

Separation of Gravity Anomaly Data considering Statistical Independence among Signals

– Application to Severely Contaminated Data Obtained by
Prototype Mobile Gravimeter –

信号データ間の統計的独立性を用いた重力異常データの分離
– 可搬型重力測定機の試作機による観測記録への適用 –

Prem Prakash Khatri
ID: 37-106939

Signature	Date	Seal
Advisor:		
Co-Advisor:		

A Thesis submitted in
partial fulfilment of the requirements for
the Master's Degree to
Department of Civil Engineering
University of Tokyo



August 10, 2012

Contents

1	Introduction and Background	1
1.1	Sub-surface investigation	2
1.1.1	Boring	2
1.1.2	Seismic method	2
1.1.3	Microtremors method	2
1.1.4	Magnetic method	3
1.1.5	Gravity method	3
1.2	Combination of subsurface survey methods	4
1.3	Mobility of Gravity method	5
1.4	Severe Noise contamination in Gravity data	5
1.5	Noise Reduction and Data separation	6
1.5.1	Low Pass Filter	6
1.5.2	Second Order Blind Identification (SOBI)	7
1.5.3	Independent Component Analysis (ICA)	7
1.6	Organization of thesis	7
2	Objectives	8
2.1	Objectives	9
2.1.1	Propose ICA for de-noising observed gravity data	9
2.1.2	Verify applicability of ICA	9
3	Gravity anomaly	10
3.1	Introduction	10
3.2	Factors governing Gravity Anomaly and Necessary corrections	11
3.2.1	Eotvos effect due to carrier motion	12
3.2.2	The Latitude correction	12

3.2.3	The Elevation correction and the Bouguer's effect	13
3.2.4	Terrain correction	13
3.3	Noise Reduction and Data separation	14
3.3.1	Low Pass Filter (LPF)	14
3.3.2	Second Order Blind Identification (SOBI)	15
3.3.3	Independent Component Analysis (ICA)	16
4	Independent Component Analysis (ICA)	18
4.1	Blind Source Separation(BSS)	18
4.2	Statistical Independence of Signals	19
4.2.1	Definition	19
4.2.2	Uncorrelated signals and independent signals	19
4.3	Independent Component Analysis (ICA)	20
4.3.1	ICA Model	20
4.3.2	Assumptions in ICA	21
4.3.3	Ambiguities of ICA	22
4.4	Non-gaussianity	23
4.4.1	Non-gaussian is independent	23
4.4.2	Measures of non-gaussianity	23
4.5	Preprocessing for ICA	25
4.5.1	Centering	25
4.5.2	Whitening	25
4.5.3	Filtering	26
4.6	ThinICA Algorithm	26
5	Time frequency Analysis	28
5.1	Introduction	28
5.2	Discrete Wavelet Analysis	28
5.3	Continuous Wavelet Transform	29
5.4	Time frequency analysis of Gravity data	30
6	Gravity Data Observation	32
6.1	Survey Site and Data observation	32
6.2	Prototype of Gravimeter (EZ-GRAV)	34
6.2.1	VSE (Analog Servo)	34

6.2.2	Accelerometer Titan	35
6.3	Synchronization of Gravimeter with GPS	35
6.4	Raw Gravity Data	36
6.4.1	Analog Servo (VSE) Data	36
6.4.2	Accelerometer Titan Data	37
6.4.3	Digital Servo Data	37
6.5	Reference Data from Gravity map	38
7	Application of Noise Reduction and Separation Methods	40
7.1	Introduction	40
7.2	FIR Low Pass Filter (LPF)	40
7.3	Discrete Hardy Wavelet Analysis(DWT) and ICA	41
7.4	Continuous Wavelet Transform (CWT) and ICA	43
7.5	LPF and SOBI	45
7.6	LPF and ICA	46
7.7	Comparison of Results	48
7.8	Conclusion and Implications	50
8	Conclusions	57
	Appendices	61
A	Discrete Analytic Signal Wavelet Decomposition	61
A.1	Introduction	62
A.2	Analytic Wavelet Transform	63
A.2.1	Analytic Signal	63
A.2.2	Discrete Analytic Wavelet Transform	64
A.2.3	Phase and Amplitude localized in time-frequency domain	64
A.3	Uncertainty in Phase	65
A.3.1	Phase in Wavelet Transform	66
A.3.2	Numerical Simulations	67
A.4	Conservation of Transmitting function	68
A.4.1	Calculation of Response Using Wavelet Coefficients and Transmitting Function	69
A.4.2	Numerical Simulation	72

A.5 Conclusion (Appendix A)	72
B Computation Results (Contd.)	74
B.1 Low Pass Filter	74
B.2 Continuous Wavelet Transform	75
B.3 Low pass filter and SOBI	75
B.4 Low pass filter and ICA	75
B.5 Comparison of Results	75

List of Figures

3.1	Density layers, Density contrasts and Gravity anomaly	11
3.2	Schematic diagram of Eotvos effect showing a variation of centrifugal acceleration (CA) as a function of latitude.	13
3.3	Topographic section illustrating Bouguer and terrain.	14
4.1	Illustration of ICA.	21
5.1	Time-frequency representation of VSE (2 sec interval sampled 4096 data) using CWT.	31
6.1	Gravity Survey Site at Toyama bay (Google map) with survey route and time. Date of survey: 2011/10/31	33
6.2	Prototype gravimeter setup (EZ-GRAV: Analog Servo, Accelerometer Titan).	36
6.3	Observed time series gravity data (Gals) by Analog Servo (VSE) starting at time 12:27:00.	37
6.4	Observed time series gravity data (gals) by Accelerometer Titan (Taurus: vertical component), starting at time 12:53:06.	38
6.5	Reference data (Eotvos effect + Free air anomaly) obtained by using Gravity map produced by AIST, Japan, starting time: 12:17:00.	39
7.1	a) Top: LPF applied to VSE data (cut-off 100 sec) and b) Bottom: Analog servo (VSE) data.	41
7.2	a) Top: LPF applied to Taurus-Z data (cut-off 100 sec) and Bottom: Accelerometer Titan (Taurus-Z component.	42
7.3	Reconstructed signals after applying Hardy wavelet low pass filter (Scale < 6 or around 60 sec cut-off period); a) Top: VSE data and Bottom: Taurus Z-component. Time starts from 13:30:00 to 15:22:00.	43

7.4	Output by ThinICA: Same as input (No separation); a) Top: Corresponds to Taurus-Z data and Bottom: Corresponds to negative multiple of input VSE.	44
7.5	Input data to SOBI: VSE and Taurus-Z data after applying FIR LPF (cutoff = 100 sec, Time: 13:30:00 to 15:22:00.	46
7.6	Output by SOBI for the input in previous figure 7.5.	47
7.7	Input sets of data (VSE and Taurus-NS, EW, Z) for ICA.	48
7.8	Output by NG-FICA algorithm: No difference from input.	49
7.9	Output by ThinICA: refer figure 7.5 for input (same input given to SOBI).	50
7.10	Comparison of two de-mixed (separated) data (using demixing matrix for CWT coefficients input to ThinICA) with LPF VSE data and Reference data: Demixed2 data shows some improvement as the reduction in jump at around 6000 seconds (or time 14:50:00) is observed.	51
7.11	Comparison of outputs by SOBI (Red), ThinICA (Blue) with Reference data (Green): SOBI is better at ship-stoppage time (3120 to 4800 seconds or 14:22:00 to 14:50:30) and ThinICA is better after ship motion starts at 4800 seconds (i.e., at time 14:50:30.	52
7.12	Comparison of outputs by ThinICA (Blue), SOBI (Green), VSE LPF(light green) with Reference data (Red). VSE LPF data is included to above figure 7.11: ICA separated data is far improved than LPF data.	52
7.13	Truncated Input LPF VSE and LPF Taurus-Z data: Ship stoppage time (14:22:00 to 14:50:30) removed.	53
7.14	Output by ThinICA for truncated input with Ship stoppage time (14:22:00 to 14:50:30) removed.	54
7.15	Output by SOBI for truncated input with Ship stoppage time (14:22:00 to 14:50:30) removed.	55
7.16	Comparison of output by ThinICA and SOBI with VSE LPF and Reference data (mGals), for truncated input with ship stoppage time removed: Both SOBI and ThinICA show improvements than LPF data with similar performance.	55
7.17	Reference data in frequency domain: Frequencies less than 0.01 Hz are observed to be dominant.	56
A.1	Time history of an analytic Hardy wavelet signal.	65

A.2	Fourier transform of analytic Hardy wavelet signal.	66
A.3	Discrete Analytic Hardy wavelet coefficients representing time-frequency characteristics of a time series signal.	67
A.4	Comparison of time series of an original and reconstructed signal after the Fourier phase of 2 percent of the total or 25 percent of the dominant Fourier coefficients are contaminated by noise. Reconstructed wave is accompanied by the ripple which did not exist in the original signal.	69
A.5	Comparison of time series of the original and reconstructed signal after the phase of 2 percent of total or 20 percent of the dominant analytic Hardy wavelet coefficients are contaminated by noise. Reconstructed signal is more affected locally compared to the reconstructed signal in Figure A.4.	70
A.6	Comparison of time histories of responses of a linear system calculated by the presented wavelet scheme and the time-integration scheme. They show a very good agreement.	73
B.1	a) Top: FIR LPF Taurus EW data with cut-off period 200 sec and b) Bottom: Accelerometer Titan (Taurus EW-component).	74
B.2	a) Top: FIR LPF Taurus NS data with cut-off period 200 sec and b) Bottom: Accelerometer Titan (Taurus NS-component).	75
B.3	a) Top: FIR LPF Taurus-Z with cut-off period 200 sec b)Bottom: Accelerometer Titan (Taurus-Z component.	76
B.4	a) Lower: Digital Servo data and b) Upper: FIR LPF digital servo data with cut-off period 100 sec.	76
B.5	Input CWT coefficients (upper left), close view of input (upper right), Output by ICA (lower left) and the de-mixing matrix (lower right).	77
B.6	Input for SOBI after filtering 200sec lpgf.	78
B.7	SOBI outputs after filtering 200sec lpgf.	78
B.8	Input for ICA after applying 200sec lpgf.	79
B.9	Output by ThinICA (for four inputs (200 sec lpf): Neither component is very good.	79
B.10	Output by Power ICA(for four inputs (200 sec lpf): All outputs are same: comparison shown later: Not very good.	80
B.11	Input to ThinICA vse and tauZ data (lpf 200sec): Time 13:00 till end.	80

B.12 Output by ThinICA for vse and tauZ data (lpf 200sec): 13:00 till end: Performance is not good.	81
B.13 Input to ThinICA(VSE and Taurus Z) LPF(200 sec). Time: 13:30:00 to 15:30:00 truncated.	81
B.14 Output by ThinICA for 13:30:00 to15:30:00 truncated lpf(200 sec) data (input vse and tauZ: One output (top) is good (same output for 100 sec cutoff is better than this).	82
B.15 Input for ICA (200sec lpf) and truncated four sets of data 13:30 to 15:30).	82
B.16 Output by ICA for input (200sec lpf) and truncated four data 13:30 to 15:30: Third output is good.	83
B.17 Four Outputs by Power ICA (input: four 200sec lpf data) compared with reference data.	83
B.18 Four Outputs by SOBI (input 200sec lpf) compared with reference data 13:00:00: Fourth output follows some trend of reference but not very good.	84
B.19 Four Outputs by Thin ICA (input 200sec lpgf) compared with reference data 1300 onwards: Fourth output is good.	84

List of Tables

- 6.1 Features of Prototype EZ-GRAV 34
- 6.2 Basic specification of accelerometer VSE 35

Acknowledgements

First of all I would like to express my sincere gratitude to my advisor Prof. Riki Honda for all his hard work for me. He has thoroughly helped me to refine this thesis. I am very much inspired from his art of scientific reporting and I am gratefully learning it from him. He is a great teacher and a great person to me. Whether it is for a simple latex or a linux command or it is a difficult concept in signal processing, he has always helped me with equal calmness and kindness. My mere words are surely not enough.

Secondly, I would like to thank my co-advisor and Project team leader Prof. Hitoshi Morikawa at Tokyo Tech. for including me in his project. My sincere thanks to him for sharing me his knowledge and relevant information about the project in a friendly manner. He was very kind on replying my all sorts of questions.

I want to pay my sincere thanks to my co-advisor Associate Prof. Tomonori Nagayama for his valuable questions, comments and suggestions. He helped me to deepen my insight in my subject. I would also like to acknowledge all the project team members; Prof. Masumoto (Toyama University) for including me in the field survey, Prof. Chiba (Toyama) for allowing us to perform survey in his ship. Many many thanks to Prof. Saeki (Tokyo University of Science) for sharing his knowledge to me about the subject during the field trip. Heartily thanks to Mr. Satoshi Tokue and Ms. Ogura at Tokyo Tech. for retrieving and preparing data for further processing and also their kind company in the field trip.

I can not remain without thanking my former lab-mate Dr. Takashi Miyamoto. He has kindly given a lot of his time to clean my rust in mathematics and computations in my initial days of research life. He also gave his valuable time to check my programs in the works presented in Appendix A.

