



**TITLE OF FYP PROPOSAL Nurul Nisa' Khairol Azmi**

**(GS123456)**

**Nora Basir (GS111111)**

**Noorezatty Mohd Yusop (GS101010)**

**Supervisor:**

**Jamilah Othman December 2019**

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Background of Study . . . . .	1
1.2	Problem Statement . . . . .	2
1.3	Research Objectives . . . . .	2
<b>2</b>	<b>Literature Review</b>	<b>4</b>
2.1	Introduction . . . . .	4
2.2	Properties of Smoother . . . . .	4
2.3	Conclusion . . . . .	5
<b>3</b>	<b>Methodology</b>	<b>7</b>
3.1	Introduction . . . . .	7
3.2	Compound Smoother-4253HT . . . . .	7
	<b>References</b>	<b>12</b>

# List of Tables

3.1	Summary of performance of modified 4253HT in extracting signal of sinusoidal function with noise added . . . . .	11
-----	------------------------------------------------------------------------------------------------------------------	----

# List of Figures

3.1	Algorithm of 4253HT . . . . .	9
3.2	Signal of Doppler . . . . .	10

# Chapter 1

## Introduction

### 1.1 Background of Study

Smoothing is more of a curve fitting whereby the main purpose is tracing the trend from a set of data series blurred by noise. In a data series, trends provide the direction to choose appropriate method of estimation. Smoothing data series does not necessarily have to be well fitted, but most importantly it has an ability to reduce noise so that overall picture regarding global behavior of data series can be captured. The pattern extracted from smoothing process is able to provides some guideline on a suitable modelling estimation for forecasting purpose. Smoothing does not only helps in curve fitting but also very useful in determining future values by eliminating non-well behave noise.

Smoothing by definition varies according to the fields of interest. Some studies, use the term filtering to refer to smoothing for example (Ataman, Aatre, & Wong, 1981), (Bovik, Huang, & Munson Jr, 1983), (Gabbouj, Coyle, & Gallagher Jr, 1992), (Zeng, 1994) and (Miao & Jiang, 2013). In order to avoid any confusion, the term smoothing is used consistently throughout this research. The main concern of smoothing is to

capture underlying pattern by removing unwanted noise from the data series.

## 1.2 Problem Statement

Linear smoother is optimal to eliminate Gaussian noise and track trends that are common in practice, (Bernholt, Fried, Gather, & Wegener, 2006). However, noise of high volatility tends to mask the general picture of a data series. The existence of non well-behaved noise violated the assumptions of linear model. Usually, least square estimation which is well known for its poor performance in the presence of outliers or long-tailed distribution data is used.

According to (Venetsanopoulos & Pitas, 1990), linear smoothers also have a high tendency to blur important features and lack of the ability to remove impulsive noise. Not only that, linear smoothers are highly vulnerable to outliers and could not deal well with nonlinearity in a data series. Blurry edge which leads to the lost of important information is actually due to the sudden changes in a series, (Bernholt et al., 2006).

Due to its ability to remove non-Gaussian noise from a data series, median smoother is usually the favored smoothing tools. Unfortunately, median smoother tends to over smoothed a data series since it eliminates Gaussian noise too.

## 1.3 Research Objectives

Since there is an opportunity for improvement in compound smoother, some modifications to running median of span size 42 is suggested in this study. The existing study only focuses on noise with long tailed distribution. The pattern with small portion of contaminated can easily be observed with naked eyes. Unfortunately, for

data with high fluctuation, the signal might mix up with heavy noise, making it hard to capture any possible trends. In this research, the performance of smoothers in highly volatile data is compared and evaluated. This research provides some values added to the existing study and also motivates future research to expand the idea this study addresses for a better solution. Guided by the earlier discussion, the purposes of study are summarized as follows;

1. To modify existing compound smoothers
2. To determine the stability of modified compound smoothers towards block pulse.
3. To evaluate the performance of modified compound smoother via simulation procedure with higher percentage of contaminated normal noise for sinusoidal, Doppler, Bumps, Blocks and Heavy Sine function.
4. To formulate a strategy of forecasting by extracting deterministic components in data series.
5. To apply the proposed modified smoother to financial, environment and agriculture data.

