. "An enhancement in adaptive median filter for edge preservation." *Procedia Computer Science* 48 (2015): 29-36. https://doi.org/10.1016/j.procs.2015.04.106

[11] Hird, Jennifer N., and Gregory J. McDermid. "Noise reduction of NDVI time series: An empirical comparison of selected techniques." *Remote Sensing of Environment* 113, no. 1 (2009): 248-258. https://doi.org/10.1016/j.rse.2008.09.003

[12] Mozny, Martin, Radim Tolasz, Jiri Nekovar, Tim Sparks, Mirek Trnka, and Zdenek Zalud. "The impact of climate change on the yield and quality of Saaz hops in the Czech Republic." *Agricultural and Forest Meteorology* 149, no. 6-7 (2009): 913-919. https://doi.org/10.1016/j.agrformet.2009.02.006

[13] Ashelford, Kevin E., Martin J. Day, Mark J. Bailey, Andrew K. Lilley, and John C. Fry. "In situ population dynamics of bacterial viruses in a terrestrial environment." *Applied and Environmental Microbiology* 65, no. 1 (1999): 169-174. https://doi.org/10.1128/aem.65.1.169-174.1999

[14] Reynolds, Richard W. *A monthly averaged climatology of sea surface temperature*. Vol. 31. US Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, 1982. https://doi.org/10.1175/1520-0442(1988)001<0075:ARTGSS>2.0.CO;2

[15] Polasek, Wolfgang. "Exploring business cycles using running medians." *Computational Statistics & Data Analysis* 2, no. 1 (1984): 51-70. https://doi.org/10.1016/0167-9473(84)90032-x

[16] Bristol, David R. "Delta: The true clinically significant difference to be detected." *Drug information journal* 29, no. 1 (1995): 33-36. https://doi.org/10.1177/009286159502900105

[17] Jin, Zhenyu, and Bing Xu. "A novel compound smoother—RMMEH to reconstruct MODIS NDVI time series." *IEEE Geoscience and Remote Sensing Letters* 10, no. 4 (2013): 942-946. https://doi.org/10.1109/lgrs.2013.2253760

[18] Janosky, J. E., T. R. Pellitieri, and Q. M. Al-Shboul. "The need for a revised lower limit for the 4253H, Twice nonparametric smoother." *Statistics & probability letters* 32, no. 3 (1997): 269-272. https://doi.org/10.1016/s0167-7152(96)00083-1

[19] Sargent, Jonathan, and Anthony Bedford. "Improving Australian Football League player performance forecasts using optimized nonlinear smoothing." *International Journal of Forecasting* 26, no. 3 (2010): 489-497. https://doi.org/10.1016/j.ijforecast.2009.10.003

[20] Shitan, Mahendran, and Turaj Vazifedan. *Exploratory data analysis for almost anyone*. UPM Press, 2011.

[21] Mills, Terence C. *Time series techniques for economists*. Cambridge University Press, 1990.

[22] Husain, Q. N. "Modification of Tukey's Smoothing: Techniques for Extreme Data." *University Putra Malaysia* (2017).

[23] Conradie, W. J., T. De Wet, and Maria Dorothea Jankowitz. "Performance of nonlinear smoothers in signal recovery." *Applied Stochastic Models in Business and Industry* 25, no. 4 (2009): 425-444. https://doi.org/10.1002/asmb.774

[24] Donoho, David L., and Iain M. Johnstone. "Ideal spatial adaptation by wavelet shrinkage." *biometrika* 81, no. 3 (1994): 425-455. https://doi.org/10.1093/biomet/81.3.425

[25] Yi, Ting-Hua, Hong-Nan Li, and Xiao-Yan Zhao. "Noise smoothing for structural vibration test signals using an improved wavelet thresholding technique." *Sensors* 12, no. 8 (2012): 11205-11220. https://doi.org/10.3390/s120811205

[26] Wicklin, Rick. *Simulating data with SAS*. SAS Institute, 2013.

[27] Sudan, Sudan. "Median filter performance based on different window sizes for salt and pepper noise removal in gray and RGB images." *International Journal of Signal Processing, Image Processing and Pattern Recognition* 8, no. 10 (2015): 343-352. https://doi.org/10.14257/ijsip.2015.8.10.34

[28] Jankowitz, Maria Dorothea. "Some statistical aspects of LULU smoothers." PhD diss., Stellenbosch: University of Stellenbosch, 2007.