My cheerful and helpful IP lab members helped to make my research life more interesting and joyful. My special thanks to Dr. Shunshaku Komatsuzaki for all his enthusiastic support in IP lab. I would also thank Team Honda members Tauqir Ahmed, Hoang Chi Thanh, Yusuke Kazama et al. for their all kinds of support and company.

I think I have not only learned my subject but also the etiquette and sense of responsibility a person should possess from the Japanese people and society here. Thanks to all of them. All my international and national friends in Todai made my Tokyo life cheerful. I am proud that I have earned very good friends here. I feel rich to be in the network of such a highly civilized people.

My family members have been very kind, patient and loving to me. I could learn and work here with their constant inspiration and support from back home.

I would like to acknowledge Asian Development of Bank (ADB-JSP, Japan) for providing me scholarship to pursue my two year master's course in the University of Tokyo. I hope I could pay the returns to Asia one day.

Lastly, Thanks to God for everything.

Abstract

The ground motion (GM) characteristics are affected by local subsurface structure. Gravity method is one of the useful methods to know the information on subsurface structure. The gravity anomaly data obtained by gravity survey can be correlated with the lateral variation of subsurface rock densities. For gravity survey, spring type gravimeter has been used so far. This gravimeter gives accurate resolution but they are very expensive and difficult to handle. Recently, Team Morikawa have developed a prototype mobile gravimeter that uses Force-Balanced (FB) accelerometer. This prototype is light weight, compact, easy to handle and inexpensive. It also offers the resolution that is good enough for preparing gravity map for subsurface modelling. However, unlike the conventional spring-type gravimeter, this newly developed FB gravimeter is highly sensitive to high frequency noise. The observed data by this gravimeter are easily contaminated by various kinds of disturbances in a small size carrier like engine vibration, carrier acceleration, wind velocity and carrier tilting accompanied by sensor drifts, electrical noise etc. The amplitudes of such noises can be upto 100,000 times larger than the gravity anomaly. In order to extract the gravity anomaly from such observation, data processing is essential.

Conventionally, the data was observed in a large carrier (ship) on a more stable environment and the sensor was not sensitive to high frequency noise, so the noise contamination was not severe. The data processing techniques like low pass filtering and Second order statistics method (such as SOBI) were used. However, in case of severely contaminated data, low pass filtering might not be enough. SOBI is an advanced blind source separation (BSS) technique that separates source and noise blindly by exploiting the statistical property of data. It separates the target source by assuming that source and unwanted data are un-correlated at various time-lags. The gravity anomaly and other noises are generated from independent physical sources. It can be safely assumed that gravity anomaly and other data are independent but, it can not be strictly claimed that they have no correlation. So, further improvement than second order statistics method is desired. As a scheme of considering independence of signals to blind source separation, Independent Component Analysis (ICA) has been used in the field of BSS since 1990's. It separates the sources by maximizing the independence of linearly transformed observed

signals. Both mixing matrix and source signals are identified when only the mixed data are available. Further, independence between signals has nothing to do with their amplitudes. The huge difference in amplitudes among gravity anomaly and noise does not affect their independence. So ICA is suitable for our purpose. ICA renders ambiguity in amplitude of separated signal but this problem has little significance in our case since an appropriate scalar multiple can be estimated with the help of information of gravity at few known points. Thus it is proposed to use ICA for separating gravity anomaly data from its mixture with several noises.

The survey data is observed at Toyama bay, Japan. The National Institute of Advanced Industrial Science and Technology (AIST), Japan has provided the gravity map for the same place. This map is used to calculate the reference data that facilitates us to verify the performance of the proposed scheme. The prototype gravimeter consisted of group of sensors. Since ICA requires at least two sets of data, the major data obtained by Analog servo (VSE) was combined with data by other sensors as supplementary data. Following Team Morikawa's approach, the performance of various sensors are compared.

The application of low pass filtering(LPF) as a pre-processing to ICA is realized to be important. The presence of high frequency noise in the data is found to be unfavourable for the separation of gravity anomaly data. Both SOBI and ICA work only after the application of LPF. The choice of an appropriate cut-off filter was also observed to affect the results. The combination of VSE data and vertical component of Accelerometer Titan (Taurus-Z) as an input to ICA gives good result. When other horizontal components were used with VSE data the results are not satisfactory. Further, ICA is found to perform better at certain conditions of data acquisition environment. At the portions when ship motion is unidirectional the trend of ICA separated data is harmonious with reference data. When the ship velocity was lesser while proceeding towards the sea, the ICA result is matching very well with reference data. When the ship was highly unstable during ship stopping time ICA result are deviating away from the reference data. At other relatively stable sections the ICA separated data follows the trend of reference data well.

The separation of input data by ICA into different output components verifies that the source gravity anomaly and other data are independent. Thus it satisfies our assumption. The harmony of ICA separated data with trend of reference data at major sections verifies the applicability of ICA, under certain data acquisition environments. The accuracy of properly separated data by ICA is good enough for preparing gravity map for the purpose

of subsurface modelling. However, there is still a room for further improvement. An effort is made to study time-frequency characteristics of data without observing any clear merit so far. The further improvement in methodology is considered to be the part of future works.

Based on the results and considering the applicability of ICA so far, it can be concluded that a positive sign is observed for the improvement of mobility of gravity method.

Chapter 1

Introduction and Background

On February 2011, a powerful M 6.3 earthquake hit Christchurch, New Zealand that severely damaged the structures rendering 185 people killed. This incident was called as one of the deadliest in the history of New Zealand. This disaster signifies the essence of careful seismic design. It has been observed in several cases that the severely damaged area are concentrated at a certain place while the other nearby places experience lesser damages. The concentration of damage is due to the ground motion characteristics and that is affected by the local subsurface structure. The reliability in seismic design of structures can be improved by improving the accuracy in simulation of ground motion characteristics and that can be done by improving the modelling of subsurface structure. The strong ground motions simulation by 3D finite-difference methods using the subsurface model can be referred to [13].

This chapter discusses on the various subsurface exploration methods that can be combined to estimate the complex subsurface model. The combination of number of methods is worthwhile for estimating a more realistic model but it is also associated with the surge in total cost and time. So, improving the mobility and reducing the cost of gravity method can be very useful. However, the mobile gravimetry is associated with a huge challenge in data processing. The contamination of observed gravity data is highlighted followed by a brief introduction to conventional noise reduction and signal separation techniques along with the proposed signal separation method.

1.1 Sub-surface investigation

The modelling of sub-surface structure can be done by combining several sub-surface survey methods such as: a) Boring, b) Seismic method, c) Microtremors method, d) Magnetic method, e) Gravity method etc. The following paragraphs briefly explain these methods and the details of Gravity method are presented in chapter 3.

1.1.1 Boring

Boring deals with excavation or drilling the earth. The accurate information about the depth of various rock strata is obtained and the soil samples can be taken to the lab to estimate their densities. However, the operation cost increases with depth of drilling and it certainly consumes immense time to perform survey at numerous points and obtain data to cover a wide area for mapping gravity data.

1.1.2 Seismic method

Seismic method [11] involves the use of a source of seismic energy, propagation of elastic waves through the sub-surface strata and reception of the reflected waves by an array of seismic detectors. The acoustic waves generated on the earth's surface by the explosion are reflected back from noticeable strata boundaries and reach the surface of the earth after varying intervals of time, depending on the distance and the nature of the subsurface traversed. The detectors transduce the received waves into representative electrical signals from which data can be deduced concerning the geological subsurface conditions.

This method can cover larger area compared to Boring, but it is still very expensive and time consuming. Roughly, the relative cost per unit area of Magnetic, Gravity and Seismic method including both field works and data processing stand around the ratio of 1 to 10 to 100 [16].

1.1.3 Microtremors method

The surface of the earth is always in motion at seismic frequencies, even without earthquakes. These constant vibrations of the Earth's surface are called microseisms or microtremors. These ground vibrations have displacement amplitude about 0.1-1 micrometer (or 0.00001-0.0001 cm), and velocity amplitude are 0.001-0.01 cm/s and that can be detected by seismographs with high magnification [20]. These seismographs are located and

arranged in arrays to cover an area of interest for recording the subtle vibrations. Records of microtremors clearly show that microtremors are highly variable, irregular, vibratory phenomena, both temporally and spatially. However, in elasticity theory, microtremors are assemblages of body waves and surface waves [22]. This vibratory phenomenon comprises a set of stationary and stable spectra with very little variation, within the temporal and spatial extent of one hour and 1-2 km radius, respectively.

The basic principle of the micro-tremor survey method is described as detecting the form of dispersion of the surface wave contained in the microtremors, i.e., determining the relationship between phase velocity and frequency (or period) [14]. Here, the band of frequencies relates to the range of the depth of investigation. The longer the period used, the deeper the depth of investigation becomes.

However, this method is also very expensive since numerous number of seismograms are required at a single time for the microtremor recordings. Further, the subsurface model developed by this method can be less accurate than that obtained by seismic method.

1.1.4 Magnetic method

Magnetic method [21] involves the observation of variation in magnetic field along the earth's surface using a magnetometer. The largest component (80-90 percent) of the earth's magnetic field originates from liquid iron in the earth's outer core [30]. The magnetic fields external to the earth also have effect on magnetic measurements that must be removed during data processing. The remaining component of the earth's field originates in iron-bearing rocks in the earth's crust. The continental terrain boundaries are commonly recognized by the contrast in magnetic fabric across the line of contact [12]. With the advent of HRAM surveys and the sub nano Tesla resolution they offer, it is now possible to map intra-sedimentary faults by identifying their small and complex magnetic anomalies that occur where there are "marker beds" containing greater than average quantities of magnetite. Displacement of these marker beds generates subtle (a few tenths to about 10 nano Tesla at 150 m elevation) anomalies that can be used to trace corresponding fault systems.

1.1.5 Gravity method

The gravitational field varies subtly from place to place along the earth's surface. The acceleration due to gravity 'g' changes by a small amount along the surface and the phe-

nomenon is called as gravity anomaly. The gravity anomaly and subsurface rock densities have positive correlation with each other. Owing to the subtlety of gravity anomaly data, the instrument required for the data observation, or gravimeter, should be very sensitive. For the purpose of subsurface structure modelling for earthquake ground motion simulation, the required resolution of gravity map is 1 milligal. To obtain this resolution in the gravity map, the sensor should offer the resolution in the order of 0.01 milligal.

The gravity anomaly data thus observed can be utilized to model the two layered subsurface structure. The low period (or high frequency) components of the data correspond to the properties of shallow depth (1-2 kms) and the high period (or low frequency) components of the data correspond to the properties of basement. Since the model obtained by this method is very simple, the combination of this method with the other methods can prove to be very useful for the estimation of more accurate and complex subsurface structure.

The further details of gravity method are discussed in chapter 3.

1.2 Combination of subsurface survey methods

All of the above sub-surface investigation methods have their own advantages and limitations. By utilizing the information obtained by multiple methods modelling sub-surface structure is done. The subsurface structure model thus obtained is more complex and realistic than compared to simple models obtained by using individual methods. As an example, magnetic method is useful for locating faults, microtremors method is useful for determining shear wave characteristics and gravity method is useful for detecting density contrasts and defining boundaries. Utilizing these different merits of different methods, the limitations of the individual methods can be minimized and more realistic model of underground rock structure can be determined. As a reference, the subsurface structure modelling of Izumo plain at southwest Japan is done by combining microtremors and gravity anomalies [17]. Similarly, the magnetic and gravity methods are combined Niigata, Japan for the same purpose [7].

However, the combination of several methods will considerably increase the cost. If the mobility of gravity method can be improved and economic gravimeters can be developed that can be very useful as the continuous data observation can be made economically.

1.3 Mobility of Gravity method

Prof. Morikawa (Tokyo Institute of Technology, Japan) and his team has recently developed a mobile gravimeter prototype. This gravimeter uses force balanced (FB) type accelerometer as its sensors. This is compact, light weight and inexpensive. Since multiple sets of data are required for data processing, the prototype is constituted of various sensors. These sensors have different dynamic observable ranges and different resolution or sensitivity to time. Some accuracy is compromised in comparison to the spring type gravimeters but it offers the accuracy enough for the purpose of preparing gravity map with 1 mGal contour intervals for subsurface structure modelling. The specifications and details of prototype gravimeter are presented in chapter 6. In contrast to the sensors used in spring type gravimeters, the FB accelerometer sensors are highly sensitive to the high frequency noise. As a result, the observed data is associated with the severe contamination with all kinds of disturbances.

1.4 Severe Noise contamination in Gravity data

The mobile air-borne and ship-borne gravity survey involves the use of small sized carriers that are associated with challenges like instability of carrier and disturbed motion. The tilting of carrier (by sea waves in case of sea area) leads to the affected recording of gravity data. The vertical component of the horizontal acceleration of the carrier adds to the gravity data to give distorted observed data. The horizontal or vertical accelerations due to the movement of ship can be a tenth or more of the total gravity. For the purpose of sub-surface structure estimation for earthquake ground motion evaluation, we must detect the changes of the order of 0.1 milligal or less within a background noise of the order of one hundred thousand milligals (1/10 of normal gravity or 1000 gals) [16]. Such a high order of noise and the subtlety of required data is the biggest challenge in air-borne or ship-borne gravity survey. The other factors responsible for occurrence of such severe noise in the observed data are engine vibration, carrier acceleration, wind-turbulence, sensor drift, temperature, electrical noise etc. The frequency of time-series gravity anomaly data is the function of the horizontal extent of the gravity disturbance and the speed of the carrier which controls the time required to pass over those disturbances [16]. The frequencies of noises are basically the function of motion of carrier and are anticipated to be higher than the frequency of gravity anomaly data. The problem is basically filtering out the high-

frequency noise from the anticipated low-frequency gravity anomaly data. However, the drift of the sensor data due to temperature changes can be of relatively lower frequency. So filtering out certain range of frequencies is not sufficient since the noises can be within the data with variety of frequency ranges.