# Chapter 2

## Literature Review

### 2.1 Introduction

Many theories and research which extended from Tukey's (1977) idea, evolved over the past few decades. Although the literatures cover a variety Tukey's approaches in method of estimation, these reviews only focused on smoothing techniques that have emerged since they have been introduced. In this chapter, some properties of good smoother are discussed in terms of monotonicity, effectiveness, consistency and stability. A review on median smoother particularly on its general behavior towards Gaussian and non-Gaussian noise and deterministic properties is included. A brief summary regarding various type of compound smoothers are also discussed. This chapter also highlights the types of means used in the modification of compound smoother.

### 2.2 Properties of Smoother

A few issues or concerns are regarding the properties of smoothers and measuring the smoothness of a smoothed series. Estimating signal by smoother does not require strict mathematical or statistical assumptions to be full filled. (Jankowitz, 2007)



discusses the properties of good smoother extensively as follows;

- **Effectiveness**

An effective smoother, let say  $S$ , if for each  $X_t$ ,  $S(X_t)$  plays as a signal and  $[(I - S)X_t = X_t - S(X_t)]$  is noise where  $I$  is the identity operator, (Rohwer, 2005). The main objective of smoothers is to ensure that it is an effective smoother. In real situation, determining whether a smoother is effective can be very difficult to be done since a signal is unknown. So, the main purpose is to obtain a good estimator for the signal. To do this, the unwanted noise can be reduced for any given signals. In the case of this study, the effectiveness of smoother is measured via simulation studies. The procedures of simulations are elaborated in Chapter 3.

- **Consistency**

Consistent smoother describes a smoother with the ability to maintain the main features of a signal and equating noise to 0, (Rohwer, 2005). The terms of idempotency and co-idempotency are highly related and defined as

## 2.3 Conclusion

Some elements of smoother that have been discussed were incorporated with the proposed modified compound smoother, and discussed later in Chapters 4 and 5. The median smoother is described extensively in terms of strength and weakness. Deliberation on some statistical and deterministic properties of median smoother for odd span size, is also included. The idea of compound smoother is extended for

improvement and also to provide more options for further analysis. A review on the types of means provides an insight into possible improvement for modification.

# Chapter 3

## Methodology

### 3.1 Introduction

In this chapter, the method of compound smoother specifically 4253HT is discussed extensively. Simulation was performed by generating deterministic functions added with noise. The procedure of simulation is based on (Donoho & Johnstone, 1994) and (Conradie, De Wet, & Jankowitz, 2009). The existing procedure only take into consideration of extracting Gaussian noise or 10% contaminated Gaussian noise. In this study, performance of the compound smoother is widened by measuring the success of recovering the signal from heavy noise. The amount of contaminated normal from non-Gaussian noise was increased to 25%, 50% and 75%.

### 3.2 Compound Smoother-4253HT

Compound smoother of 4253HT is a combination of algorithm consisting of running median of span size four, two, five and three, followed by Hanning and re-smoothed the rough. Let a temporal data  $\mathbf{X}$  be a doubly-infinite sequence of real data  $\{X_{-N}, \dots, X_{t-1}, X_t, X_{t+1}, \dots, X_N\}$ , (Mallow, 1980). Let  $S$  be a smoother that works on  $\mathbf{X}$  to generate a data series  $S(\mathbf{X}_t)$ , smoothed values.

The computation of compound smoother started with a median smoother of window size four, re-centered by median smoother of window size two. Then, the smoothed values are re-smoothed by a running median of span size five and next by running median of span size three. Subsequently, the values are computed using running weighted average. The result of this smoothing is polished by computing the rough or residual, applying the same algorithm of smoothing and adding the result to first pass. Figure 3.1 shows the algorithm involve in compound smoother 4253HT.

In general, the flow of computation are as follows;

1. Perform running medians of span size four;

$$\begin{aligned}
S_1(\mathbf{X}_t) &= \text{median}(X_{t-2}, X_{t-1}, X_t, X_{t+1}) \\
&= \text{median} [X_{(t-2)}^*, X_{(t-1)}^*, X_{(t)}^*, X_{(t+1)}^*] \\
&= \text{mean} [X_{(t-1)}^*, X_{(t)}^*] \\
&= \frac{1}{2} [X_{(t-1)}^* + X_{(t)}^*] \tag{3.1}
\end{aligned}$$

where  $X_{t-i}^*$  are the ordered observations in a window of size 4,  $i = \{-2, -1, 0, 1\}$ .