1.5 Noise Reduction and Data separation

The observed gravity anomaly data is in the mixed form of true gravity anomaly data and the unwanted acceleration components of carrier motion (called as noise here). Since the useful data is subtle, the task of separating it from the mixture is sensitive. The noise reduction and separation techniques should be carefully applied in order to obtain the true data. Currently, several methods such as simple low pass filter and second order statistics methods (such as SOBI) are used. But the noise contamination is severe, so further improvement in the noise reduction methodology is expected. Since the gravity anomaly and noises are originated from the independent physical processes, it is expected that the source data and noises are statistically independent. The advanced blind source separation techniques utilize this simple condition of statistical independence among source and noise and Independent Component Analysis (ICA) is gaining popularity for considering independence.

1.5.1 Low Pass Filter

A signal or data is composed of several spectral components. Considering the fact that the higher frequency components correspond to noises, they can be filtered out keeping only the lower frequency components by applying a low pass filter. Filtering is a convenient technique that has been used in the field of signal processing for a long time. However, there is always a risk of losing the useful information while filtering out a certain range of spectral components. Moreover, if there are unwanted information in lower spectral range, the true data is still latent mixed with the unwanted data. The mobile gravimetry is associated with noises at various ranges of frequencies, filtering alone might not be enough.

1.5.2 Second Order Blind Identification (SOBI)

SOBI is an advanced blind signal separation technique [1] that exploits the fact that the true source and noises are not correlated at all the time-lags. Correlation is measured by second order moments so they consider the linear dependency between the variables under consideration. However, un-correlation is the weaker form of independence i.e., Un-correlation does not always imply independence although independence always implies un-correlation. Although SOBI is a good method, it might not offer the representation of independence between the source data and noises at all circumstances.

1.5.3 Independent Component Analysis (ICA)

ICA is another advanced blind signal separation technique [4] that considers the independence among the source data and noises. The observed data is assumed to be in a linearly mixed form of the source data and noises. Since the mixing parameters and source data both are determined with the help of only observed data, this method is known as a blind technique. Independence is measured by the fourth order cumulants, so they consider the non-linear dependence between the variables and thus can consider the independence more strongly than by SOBI.

We propose to use ICA for the acquisition of gravity anomaly data. ICA is explained in detail in chapter 4.

1.6 Organization of thesis

This thesis consists of eight chapters. Chapter 1 presents the introduction and background for the research, starting from various subsurface exploration methods, highlighting the significance of mobility of gravity method followed by challenges in mobile gravimetry and noise reduction techniques to cope with the challenge. Chapter 2 discusses on the objectives of research. The detail description of gravity method, gravimeters and gravity data processing or de-noising techniques are included in Chapter 3. Chapter 4 illustrates ICA in detail and Chapter 5 introduces the concept of time-frequency analysis using continuous wavelet transforms for further analysing the data. Chapter 6 is devoted to data observation. The application of noise reduction techniques, results and their comparison with the reference data are shown in Chapter 7. Finally chapter 8 concludes the thesis relating the objectives and results.

Chapter 2

Objectives

The numerous efforts have been made in order to improve the reliability of seismic design, but it is yet not enough. The key in reliable seismic design is the realistic ground motion simulation and that requires the realistic modelling of subsurface structure. Given the huge scale of earth and impassable solid subsurface, to follow non-destructive or passive subsurface survey methods is the best ways to deal with and Gravity method is one of the useful methods of such type. Further, the combination of more information from multiple methods contributes to improve accuracy of subsurface model. But the associated increased cost can not be underestimated. Thus, for the economic combination of multiple subsurface survey methods, the development of inexpensive mobile gravimeter would be helpful. Recently, Prof. Morikawa et al. have developed a new prototype mobile gravimeter. Mobile gravity survey involves the use of small carriers that are severely unstable during the motion. The various factors like engine vibration, tilting of carrier, sensor drifts etc. contribute to the inevitable distortion in the gravity data. The only option to cope up with this challenge is to apply an appropriate data processing methodology. Because of the immense severity in contamination of data, the conventional methods like filtering and second order statistics methods might not be enough. So the objective in this thesis is to propose an appropriate methodology for processing these data to extract gravity anomaly data.

2.1 Objectives

2.1.1 Propose ICA for de-noising observed gravity data

The challenge of the problem lies in the fact that the distorted (noisy) observed data is the only thing we possess and we are ignorant on the mixing parameters and the source components. However, because the source components generate from totally different physical processes, we can assume that they are statistically independent to each other. To assume independence among the data is a weak assumption. With the help of observed data and such a weak assumption, an effort is made to extract the true gravity anomaly data.

Independent Component Analysis (ICA) has been gaining popularity as a method to consider independence among source signals. It is proposed to use ICA for the purpose of separating gravity data from the observed distorted data. Independence is attained by maximizing non-gaussianity of arbitrarily transformed observed data (refer chapter 4 for details). Non-gaussianity is measured by the higher order cumulants such as fourth order in kurtosis. Since the higher order cumulants represent non-linear dependency between the random variables, they can consider independence more strictly compared to second order statistics that uses second order cumulants. The second order statistics method can assure only linear independence while the proposed method ICA can assure non-linear independence between the random variables.

2.1.2 Verify applicability of ICA

The second objective is to verify the applicability of ICA by utilizing it for the real world observed data. The separation of mixed observed data into the independent components will verify the assumption that source gravity data and noises are statistically independent. Once the independence between gravity data and noises are verified. The applicability of ICA for the purpose of gravity data processing is verified by comparing its performance with the available accurate reference data obtained from gravity map produced by AIST, Japan.

Chapter 3

Gravity anomaly

3.1 Introduction

This chapter is focused on Gravity anomaly and data processing. The contents related to Gravity anomaly are mostly referred to [16].

According to the Newton's law of gravitation, the gravitational force between two objects is directly proportional to the product of their masses and inversely proportional to the square of distance between them. So the gravitational field of earth changes depending upon the change in amount of mass and distance from the point of consideration on or above earth's surface to the center of earth. These quantities vary as a result of the variation in latitude, altitude, densities of the subsurface rocks etc. Since the variations in mass and distance due to these factors are subtle relative to the huge mass and radius of earth, the magnitude of gravity anomaly is also subtle. The acceleration due to gravity 'g' varies with the lateral variation in density of rocks. Provided the earth crust materials were in layers of horizontally uniform density, the gravity anomaly would be zero irrespective of the radial or vertical variation in density. However, in reality, the layers of uniform densities might be lying in an abruptly complex manner. This concept is represented in the schematic diagram (3.1) [16] If the layers 1,2,3,4 with increasing density values lay flat uniform laterally, there will be no gravity anomaly and the 'g' value would be constant. On the other hand, if the density contrast occurs laterally as a result of structural uplift as seen in the central portion of the figure, gravity anomaly is observed. Since the density contrast is positive, positive gravity anomaly is observed in the gravity profile. Thus, the density of the various components of the geologic column and the resulting density

contrasts among the rocks produced by structures developed in these rocks are related to gravity anomalies. And, so the structure of sub-surface rocks is estimated by utilizing this relation. However, the observation of gravity anomaly data alone is not enough and minimum lithological information on the rock strata or rock densities at certain regions are necessary.

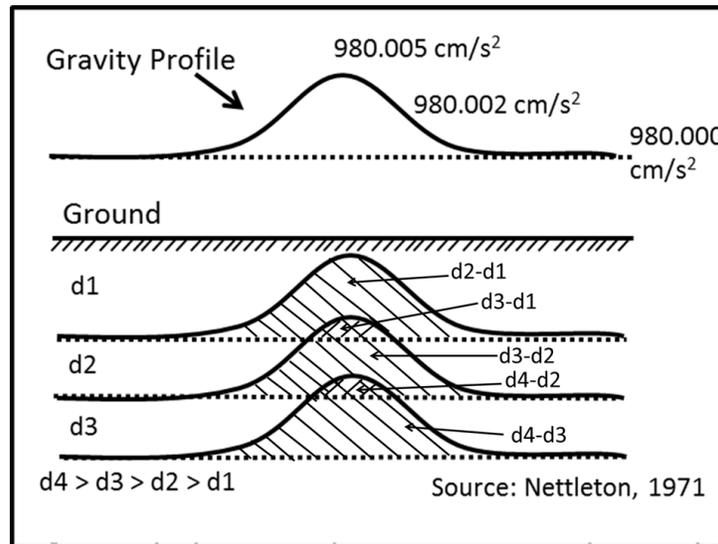


Figure 3.1: Density layers, Density contrasts and Gravity anomaly

Owing to the subtlety of gravity anomaly, the instrument required for the data observation, or gravimeter, should be very sensitive. For the purpose of subsurface structure modelling for earthquake ground motion simulation, the required resolution of gravity map is 1 milligal. To obtain this resolution in the gravity map, the sensor should offer the resolution in the order of 0.01 milligal. The raw data observed by a gravimeter consists of the gravity anomaly affected due to several factors along with the added unwanted noises. The factors affecting gravity anomaly are discussed in following subsection.

3.2 Factors governing Gravity Anomaly and Necessary corrections

There are several factors that are responsible for the occurrence of Gravity anomaly. In order to map the gravity anomaly due to variation in rock density in a spatial plan, the data should be corrected for the factors that are significant and are well known. The

effects due to all the factors except subsurface densities can be corrected and the resultant gravity anomaly data can be correlated to the variation in densities of subsurface rock strata [16]. The gravity anomaly and subsurface rock densities have positive correlation with each other. The several factors and necessary corrections are described in following subsections. The most contents of this section are referred to [16].

3.2.1 Eotvos effect due to carrier motion

Eotvos effect is known as the change in perceived gravitational force that is caused by the change in gravitational acceleration resulting from eastbound or westbound velocity of an object (ref: wikipedia). Every object on the surface of earth is subjected to Coriolis acceleration that occurs due to the rotation of earth. The vertical component of this acceleration acts outward the earth and it can be visualized as the centrifugal acceleration due to earth's rotation. This outward acceleration is a maximum at equator and decreases to zero at the poles. The earth is continuously rotating towards the east. If a vehicle moves towards east along the equatorial line its speed adds up with the earth's rotating speed and the centrifugal acceleration increases and the decrease in gravitational force is perceived. This is negative Eotvos effect and the correction should be positive. On the other hand, if the vehicle moves westbound the net centrifugal acceleration decreases and the increment in gravitational force are perceived. This is positive Eotvos effect and the correction is negative. The Eotvos effect can be calculated for varying velocities V at all angle of latitudes (φ) and course azimuth or heading α as: $E = 7.49V \cos(\varphi) \sin(\alpha)$. Also refer figure 3.2.

In this paper, all the data are shown including the Eotvos effect, and the necessary corrections are supposed to be done after the successful extraction of gravity anomaly data from the observed data.

3.2.2 The Latitude correction

This correction is necessary if the carrier for gravity survey moves from one angle of latitude to other. This correction removes the effect of the increase of gravity from the equator to the poles that occurs because of the combined effect of earth's rotation and the fact that the equatorial radius is 13 miles greater than the polar radius. The gradient of this effect is $1.307 \sin 2\varphi$ milligal/mile, where φ is the latitude. However, in this particular research work, the angle of latitude travelled is very small so the effect is negligible.

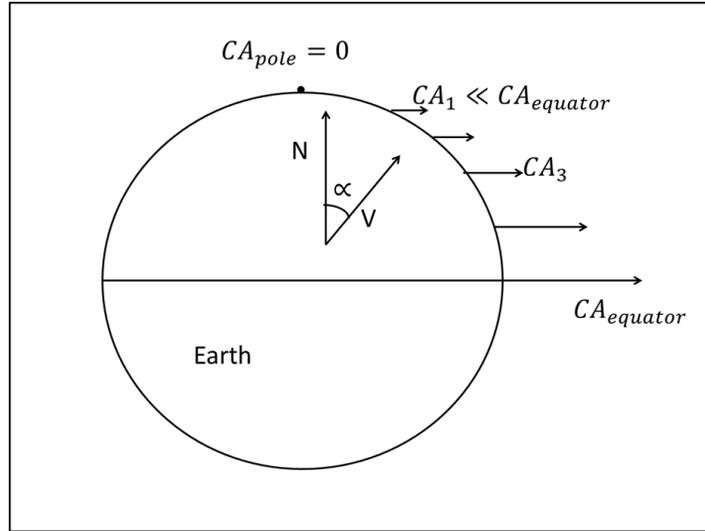


Figure 3.2: Schematic diagram of Eotvos effect showing a variation of centrifugal acceleration (CA) as a function of latitude.