2. Repeat the running medians of span size five and then running median of span size three. A span of five periods are as follows;

$$S_3(\mathbf{X})_t = \text{median} [S_2(\mathbf{X}_{t-2}), S_2(\mathbf{X}_{t-1}), S_2(\mathbf{X}_t), S_2(\mathbf{X}_{t+1}), S_2(\mathbf{X}_{t+2})]. \tag{3.2}$$

### Simulation Procedure for Doppler Signal

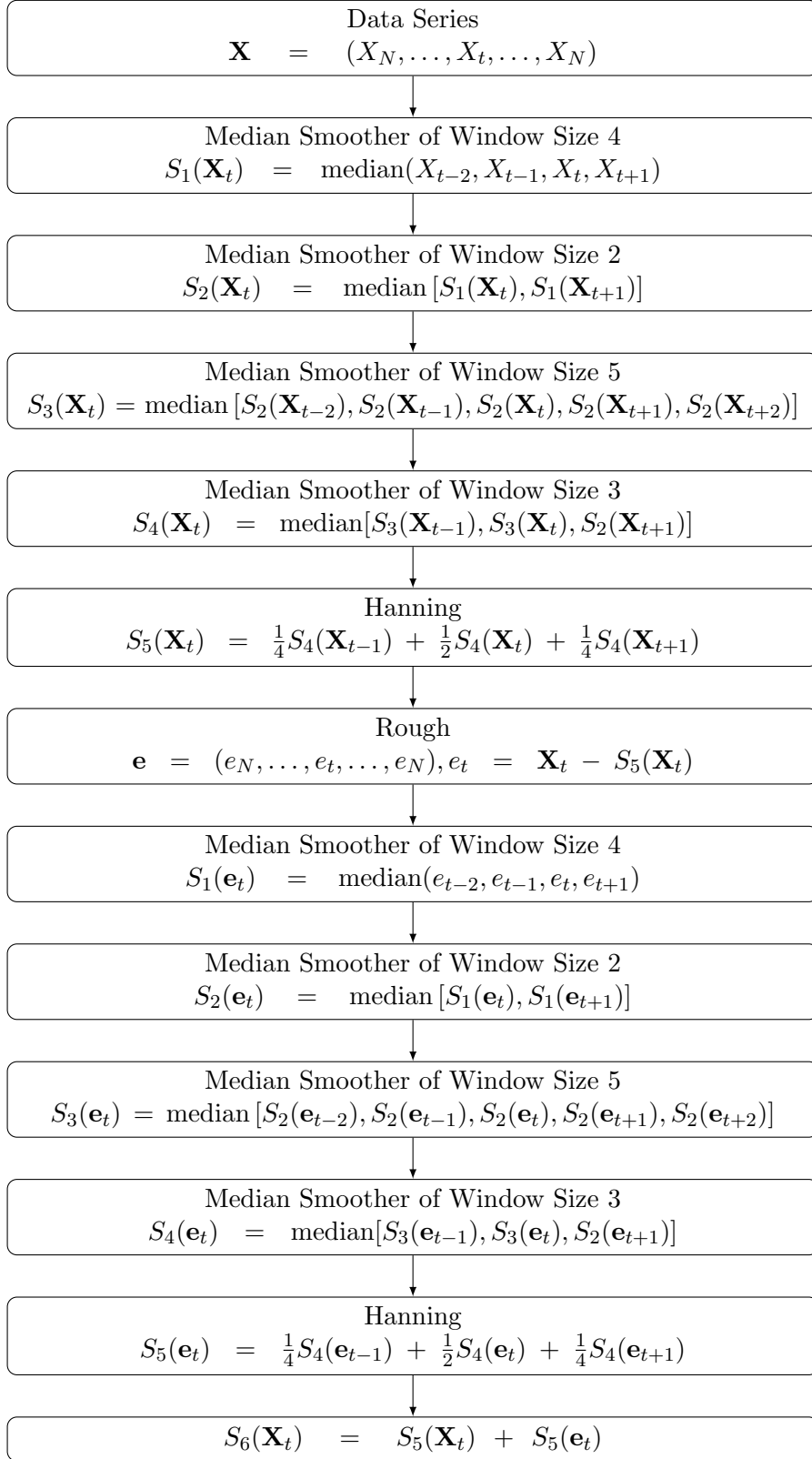


Figure 3.1: Algorithm of 4253HT

Doppler is a sine function begin with small and fast waves which extendedly become larger and slower as  $t$  increases. The function of Doppler can be expressed as;

$$\mu_t = [t(1 - t)]^{\frac{1}{2}} \sin[2\pi(1 + \epsilon)/(t + \epsilon)], \quad \epsilon = 0.05. \quad (3.3)$$

Since the number of observations in each function is  $n = 2048$ , hence  $(t_1, \dots, t_n) = (\frac{1}{n}, \dots, 1)$ .

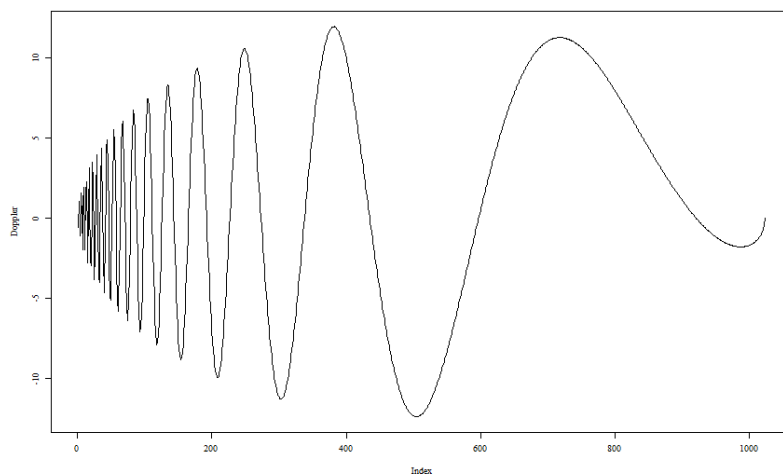


Figure 3.2: Signal of Doppler

Figure 3.2 depicts an example of Doppler function.

Table 3.1: Summary of performance of modified 4253HT in extracting signal of sinusoidal function with noise added

Measurement error	Frequency	Noise			
		10%	25%	50%	75%
Regression coefficient	Low	Geometric	Geometric	Geometric	Geometric
	Moderate	Contra harmonic	Harmonic	Harmonic	Geometric
	High	Geometric	Geometric	Geometric	Geometric
EIMSE	Low	Quadratic	Quadratic	Quadratic	Contra Harmonic
	Moderate	Contra Harmonic	Harmonic	Harmonic	Contra harmonic
	High	Contra harmonic	Quadratic	Quadratic	Contra harmonic
Variation Reduction	Low	Contra harmonic	Contra harmonic	Contra harmonic	Contra harmonic