3.2.3 The Elevation correction and the Bouguer's effect

The acceleration due to gravity decreases with the increasing distance from the center of earth. So the value decreases with the elevation from sea level in the free air. This is known as the free air effect and is 0.09406 milligal/ft. However, if the gravity measurement is done on the terrain at a certain elevation from the sea level, the attraction of the material between the measuring station and the sea level should be accounted that increases the gravity value. This latter effect is known as Bouguer's effect and is calculated as 0.01276ρ , ρ being the density of additional material between sea and the station. So the total elevation correction combines the free air effect and Bouguer's effect and is given by $E_c = (0.09406 - 0.01276\rho)h$ milligal/ft.

3.2.4 Terrain correction

Terrain correction is needed when there is substantial local topography around the station like in the mountainous regions. The Bouguer's correction just considers the station as it is on the flat land. When it is on the sloppy hill, Bouguer's correction alone is not enough because it considers the attraction due to the material that is not present at one side below the station point and at the same time fails to consider the attraction due to the material that is present at the other side above the station point as seen in point 3 in

figure 3.3 [16].

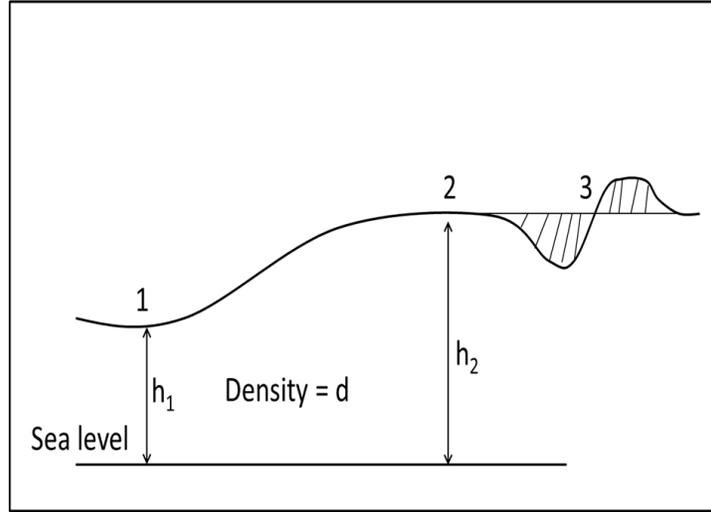


Figure 3.3: Topographic section illustrating Bouguer and terrain.

3.3 Noise Reduction and Data separation

The conventional low pass filter and SOBI are explained in this section.

3.3.1 Low Pass Filter (LPF)

The filtering of digitized data is the oldest discipline in the field of digital signal processing. It refers to the processing of time-domain signal by reducing its original spectral content, or simply filtering out the unwanted spectral components. Filtering allows certain frequency components to pass and attenuates other unwanted frequency components [24]. This is a useful method when the ranges of unwanted frequency components are obvious and known. However, in gravity survey (air-borne or ship-borne) there are several kinds of noise comprising of diverse range of frequencies. The exact ranges of frequencies are unknown for both gravity anomaly data and the other noises. We can make a rough assumption that gravity anomaly has lower frequency components and noises due to engine vibration and ship movement has higher frequency components, but there is possibility that some unwanted signals can be present in lower frequencies too, i.e. noise due to sensor drift etc.. Therefore, filtering alone may not be sufficient for this purpose.

Moreover, it is always associated with the risk of incurring a loss of useful information inherent in the removed frequency components.

Finite Impulse Response LPF (FIR LPF)

The low pass filtering involves smoothing the data by replacing every data in time series with a weighted average of that point and the points around it [4]. The number of points before and after it depend upon the choice of cut-off period and the time interval of data to be filtered. For example, if the time interval is 1 sec and cut-off period is 101 seconds, the number of data before and after the central data are 50 each.

When the original data before smoothing are considered (and not the already smoothed data) for smoothing every other data point, it is known as Finite Impulse Response (FIR) filter. If equal weights are given to all the data (linearly constant function) before and after the central data, it is an uniform filter. If the values of weights are distributed according to a gaussian function (central data given the maximum weights) it is known as a gaussian filter. FIR Low Pass Gaussian filter is used in this research.

The simple algorithm for the FIR LPF (Gaussian) is presented as follows. Let us consider a time series $y(t)$ is going to be applied FIR LPGF and y_f is the signal after filtering. Assuming a cut-off period t_c , time interval of data Δt and gaussian function f used for distribution of smoothing weights with standard deviation σ ,

$$\sigma = t_c/(2\Delta t) \quad (3.1)$$

$$f = \frac{1}{K} e^{-\frac{x^2}{2\sigma^2}} \quad (3.2)$$

$$y_f = \{y\} \times [f'_1, f'_2, \dots, f'_n] \quad (3.3)$$

where K is a constant that normalizes sum of weights into 1, range of x can be chosen according to the need (such as -3σ to $+3\sigma$), n is the number of time intervals of $y(t)$ and $[f'_1, \dots, f'_n]$ is $(n \times n)$ matrix.

3.3.2 Second Order Blind Identification (SOBI)

SOBI is an advanced blind source separation technique that exploits the time coherence of the source signals [1]. It assumes that the random variables (or true source and noises in the context of this thesis) are not correlated at all the time-lags with each other. This approach relies on stationary second-order statistics that are based on joint diagonalization of a set of covariance matrices that is based upon eigen-decompositon [1]. A stochastic

process $X(t), t \geq 0$ is said to be a stationary process if for all n, s, t_1, \dots, t_n , the random vectors $X(t_1), \dots, X(t_n)$ and $X(t_1 + s), \dots, X(t_n + s)$ have the same joint distribution. The process $X(t), t \geq 0$ is defined as a second-order stationary, or a covariance stationary if $E[X(t)] = c$ (constant) and $Cov(X(t), X(t + s))$ does not depend on time t [12].

The data model is $\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n}$, which is similar to ICA model explained in chapter 4 and that resembles to many situations of practical interest. Here, \mathbf{x} is the observed data vector, \mathbf{A} being the mixing matrix, \mathbf{s} being the source vector and \mathbf{n} being the noise vector. SOBI application is briefly explained as follows. Firstly, the whitening of observed data vector \mathbf{x} is done into whitened data vector \mathbf{z} . This means the observed data are converted into signals that are uncorrelated with each other and have unit variances each. Thus the co-variance matrix of whitened data vector \mathbf{z} is an identity matrix. Then the further sample co-variance matrices are computed for a certain set of time lags between the whitened data i.e., $z(t + \tau)$ and $z(t)$. for this whitened data vector \mathbf{z} . Then, these co-variance matrices are jointly diagonalized. This means all the sets of matrices are converted into identity matrices jointly. The diagonalizer matrix is combined with whitening matrix to determine the mixing matrix \mathbf{A} and finally the source data vector is estimated by inversion. The detail algorithms are referred to [1].

Correlation is measured by second order moments i.e., product of first order variables and so they consider the linear dependency between the variables under consideration. However, uncorrelation is the weaker form of independence i.e., Uncorrelation does not always imply independence although independence always implies uncorrelation. Although SOBI is a good method, it might not offer the representation of independence between the source data and noises at all circumstances.

3.3.3 Independent Component Analysis (ICA)

ICA is another advanced blind signal separation technique [4] that considers the independence among the source data and noises. The observed data is assumed to be in a linearly mixed form of the source data and noises. Since the mixing parameters and source data both are determined with the help of only observed data, this method is known as a blind technique. Independence is measured by the fourth order cumulants, so they consider the non-linear dependence between the variables and thus can consider the independence more strongly than by SOBI.

We propose to use ICA for the purpose of de-noising the gravity anomaly data. ICA is

explained in detail in chapter 4.

Chapter 4

Independent Component Analysis (ICA)

The content in this chapter are referred to the book [4].

4.1 Blind Source Separation(BSS)

Blind Source Separation is the separation of set of source signals from the available mixed set of signals without knowing anything (or very little information) about the source signals and the mixing parameters. Let us imagine the sets of observed data $x_1(t)$, $x_2(t)$ and $x_3(t)$ with the help of three sensors where t is the time index. The number of these sensors should be at least two for the purpose of BSS. Let us denote these source signals by $s_1(t)$, $s_2(t)$ and $s_3(t)$. Assuming the source signals are linearly mixed, the observed data can be expressed as:

$$\begin{aligned}x_1(t) &= a_{11}s_1(t) + a_{12}s_2(t) + a_{13}s_3(t) \\x_2(t) &= a_{21}s_1(t) + a_{22}s_2(t) + a_{23}s_3(t) \\x_3(t) &= a_{31}s_1(t) + a_{32}s_2(t) + a_{33}s_3(t)\end{aligned}\tag{4.1}$$

where a_{ij} with $i, j = 1, \dots, 3$ are some unknown parameters that depend on the medium between the source of data and the sensor, i being the number of sensors and j being the number of source data ($i = j$ is not true always). Here, only x_i are known and both the matrix a_{ij} and source s_j are unknowns. This is the BSS problem. Assuming that mixing matrix a_{ij} is invertible, there exists a de-mixing matrix w_{ij} such that the sources s_i are separated as

$$\begin{aligned}
s_1(t) &= w_{11}x_1(t) + w_{12}x_2(t) + w_{13}x_3(t) \\
s_2(t) &= w_{21}x_1(t) + w_{22}x_2(t) + w_{23}x_3(t) \\
s_3(t) &= w_{31}x_1(t) + w_{32}x_2(t) + w_{33}x_3(t)
\end{aligned} \tag{4.2}$$

For the separation, the mixing parameters w_{ij} are arbitrarily chosen. The mixed set of signals are linearly combined with these arbitrary mixing parameters. There are certain measures of statistical independence that are used to for these transformed mixed signals $w_{ij}x_j$. Finally, the measures of independence are maximized to make them mutually independent. Once their mutual independence is maximized, they are expected to be equal to the source signals s_i .

4.2 Statistical Independence of Signals

4.2.1 Definition

If y_1 and y_2 are two scalar-valued random variables, they are said to be independent if information on the value of y_1 does not give any information on the value of y_2 , and vice versa. Let us denote the joint probability density function (pdf) of y_1 and y_2 by $p(y_1, y_2)$. The marginal pdf $p(y_1)$ is given as:

$$p(y_1, y_2) = p_1(y_1)p_2(y_2) \tag{4.3}$$

If $f_1(y_1)$ and $f_2(y_2)$ are two functions of two independent random variables y_1 and y_2 respectively, we have

$$E\{f_1(y_1)f_2(y_2)\} = E\{f_1(y_1)\}E\{f_2(y_2)\} \tag{4.4}$$

4.2.2 Uncorrelated signals and independent signals

Uncorrelatedness is the weaker form of independence. Two random variables y_1 and y_2 are said to be uncorrelated if their covariance is zero:

$$E\{y_1y_2\} - E\{y_1\}E\{y_2\} = 0 \tag{4.5}$$

If the variables are independent $E\{y_1 y_2\} = E\{y_1\}E\{y_2\}$, so they are always uncorrelated. However, the converse is not always true as uncorrelatedness does not always imply independence.

4.3 Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a statistical and computational BSS technique for separating latent source data from its mixture with other signals. ICA is a relatively new invention. It was first introduced in 1980s in the context of neural network modeling [4]. Some highly successful new algorithms were introduced in mid-1990s. It has wide applications in the fields like biomedical signal processing, audio signal separation, telecommunication, financial time series analysis etc.

This thesis introduces the application of ICA for the purpose of separating Gravity anomaly data from its mixture with several noises. The features of ICA and its principle for the separation of sources are explained in this section.

4.3.1 ICA Model

To rigorously define ICA, a statistical model can be used. The n random variables $x_i(t)$ with $i = 1, \dots, n$ are observed which are modelled as linear combination of n random variables $s_i(t)$:

$$x_i = \sum a_{ij} s_j \tag{4.6}$$

In this model, the multivariate random vector x_i is observed by sensors, while both the matrix a_{ij} and multivariate random source vector s_j are unknowns. If the matrix a_{ij} was known, it would be very convenient to calculate the source vector s_j by simply inverting the linear system. Since there are two unknowns and only one known quantity in one linear equation, the problem does not have a unique solution. In order to solve this, a certain assumption regarding the statistical property of s_j is done. It is assumed that $s_1(t)$, $s_2(t)$ and $s_3(t)$ are, at each time instant t , statistically independent.

This is not unrealistic and is a weak assumption in the case of gravity anomaly data observation, because the sources of gravity and the noises are totally different and they have no relation with each other. Further, the gravity anomaly data is expected to have lower frequencies than the noises due to carrier movement or vibration or electrical noise

and this fact also supports the assumption. Independence Component Analysis (ICA) can be used to estimate a_{ij} based on the information of their independence. This allows to separate the original signals $s_1(t)$, $s_2(t)$ and $s_3(t)$, from their mixtures $x_1(t)$, $x_2(t)$ and $x_3(t)$. The original sources, mixed signals and the separated signals can be observed in the figure 4.1 as an illustration to ICA.

The basic ICA model can be represented as:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{4.7}$$

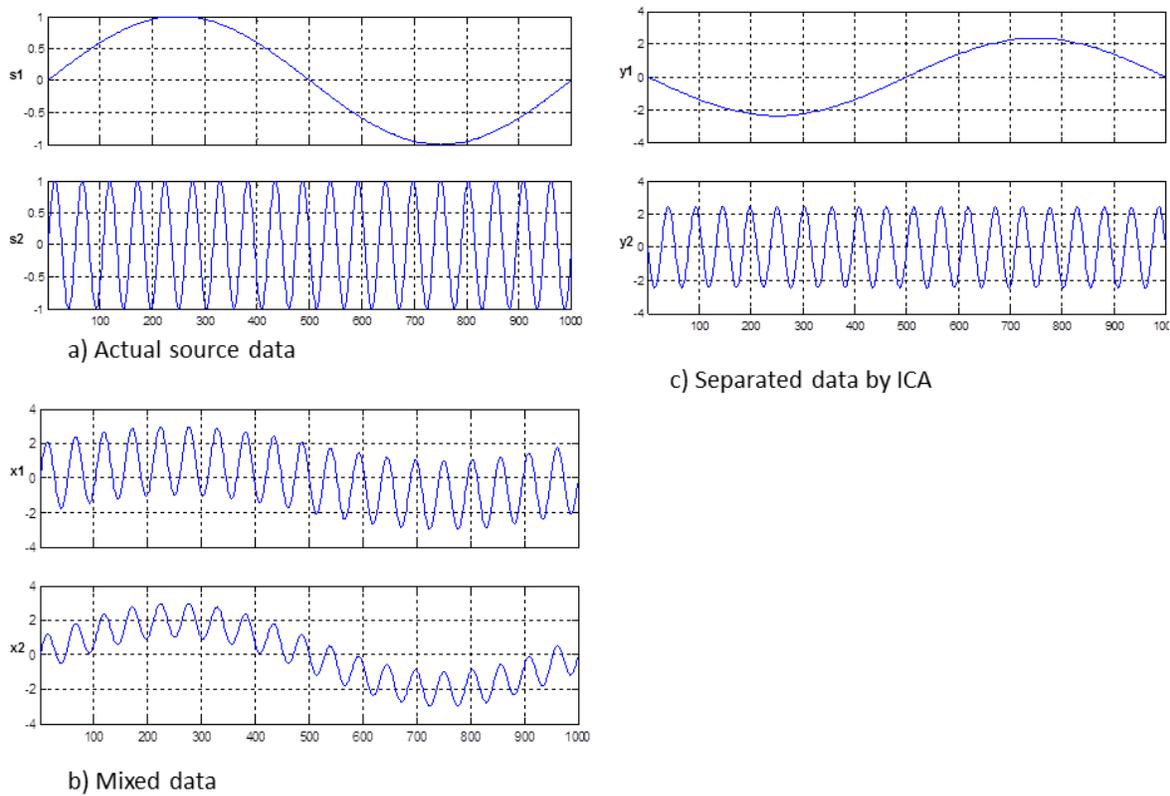


Figure 4.1: Illustration of ICA.

4.3.2 Assumptions in ICA

The independent components are assumed statistically independent. ICA is a powerful method since it rests fully on such a weak assumption. The random variables y_1, y_2, \dots, y_n

are said to be independent if information on the value of y_i does not give any information on the value of y_j for $i \neq j$. Independence can be defined by the probability densities. Let us consider $p(y_1, y_2, \dots, y_n)$ be the joint probability density function (pdf) of the variables y_i , and $p_i(y_i)$ the marginal pdf of y_i when considered alone. Then y_i are said to be independent if and only if the joint pdf is given as:

$$p(y_1, y_2, \dots, y_n) = p_1(y_1)p_2(y_2)\dots p_n(y_n) \quad (4.8)$$

The independent components must have non-gaussian distributions. The higher order information is essential for ICA estimation as will be explained in latter section. Since the higher order cumulants are zero for Gaussian distributions, ICA is impossible if the observed variables have gaussian distributions.

For simplicity, the unknown mixing matrix is assumed to be square. The square mixing matrix implies the equal number of sensors and number of independent sources. This simplifies the ICA estimation very much. By estimating the mixing matrix \mathbf{A} , its inverse matrix \mathbf{W} can be estimated and the independent components are easily computed as:

$$\mathbf{s} = \mathbf{A}^{-1}\mathbf{x} = \mathbf{W} \mathbf{x} \quad (4.9)$$

It is also assumed that mixing matrix is invertible.

4.3.3 Ambiguities of ICA

The variances (energies) of the independent components cannot be determined. Since both \mathbf{s} and \mathbf{A} of ICA model in equation (4.5) are unknown, the infinite number of solutions for \mathbf{s} and \mathbf{A} are possible that give the same product \mathbf{x} . Any scalar multiplier in one of the sources s_i could be cancelled by dividing the corresponding column a_i of \mathbf{A} by the same scalar α_i :

$$x = \sum_i \left(\frac{1}{\alpha_i} a_i\right) (s_i \alpha_i) \quad (4.10)$$

This problem is fixed by normalizing the independent components into unit variance, i.e., $E\{s_i^2\} = 1$. This still leaves the ambiguity of the sign, so it might often be necessary to multiply the independent components with -1 , which does not affect the model. The order of the independent components cannot be determined.

4.4 Non-gaussianity

4.4.1 Non-gaussian is independent

The central limit theorem states that under certain conditions, the distribution of a sum of independent random variables tends towards a Gaussian distribution. Thus the sum of two independent random variables usually has a distribution that is closer to gaussian than any of the two original random variables. Conversely, if such data close to gaussian are demixed by maximizing the non-gaussianity we tend to obtain the independent data. Let us assume that the data vector \mathbf{x} is a mixture of independent components as shown in equation (4.7). According to equation (4.9) the source independent components are again the linear mixture of $\{x_i\}$. Let us denote the estimated signal vector as $\mathbf{y} = \mathbf{W} \mathbf{x}$, where \mathbf{W} is a de-mixing matrix to be determined. If \mathbf{W} is equal to inverse of \mathbf{A} , \mathbf{y} would be equal to \mathbf{s} . Following the principle of central limit theorem, we consider \mathbf{W} a matrix of vectors that maximize the non-gaussianity of $\mathbf{W}\mathbf{x}$ in order to determine \mathbf{y} as close to \mathbf{s} .

4.4.2 Measures of non-gaussianity

In order to maximize the non-gaussianity we require a quantitative measure of non-gaussianity of a random variable, say y . For simplification, y is assumed to have zero mean and unit variance.

4.4.2.1 Kurtosis

The classical measure of non-gaussianity is kurtosis or the fourth-order cumulant. It is defined as

$$Kurt(y) = E\{y^4\} - 3(E\{y^2\})^2 \quad (4.11)$$

Since y has unit variance, the relation simplifies to

$$Kurt(y) = E\{y^4\} - 3 \quad (4.12)$$

The kurtosis is zero for a gaussian random variable. For most nongaussian random variables, kurtosis is nonzero. Random variables with negative kurtosis are called subgaussian, and those with positive kurtosis are called supergaussian. However, nongaussianity

is typically measured by the absolute value of kurtosis. In practice we start from some weight vector matrix \mathbf{W} , compute the direction in which the kurtosis of vector $\mathbf{y} = \mathbf{W}\mathbf{x}$ is growing most strongly based on the available sample vector \mathbf{x} , and use a gradient method for finding a new vector matrix \mathbf{W} . However, kurtosis also has some drawbacks. Its value may depend on only few observations in the tail of the distribution [23], which may turn erroneous.

4.4.2.2 Negentropy

Negentropy is based on the information-theoretic quantity of entropy. The entropy of a random variable can be interpreted as the degree of information that the observation of the variable gives. Higher entropy is characterized by higher degree of randomness or unpredictability of the variable. For a random vector \mathbf{y} with probability density $p(\mathbf{y})$, the differential entropy H is defined as:

$$H(\mathbf{y}) = - \int f(\mathbf{y}) \log f(\mathbf{y}) d\mathbf{y} \quad (4.13)$$

According to information theory, a gaussian variable has the largest entropy among all random variables of equal variance [29], [6]. This suggests that gaussian distribution has the highest degree of randomness of all distributions and naturally the entropy is smaller for distribution that are clearly concentrated on certain values, i.e., such as supergaussian distribution. Negentropy is the modified version of the definition of differential entropy represented by J as:

$$J(\mathbf{y}) = H(\mathbf{y}_{gauss}) - H(\mathbf{y}) \quad (4.14)$$

where \mathbf{y}_{gauss} is a gaussian random variable of the same covariance matrix as \mathbf{y} . Negentropy is zero if and only if \mathbf{y} has a gaussian distribution, otherwise, it is non-negative. Although Negentropy is a robust measure of nongaussianity it has difficulties in computations. For simplification, approximations of negentropy are made using higher order moments such as [18]:

$$J(y) \approx \frac{1}{12} E\{y^3\}^2 + \frac{1}{48} kurt(y)^2 \quad (4.15)$$

The random variable y is assumed to be of zero mean and unit variance.

4.5 Preprocessing for ICA

Before applying an ICA algorithm on the observed data, it is useful to do preprocessing. Centering, whitening and filtering are the basic preprocessing techniques.

4.5.1 Centering

As self-explained by the word, centering refers to centering the observed vector \mathbf{x} by subtracting it with its mean vector $\mathbf{m} = E\{\mathbf{x}\}$ order to make \mathbf{x} a vector with zero mean variables. After estimating the source vector \mathbf{s}_c with such centered data, the estimation can be completed by adding back the mean vector $\mathbf{A}^{-1}\mathbf{m}$ to \mathbf{s}_c to get \mathbf{s} .

4.5.2 Whitening

Whitening refers to the transformation of the observed vector \mathbf{x} linearly to make it white vector $\tilde{\mathbf{x}}$. A white vector has its components uncorrelated and their variances equal to unity. Naturally the covariance matrix of $\tilde{\mathbf{x}}$ is an identity matrix:

$$E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} = \mathbf{I} \quad (4.16)$$

Whitening can be done by using the eigen-value decomposition (EVD) of the covariance matrix $E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} = \mathbf{O}\mathbf{D}\mathbf{O}^T$, where \mathbf{O} is the orthogonal matrix of eigenvectors of $E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\}$ and \mathbf{D} is the diagonal matrix of its eigenvalues, $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$. Whitening is done as:

$$\tilde{\mathbf{x}} = \mathbf{O}\mathbf{D}^{-\frac{1}{2}}\mathbf{O}^T\mathbf{x} \quad (4.17)$$

Where the matrix $\mathbf{D}^{-\frac{1}{2}} = \text{diag}(d_1^{-\frac{1}{2}}, \dots, d_n^{-\frac{1}{2}})$ and now $E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} = \mathbf{I}$. Whitening transforms the mixing matrix into a new one, $\tilde{\mathbf{A}}$. From equations (4.7) and (4.17):

$$\tilde{\mathbf{x}} = \mathbf{O}\mathbf{D}^{-\frac{1}{2}}\mathbf{O}^T\mathbf{A}\mathbf{s} = \tilde{\mathbf{A}}\mathbf{s} \quad (4.18)$$

Whitening makes the new mixing matrix $\tilde{\mathbf{A}}$ an orthogonal:

$$E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T\} = \tilde{\mathbf{A}}E\{\tilde{\mathbf{s}}\tilde{\mathbf{s}}^T\}\tilde{\mathbf{A}}^T = \tilde{\mathbf{A}}\tilde{\mathbf{A}}^T = \mathbf{I} \quad (4.19)$$

Whitening reduces the mixing matrix into orthogonal and thus eliminates the number of parameters to be identified from n^2 in case of original mixing matrix \mathbf{A} to $n(n-1)/2$ in case of orthogonal matrix $\tilde{\mathbf{A}}$. It can be said that whitening solves half of the problem of ICA.

4.5.3 Filtering

When we are certain that the observed data consists of unwanted data at certain range of frequencies it might be useful to filter out them before processing by ICA. The time filtering of observed data vector \mathbf{x} is done by multiplying it by a filtering matrix \mathbf{F} as:

$$\mathbf{x}^* = \mathbf{x}\mathbf{F} = \mathbf{A}\mathbf{s}\mathbf{F} = \mathbf{A}\mathbf{s}^* \quad (4.20)$$

which shows that ICA model still remains valid, with the same mixing matrix after applying a filter. So, the linear filtering of the signals is allowed [4].

4.6 ThinICA Algorithm

The contents in this section are referred to [27]. ThinICA is one of the several algorithms used in ICA. It uses a multivariate contrast function for the blind signal extraction of a selected number of independent components from their linear mixture. All the independent components can also be separated if needed. It combines the robustness of the joint approximate diagonalization techniques (such as SOBI) with the flexibility of the methods for blind signal extraction [27]. By maximizing the contrast function it gives two options: a) Hierarchical extraction based upon thinQR factorizations or b) Simultaneous extraction based upon thin Singular Value Decomposition (SVD) factorizations [27].