	Moderate	Contra harmonic	Contra harmonic	Contra harmonic	Contra harmonic
	High	Geometric	Geometric	Geometric	Geometric

# References

- Ataman, E., Aatre, V., & Wong, K. (1981). Some statistical properties of median filters. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(5), 1073–1075.
- Bernholt, T., Fried, R., Gather, U., & Wegener, I. (2006). Modified repeated median filters. *Statistics and Computing*, 16(2), 177–192.
- Bovik, A. C., Huang, T. S., & Munson Jr, D. C. (1983). A generalization of median filtering using linear combinations of order statistics. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31(6), 1342–1350.
- Conradie, W., De Wet, T., & Jankowitz, M. D. (2009). Performance of nonlinear smoothers in signal recovery. *Applied Stochastic Models in Business and Industry*, 25(4), 425–444.
- Donoho, D. L., & Johnstone, J. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3), 425–455.
- Gabbouj, M., Coyle, E. J., & Gallagher Jr, N. C. (1992). An overview of median and stack filtering. *Circuits, Systems and Signal Processing*, 11(1), 7–45.
- Jankowitz, M. D. (2007). *Some statistical aspects of lulu smoothers* (Unpublished doctoral dissertation). Stellenbosch University, Cape Town.
- Miao, Z., & Jiang, X. (2013). Weighted iterative truncated mean filter. *IEEE Transactions on Signal Processing*, 61(16), 4149–4160.
- Rohwer, C. (2005). Lulu-smoothers, signals and ambiguity. *Nonlinear Smoothing and*



*Multiresolution Analysis*, 21–30.