The signal model is identical to the general ICA model. The basic principle of source estimation is similar to the general principle of ICA described in section 4.4 and 4.5. The difference of ThinICA is the measure of independence. It estimates the independent components by jointly maximizing a weighted square sum of cumulants of fixed order $q \geq 2$, determined by tuples $\tau = (t_1, \dots, t_q)$ contained in the set Ω . The contrast function is given as [27]

$$\psi_{\Omega}(y) = \sum_{\tau \in \Omega} \omega_{\tau} |Cum(y(t_1), \dots, y(t_q))|^2$$

$$\text{subject to } \|U\|_2 = 1 \quad (4.21)$$

where ω_{τ} are positive weighting terms given after summing up the cumulants at q different times and \mathbf{U} is a unitary matrix. $\mathbf{y}(t)$ denotes output or estimated source i.e.,

$$\mathbf{y}(t) = \mathbf{U}\mathbf{z}(t) = \mathbf{U}\mathbf{W}\mathbf{x}(t) = \mathbf{U}\mathbf{W}\mathbf{A}\mathbf{s}(t) \quad (4.22)$$

The notations carry the usual meanings. The sources are estimated by either hierarchically maximizing or simultaneously maximizing the contrast function. For simplifying the optimization of contrast function in equation 4.21, alternative similar contrast functions are utilized. Refer to [27]. The algorithm can be summarized as follows. The observed data are first whitened. An arbitrary unitary matrix is chosen for initialization. The square matrices are formed following the contrast function. These matrices are diagonalized either by hierarchical or simultaneous approaches to determine new unitary matrix. Unitary matrix is updated. And finally the independent sources are estimated using equation 4.22.

When the number of independent components to be extracted are equal to the number of sources and $q = 2$, it is equivalent to SOBI, based on the joint approximate diagonalization (JADE) of a certain set of cumulant slices. However, implementation is different. In other cases, it is superior to other algorithms since it offers both the extraction of selected number of components or separation of all the components. The ThinICA algorithms maximize the contrast function by combining simultaneously the several advantages of powerful techniques like Fast-ICA, JADE and SOBI.

As for this research, several other ICA algorithms are utilized such as Natural Gradient-Flexible ICA, PowerICA, Fixed Point ICA etc using ICALAB Toolbox in Matlab [3, 2].

Chapter 5

Time frequency Analysis

5.1 Introduction

The Fourier transform is used as a common tool for representing frequency characteristics of a time series signal. However, the Fourier transform doesn't provide the time information, so it is inadequate. The variation of frequency characteristics with time can be of major interest in many practical cases. So the joint time-frequency [28, 15] representation is needed. Wavelet analysis is becoming a common tool for studying time frequency characteristics of a time series data [9].

In order to study the time variation of frequency characteristics of observed gravity data and apply time-frequency filtering, wavelet transform is used here.

This chapter briefly introduces the Discrete wavelet analysis and Continuous wavelet analysis that are applied for the time-frequency analysis of gravity data.

5.2 Discrete Wavelet Analysis

The wavelet transform decomposes a time series signal into wavelet coefficients at time-frequency domain. The discrete wavelet transform (DWT) of a time series $x(t)$ is given as

$$\tilde{X}(\tau, s) = \sum_{n=0}^{(N-1)} x(n\Delta t) \psi\left(\frac{n\Delta t - \tau}{s}\right) \quad (5.1)$$

where s denotes scale, τ denotes time-shift, $n = 0, 1, 2, \dots, (N - 1)$ being the time index, N being the total number of discrete data and time $t = n\delta t$.

It decomposes a time series signal into wavelet coefficients at different scales (or frequencies) and at various time-shifts as shown in Fig. A.3 in Appendix A. ψ denotes an orthogonal wavelet function as shown in Fig. A.1 in Appendix A. The wavelet function is normalized to unit energy. The inverse wavelet transform is given as

$$x(t) = \sum_{\tau} \sum_s \tilde{X}(\tau, s) \psi\left(\frac{n\Delta t - \tau}{s}\right) \quad (5.2)$$

A scale s represents the range of frequencies. If the total number of time series data is 2^m , the number of scales is limited to m discrete set of scales [19, 9]. If these number of scales are numbered from $j = 1$ to m , the total number of wavelet coefficients at each scale is given by 2^{j-1} . This means that the number of coefficients at higher scales are higher than at lower scales. Further, the higher scales represent the larger range of frequencies. So the frequency resolution is finer and time resolution is coarser towards lower scales. Similarly, the frequency resolution is coarser and time resolution is finer towards the higher scales. The advantage of DWT is that the time series signal can be reconstructed by inverse transform using equation 5.2. This allows us to reconstruct the time series signal after discarding the unwanted coefficients at desired time and scales. However the coarser time and frequency resolutions might prevent some freedom in data processing. For example, the useful information at lower frequencies might be missing due to coarse time resolution. In order to utilize the finer time and frequency resolution, continuous wavelet transform is useful.

The further details about discrete Hardy wavelet analysis are presented in a related paper by this author in Appendix A.

5.3 Continuous Wavelet Transform

The continuous wavelet transform (CWT) of a time series signal $x(t)$ is defined as the convolution of $x(t)$ with a scaled and translated wavelet function ψ_c as [9]

$$\tilde{X}(\tau, s) = \sum_{n=0}^{(N-1)} x(n\delta t) \psi\left(\frac{n\delta t - \tau}{s}\right) \quad (5.3)$$

where δt is the time interval and N is the total number of time series data. In fourier domain, it is given as

$$\tilde{X}(\tau, s) = \sum_{k=0}^{(N-1)} \hat{x}_k \hat{\psi}(s\omega_k) e^{i\omega_k n \delta t} \quad (5.4)$$

where $k = 0, \dots, N - 1$ is the frequency index, \hat{x}_k is the discrete Fourier transform of $x(t)$ and ω_k is the angular frequency given by

$$\omega_k = \frac{2\pi k}{N\delta t} \quad (5.5)$$

Unlike in DWT, the wavelet function in CWT is non-orthogonal (such as Morlet wavelet), so the signal reconstruction by inverse transform is not possible.

In CWT, a set of scales s can be chosen according to the need, so the frequency resolution can be made finer than in DWT. The time resolution in all scales are equivalent to original time series data and the continuous variation in wavelet spectrum is observed.

Representing scales as fractional powers of two [9]:

$$s_j = s_0 2^{j\delta j} \quad (5.6)$$

$$J = \delta j^{-1} \log_2(N\delta t/s_0) \quad (5.7)$$

where s_0 is the smallest resolvable scale and J determines the largest scale, $j = 0, 1, \dots, J$. The s_0 should be chosen so that the equivalent Fourier period is $N\delta t$. The choice of a δj affects the scale (frequency) resolution. Smaller values of δj give finer resolution and vice versa. For example, Let us consider $N = 4096$ and $\delta t = 2sec$. Let us choose $\delta j = 0.125$. Then from equation 5.7, the total number of scales becomes $J = 88$. The time-frequency representation by CWT can be visualized in the figure 5.1 below.

5.4 Time frequency analysis of Gravity data

The gravity survey was performed in a ship. The motion of ship affects the data observation system. Depending upon the frequency of tilting or frequency of vibration of ship, the observed data is affected accordingly. The ship tilting frequency varies depending upon the variation in wind velocities and sea wave frequencies. In order to study the time varying frequency characteristics and apply the joint time-frequency filtering, DWT and CWT are used.

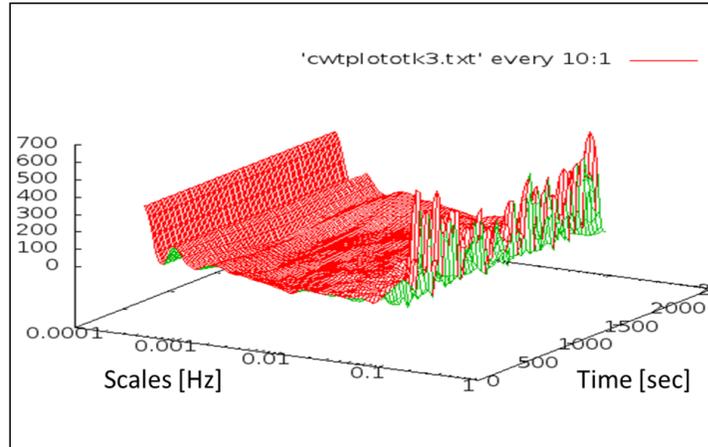


Figure 5.1: Time-frequency representation of VSE (2 sec interval sampled 4096 data) using CWT.

The discrete wavelet analysis is used for the purpose of applying joint time-frequency filter to the observed gravity data as a pre-processing to Independent Component Analysis (ICA). The orthogonal Hardy wavelet function [26] is used that decomposes the time series gravity data into Hardy wavelet coefficients. The coefficients at certain time and scales are only chosen and inverse transform is applied to reconstruct the signal. The sets of time-frequency filtered reconstructed signals are then applied to ICA for further processing. The procedures and results are discussed in chapter 7.

The gravity anomaly data is expected to be dominant at lower frequencies. Extracting the coefficients at lower frequencies or scales is the main interest. DWT offers limited time resolution at lower frequencies and the frequency resolution is also coarse. So, the finer time and frequency resolutions offered by CWT at lower scales might be more advantageous. The time series gravity data is decomposed into continuous wavelet coefficients. The relation in equation 5.6 is utilized to select the desired number of scales corresponding to desired cut-off frequency. For example, scales 1-43 correspond to cut-off period 200 seconds and scales 1-51 correspond to cut-off period 100 seconds. The coefficients at certain time and scale are selected. Since the reconstruction of time series signal is not possible in CWT, the coefficients are directly processed by ICA. The detail procedure and results are discussed in chapter 7.

Chapter 6

Gravity Data Observation

The site for gravity survey was located at Toyama Bay of Toyama Prefecture, Japan. The survey team along with the author visited the site for data observation on Oct 31, 2010. A small ship was used for the survey. The instrument setup consisted of prototype gravimeter setup synchronized with accurate Global Positioning System (GPS hereafter). The base station of GPS was setup at the bay and another GPS was setup in the moving ship. The accelerometer sensors observed the gravity data, which were displayed or recorded in the computer. The synchronized position of the moving ship read by GPS and was recorded in the computer. The observed noisy raw data was retrieved in the office for processing. This chapter includes the description of prototype gravimeter, observed raw data and the reference data obtained using the available gravity map for Toyama bay.

6.1 Survey Site and Data observation

Toyama bay is located at around 250 kms north-west from Tokyo. The site was chosen because the gravity map developed by Government of Japan was available for that place and that could be used for obtaining the data for reference. The reference data is used for comparison with the results obtained by processing the observed raw data. In order to introduce the general condition of the observation on a carrier, the prototype gravimeter was setup on a mid-size ship. Owing to the instability issues, such ship offers the much worse data observation environment compared to a large-size ship. The gross weight of ship was 231 tonnes, its length was 53.59 m and its maximum speed being 15 knots. The total survey length was around 20 kms. The survey site and related information can

be observed in the figure (6.1). The ship started moving at time 13:30:00 and followed the South-North route until time 14:22:00 and covered the distance around 10 kms as shown in the map. Thus the average velocity while going was around 11 km/hr. From time 14:22:00 to 14:50:00, the ship stopped for some other purposes not related to gravity survey. It started moving back towards the bay at time 14:50:00 and reached back to bay at 15:22:00. The average speed while returning back was thus around 18 km/hr. The ship followed almost the same route while getting back to bay. The total time taken for the ship navigation was around 3 hours. The ship had generally a variable velocity. The VSE data observation was started well before the ship started moving i.e., from 12:27:00. The digital servo data observation started from 12:39:00 and the accelerometer titan started observing data from 12:53:00. The data from time 13:30:00 to 15:22:00 is of main concern for further processing.

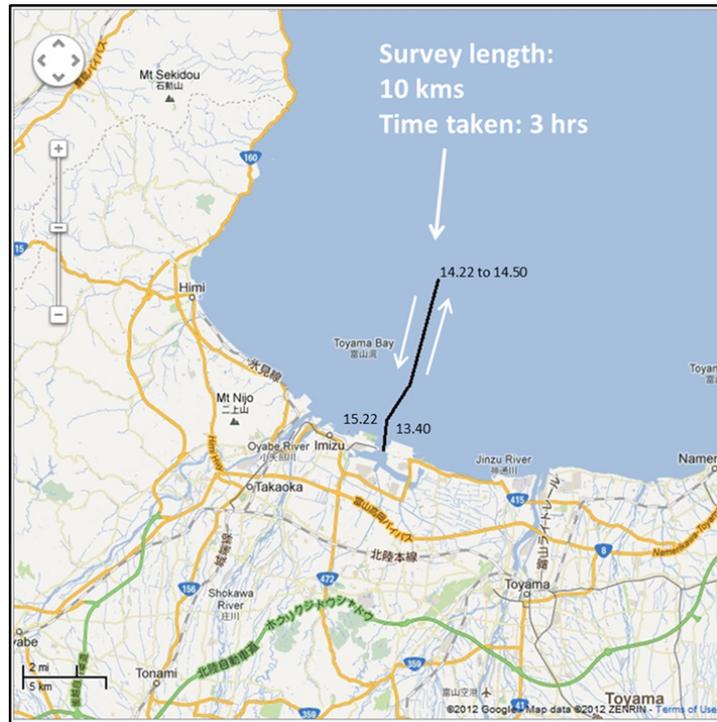


Figure 6.1: Gravity Survey Site at Toyama bay (Google map) with survey route and time.
Date of survey: 2011/10/31

6.2 Prototype of Gravimeter (EZ-GRAV)

In order to examine the performance of a force-balanced-type accelerometer, a prototype was developed, named as EZ-GRAV [8]. Our purpose is to apply the gravity anomaly data for modelling a sub-surface structure. This is further used for the estimation of earthquake ground motions that require a gravity map with contour lines of one milligal interval. Thus, our target accuracy of the gravity survey is in the order of sub milligal, i.e. from 10^{-5} to 10^{-4}cm/s^2 . To obtain this accuracy, the sensor unit have to satisfy the resolution of at least 10^{-5}cm/s^2 since the observed data is contaminated by severe noise and the subtle gravity anomaly data demands the highest possible degree of precision. EZ-GRAV used VSE-156SG (hereafter VSE) by Tokyo Sokushin Company Limited as its sensor and included gradiometers and accelerometer with three components, i.e., one vertical and two horizontal. The observed data is recorded in digital format with 24 bit and 0.01 second interval. All the sensors are fixed on an aluminium thick plate and set inside a constant temperature reservoir, because the devices are sensitive to temperature and its fluctuation is too larger than the variation of gravity. The temperature is controlled within $\pm 1^\circ\text{C}$. The features of prototype EZ-GRAV are listed in table (6.2) and the picture of setup can be seen on figure (6.1).

Table 6.1: Features of Prototype EZ-GRAV

Description	EZ-GRAV
Sensor for gravimeter	VSE
Accelerometer Titan	2 horizontal and 1 vertical components
Gradiometer	2 horizontal components
Recording interval	0.01 s
Max. Input Voltage	± 10 V
Constant temperature reservoir	N/A

6.2.1 VSE (Analog Servo)

The basic specification of VSE is listed in Table (6.2).

Table 6.2: Basic specification of accelerometer VSE

Description	VSE-156SG
Observable dynamic range	± 50 Gal
Maximum Output Level	± 10 V
Sensor Resolution (Accuracy)	$2\sim 10 \times 10^{-6}$ Gal

6.2.2 Accelerometer Titan

The accelerometer Titan are of two kinds, namely Taurus and Trident. Both of these sensors consists of three sensors each, oriented at three different directions, i.e., North-South (NS component), East-West (EW component) and Up-down (vertical or UD component). The NS and EW components record the acceleration due to carrier movement, while the UD component refer to gravity anomaly.

The accelerometer Taurus has observable range of $\pm 4 \times \mathbf{g} = 4 \times 980 \text{gals}$ and Trident has observable range of $\pm 0.25 \times \mathbf{g} = 0.25 \times 980 \text{gals}$. These sensors have very high observable dynamic range but they do not offer as high resolution as given by VSE. However, we require multiple sets of data for data processing by Blind Signal Separation techniques, the data observed by these sensors are also useful.

6.3 Synchronization of Gravimeter with GPS

The gravimeters were synchronized with accurate GPS in order to have control on the position of moving ship. The accuracy of GPS was in the order of few cms. Thus the observed time series data of gravity anomaly in the moving ship can be later converted into spatial resolution. The total distance covered by ship to and from Toyama bay was around $\mathbf{x} = 2 \times 10 = 20$ kms, and the total time of travel was around $\mathbf{t} = 4$ hours. So in case of real project, the \mathbf{t} hours time series data should be finally converted to \mathbf{x} km spatial data for the purpose of preparing gravity contour map. This conversion should be done after the data processing i.e., after extracting the gravity anomaly data from the noisy observation. The time series data is used for data processing because it is convenient for recording and processing.

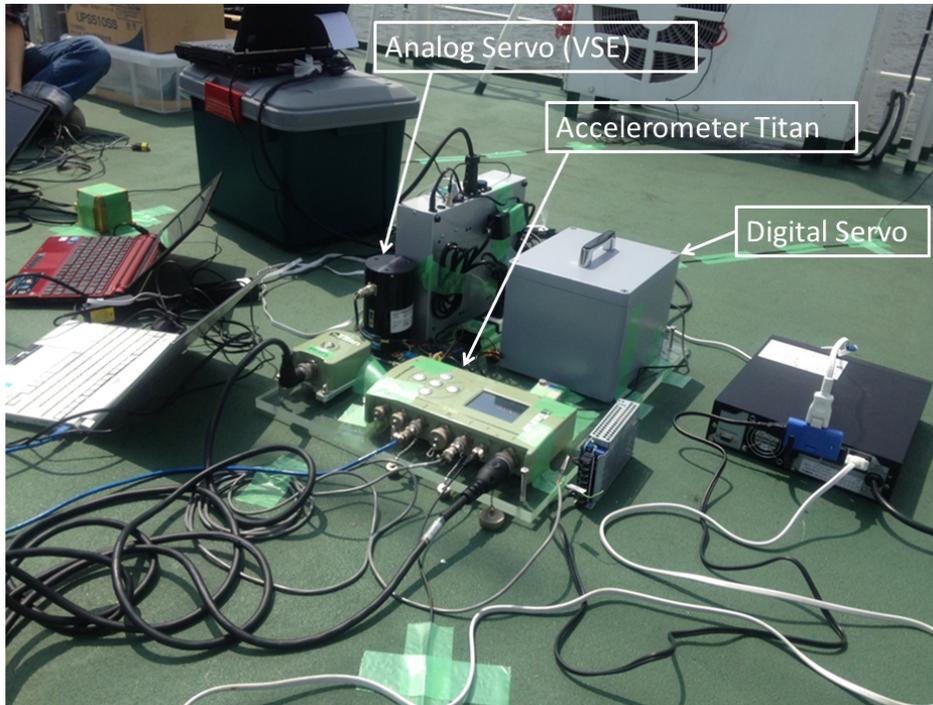


Figure 6.2: Prototype gravimeter setup (EZ-GRAV: Analog Servo, Accelerometer Titan).

6.4 Raw Gravity Data

The observed raw noisy data as expected, was contaminated with severe noise. The sources of noise were ship movement, engine vibration, sea waves, fluctuating wind speed, electrical noise etc. Further, there are several factors that govern the gravity anomaly, such as, tidal effect, topography, air pressure, sensor drift, elevation and Eötvös effect. The tidal effect are in the order of 20 microgals, and reference data are in the order of milligals, so they are neglected. Similarly, as the survey is done at sea, the effect due to elevation, topography and air pressure are assumed to be almost zero. However, the sensor drift and Eötvös effect should be considered and should be deducted from the processed data before projecting it into a gravity map. This section shows the observed time series data with two different sensors.

6.4.1 Analog Servo (VSE) Data

Owing to its high resolution, the VSE data is supposed to be the major data set amongst all the other data. The time series of recorded data can be observed in the figure (6.3)

below. The data recording was started on 12:27:00. The time interval of recorded data was 0.01 seconds.

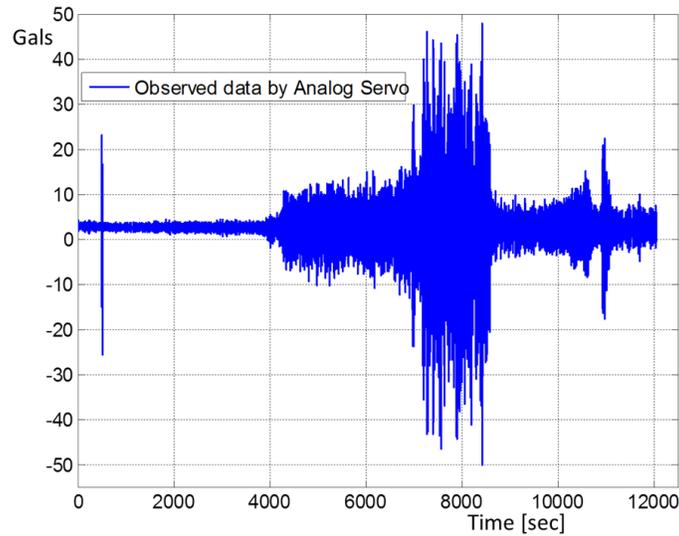


Figure 6.3: Observed time series gravity data (Gals) by Analog Servo (VSE) starting at time 12:27:00.

6.4.2 Accelerometer Titan Data

As explained in subsection (6.2.2), the accelerometer Titan consists of two sets of sensors with different observable ranges. Both of them record three components of data. These sensors started recording the data from time 12:53:06. The time interval of data was 0.01 seconds.

6.4.2.1 Taurus

The vertical (Z) component of Taurus data is shown in figure 6.4. The other components can be seen in figure (ICA input without filtering) of chapter 7.

6.4.3 Digital Servo Data

This is a different accelerometer with observable range ± 400 gals i.e., higher than Analog servo. The data observed by this sensor was not considered to be reliable. The picture of digital servo data before and after filtering is shown in Appendix B.

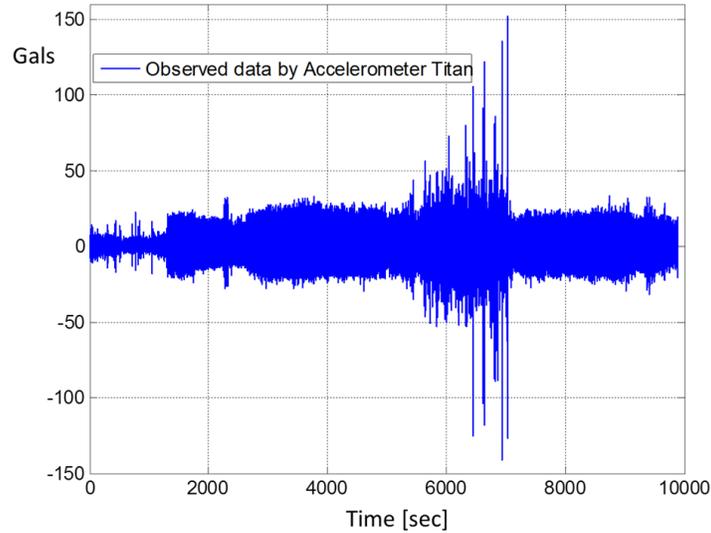


Figure 6.4: Observed time series gravity data (gals) by Accelerometer Titan (Taurus: vertical component), starting at time 12:53:06.

6.5 Reference Data from Gravity map

The reference data was obtained by adding the Eötvös effect due to the ship movement and the free air anomaly. The free air anomaly was obtained from the gravity map that was prepared by Government of Japan by using a shipboard gravity survey. This survey was performed by accurate shipborne gravimeter in a much more stable environment using a huge ship. This gravimeter was not sensitive to high frequency noise and reads only low frequency gravity data. Besides the accuracy of gravimeter, the survey was performed along the lines making a grid, and so the multiple data at points of intersection were averaged. Thus, the reference data is supposed to be accurate and reliable. And, the Eötvös effect is calculated from the position of the ship every second obtained from GPS. The reference data can be observed in the figure (6.5).

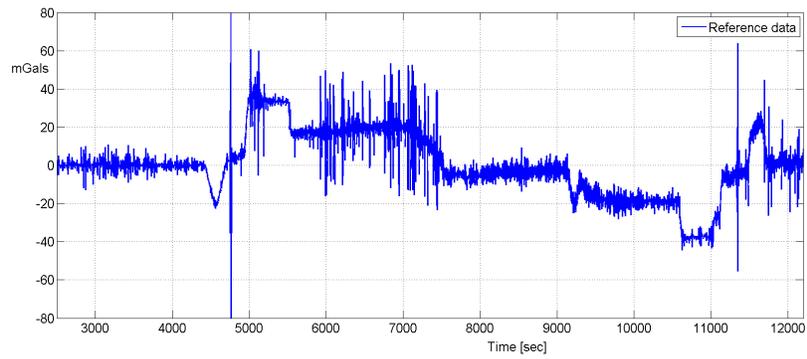


Figure 6.5: Reference data (Eotvos effect + Free air anomaly) obtained by using Gravity map produced by AIST, Japan, starting time: 12:17:00.

Chapter 7

Application of Noise Reduction and Separation Methods

7.1 Introduction

The computation results obtained by various methods of data processing are presented in this chapter. The data processing tools such as low pass filter, time-frequency analysis and blind source separation (BSS) techniques were applied that are presented in individual sections with comparison of their results in the final section. For the purpose of BSS techniques like SOBI and ICA, ICALAB toolbox [3] in MATLAB were used. There are several types of ICA algorithms [2] out of which mainly ThinICA is used with occasional use of NG-FICA and PowerICA algorithms.

7.2 FIR Low Pass Filter (LPF)

The simple finite impulse response low pass gaussian filter (hereafter FIR LPGF or simply LPF) was applied to the raw data. LPF simply filters out the high frequency components of the data. It can also be called as frequency-filter. The data processing is done for two different cut-off periods, i.e., 100 seconds and 200 seconds. The simple algorithm applied for the FIR LPGF is presented in chapter 3. The various components of data from the sensors before and after applying the filter are presented here. The two important sets of data i.e., Analog servo (VSE) and the vertical component of Accelerometer Titan (Taurus-Z component) after applying LPF with cut-off period 100 sec are only displayed in this

section. Out of all the sets of data, these two sets of data roughly followed the trend with reference data corresponding to the time when ship is moving towards the sea. At the time ship turns back, the data suddenly jumped away from the reference data and then shows somewhat opposite trend to reference data. The comparisons are presented in final section of this chapter. Since the other sets of data like Taurus NS and EW components are not very significant, they are presented in Appendix 2.