Venetsanopoulos, A., & Pitas, I. (1990). *Nonlinear digital filters* (1st ed.). New York: Springer US.

Zeng, B. (1994). Convergence properties of median and weighted median filters. *IEEE Transactions on Signal Processing*, 42(12), 3515–3518.

# Appendix A

Code: \_\_\_\_\_ Age: \_\_\_\_\_ Female  Male  Education/Profession: \_\_\_\_\_ / \_\_\_\_\_

## SOCIAL ANXIETY QUESTIONNAIRE FOR ADULTS (SAQ)

(Caballo, Salazar, Arias, Irurtia, Calderero, & CISO-A Research Team, 2010)

Below are a series of social situations that may or may not cause you UNEASE, STRESS or NERVOUSNESS. Please place an "X" on the number next to each social situation that best reflects your reaction, where "1" represents no unease, stress or nervousness and "5" represents very high or extreme unease stress, or nervousness.

If you have never experienced the situation described, please **imagine** what your level of UNEASE, STRESS, or NERVOUSNESS might be if you were in that situation and rate how you imagine you would feel by placing an "X" on the corresponding number.

### LEVEL OF UNEASE, STRESS OR NERVOUSNESS

Not at all or very slight 1	Slight 2	Moderate 3	High 4	Very high or extremely high 5
-----------------------------------	-------------	---------------	-----------	-------------------------------------

Please rate all the items and do so **honestly**; do not worry about your answer because there are no right or wrong ones. Thank you very much for your collaboration.

1. Greeting someone and being ignored	1	2	3	4	5
2. Having to ask a neighbor to stop making noise	1	2	3	4	5
3. Speaking in public	1	2	3	4	5
4. Asking someone I find attractive for a date	1	2	3	4	5
5. Complaining to the waiter about my food	1	2	3	4	5
6. Feeling watched by people I find attractive	1	2	3	4	5
7. Participating in a meeting with people in authority	1	2	3	4	5
8. Talking to someone who isn't paying attention to what I am saying	1	2	3	4	5
9. Refusing when asked to do something I don't like doing	1	2	3	4	5
10. Making new friends	1	2	3	4	5
11. Telling someone that they have hurt my feelings	1	2	3	4	5
12. Having to speak in class, at work, or in a meeting	1	2	3	4	5
13. Maintaining a conversation with someone I've just met	1	2	3	4	5
14. Expressing my annoyance to someone that is picking on me	1	2	3	4	5
15. Greeting each person at a social meeting when I don't know most of them	1	2	3	4	5
16. Being teased in public	1	2	3	4	5
17. Talking to people I don't know at a party or a meeting	1	2	3	4	5
18. Being asked a question in class by the teacher or by a superior in a meeting	1	2	3	4	5
19. Looking into the eyes of someone I have just met while we are talking	1	2	3	4	5
20. Being asked out by a person I am attracted to	1	2	3	4	5
21. Making a mistake in front of other people	1	2	3	4	5
22. Attending a social event where I know only one person	1	2	3	4	5
23. Starting a conversation with someone I find attractive	1	2	3	4	5
24. Being reprimanded about something I have done wrong	1	2	3	4	5
25. While having dinner with colleagues, classmates or workmates, being asked to speak on behalf of the entire group	1	2	3	4	5
26. Telling someone that their behavior bothers me and asking them to stop	1	2	3	4	5
27. Asking someone I find attractive to dance	1	2	3	4	5
28. Being criticized	1	2	3	4	5
29. Talking to a superior or a person in authority	1	2	3	4	5
30. Telling someone I am attracted to that I would like to get to know them better	1	2	3	4	5