Since the observed data after filtering doesn't follow the trend of reference data well, it is felt that simple filter was not enough for this data. The application of time-frequency filtering and BSS techniques are presented in following sections.

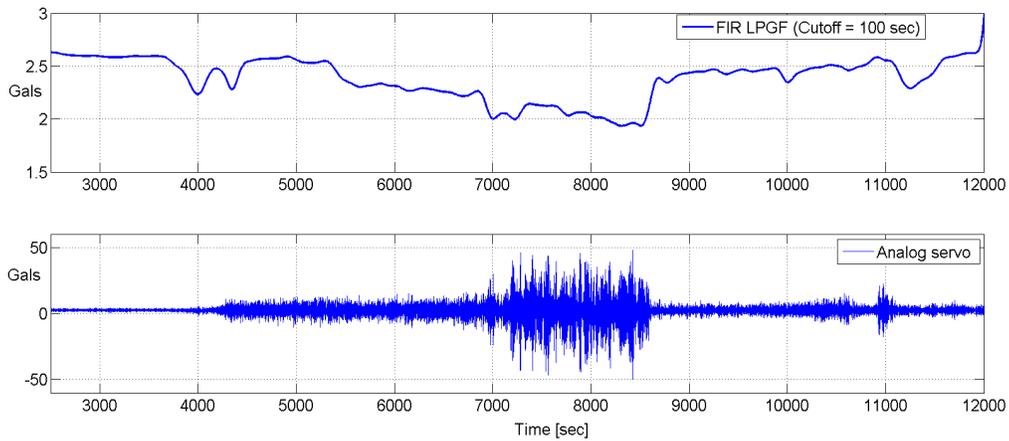


Figure 7.1: a) Top: LPF applied to VSE data (cut-off 100 sec) and b) Bottom: Analog servo (VSE) data.

7.3 Discrete Hardy Wavelet Analysis(DWT) and ICA

Since the simple frequency filter is not enough for de-noising gravity data, an endeavour is made to apply the time-frequency filter using Discrete Hardy wavelet analysis (hereafter DWT). This tool decomposes the time series data into time-frequency domain. The coefficients can be selected or removed at the required time or frequency jointly. The discrete Hardy wavelet is orthogonal so it allows the reconstruction of signal. The details on discrete Hardy wavelet analysis are presented in chapter 3 and the related paper is presented in Appendix A.

The certain number of coefficients are chosen in desired time and frequency jointly the

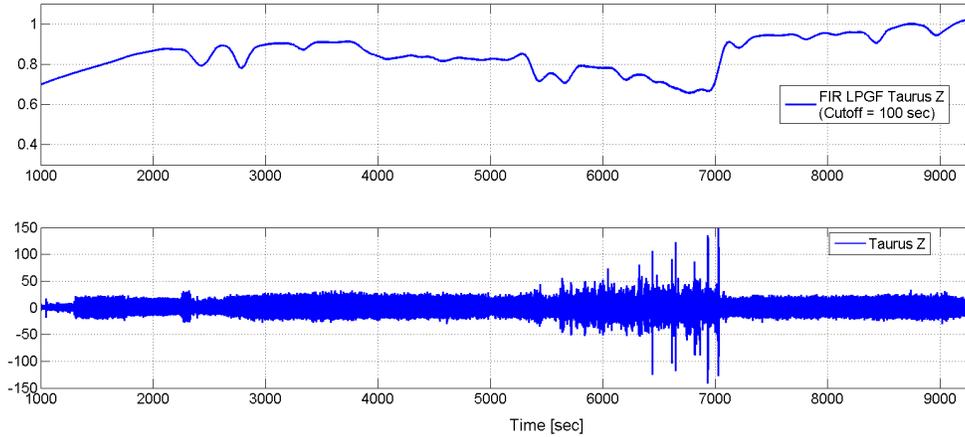


Figure 7.2: a) Top: LPF applied to Taurus-Z data (cut-off 100 sec) and Bottom: Accelerometer Titan (Taurus-Z component).

other coefficients are discarded. The signal is then reconstructed to obtain the time-frequency filtered time series. The reconstructed signal corresponding to VSE data looked similar to the one after applying LPF, so it is further processed with ICA together with reconstructed Taurus-Z data. The applied procedure is mentioned below.

The two sets of data i.e VSE and Taurus-Z are selected. They are sampled from originally 0.01 sec interval to 0.1 sec interval and truncated from 13:30:00 to 15:22:00 excluding some portion from 14:22:00 to 14:25:00. This makes the total number of data as a power of 2 i.e., ($2^{16} = 65536$ in this case). These data are decomposed into time-frequency domain by using discrete Hardy wavelet analysis. The total number of scales (or frequency ranges) is 16. The coefficients at higher frequency ranges (scales greater than 6 in this case) are filtered out (i.e.made zero). This is roughly equivalent to LPF with 60 second cut-off period. For 6-th scale, the coefficients lying within the time range corresponding to ship stoppage time, i.e., 14:22:00 to 14:50:00, are made zero. Then the inverse discrete Hardy wavelet transform is applied to obtain the reconstructed data. The reconstructed filtered data are shown in figure 7.3.

The reconstructed data corresponding to VSE is similar to LPF VSE data while Taurus-Z is not. These data are then processed by ThinICA algorithm. The results are shown in figure 7.4 below. The Hardy wavelet filtered data look unaffected by ICA. The inconvenience in processing was felt due to the course frequency resolution in DWT.

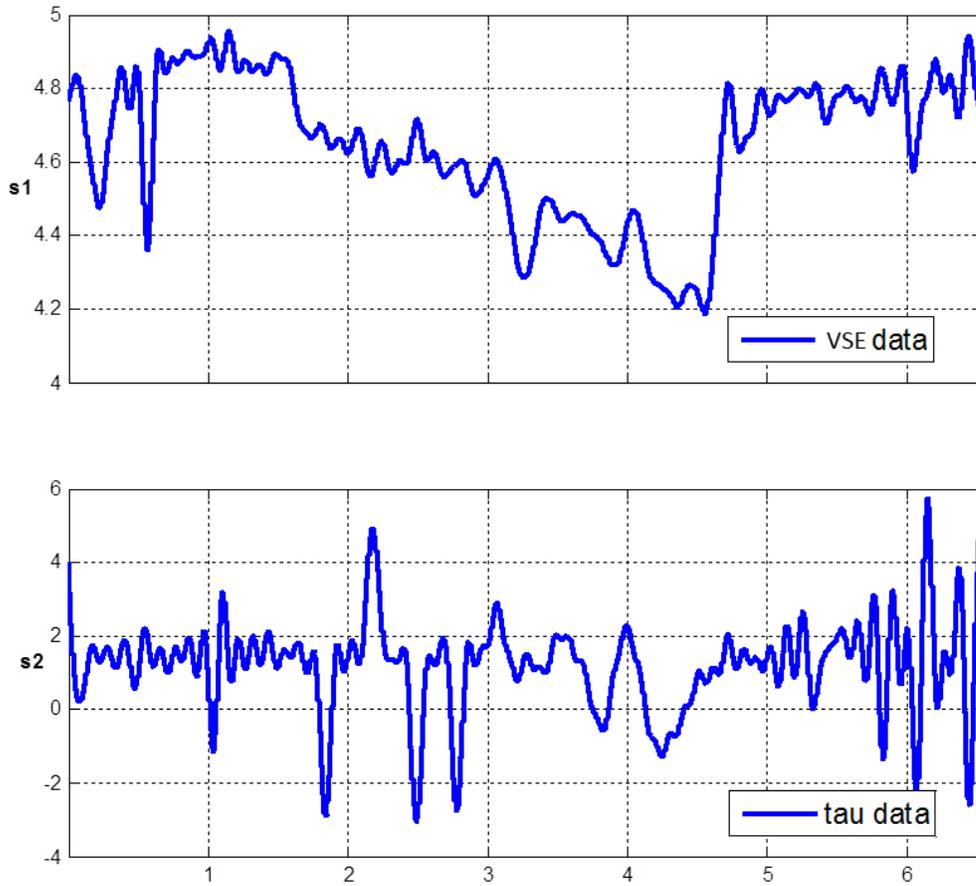


Figure 7.3: Reconstructed signals after applying Hardy wavelet low pass filter (Scale < 6 or around 60 sec cut-off period); a) Top: VSE data and Bottom: Taurus Z-component. Time starts from 13:30:00 to 15:22:00.

7.4 Continuous Wavelet Transform (CWT) and ICA

The number of coefficients corresponding to lower frequency ranges are much less in discrete wavelet analysis (DWT) while the number of coefficients are equivalent to the time series data at all frequency ranges in CWT. Since gravity data is supposed to lie at lower frequency ranges, we can utilize the enough number of coefficients in lower frequency ranges in case of CWT. Moreover, unlike in DWT, the frequency resolution is finer and can be chosen according to the need in case of CWT. An effort is done to apply CWT for time-frequency analysis of observed gravity data before processing by ICA. The procedure followed is as follows.

The VSE and Taurus data are sampled from originally 0.01 second to 2 second interval.

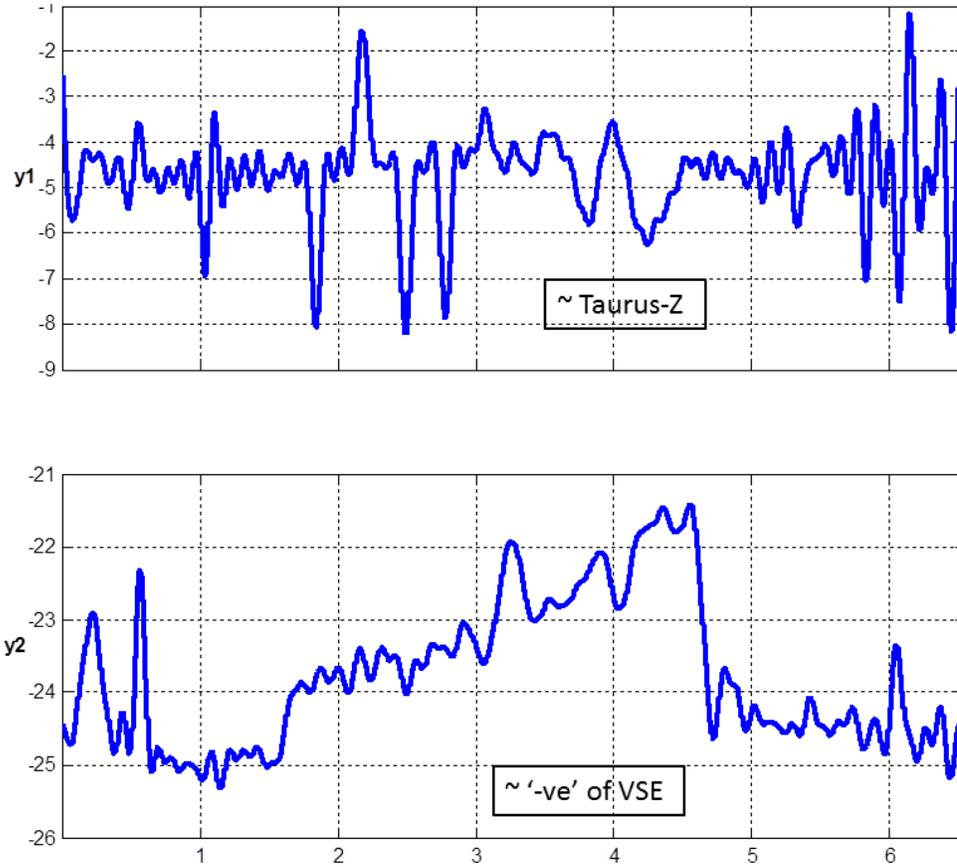


Figure 7.4: Output by ThinICA: Same as input (No separation); a) Top: Corresponds to Taurus-Z data and Bottom: Corresponds to negative multiple of input VSE.

They are truncated from time 13:05:28 to 15:22:00. This makes the total number of data as 4096 (2^{12}). These data are then decomposed by CWT using Morlet wavelet. The frequency resolution or total number of scales (frequency ranges) are chosen to be 88. The details of CWT are presented in chapter 5. Then the real part of the CWT coefficients corresponding to scales 1 to 43 are selected and listed in series (i.e. 43×4096 number of data). The coefficients higher than scale 43 are discarded (Note: Scales 1 - 43 correspond to frequency < 0.005 Hz or cut-off time 200 sec). For scales 1 - 43, the coefficients (real part) corresponding to ship stoppage time i.e., 14:22:00 to 14:50:00 are made as zero. Thus the coefficients are selected at desired time and frequencies. Unlike in DWT, the reconstruction of signal is not possible in CWT as the Morlet wavelet is not orthogonal. So the real part of CWT coefficients are directly used for processing by ICA. The two series of 43×4096 data corresponding to VSE and Taurus-Z data are then input to ICA.

The de-mixing matrix 'W' computed by ICA is utilized. This de-mixing matrix 'W' was multiplied (refer chapter 4 for ICA model and equation 7.1) to FIR LPF (cut-off 200 sec) VSE and Taurus Z data i.e., cut-off frequency equivalent to chosen CWT scales 1-43. One of the outputs by ICA shows very little improvement to the input data when compared with reference data and VSE LPF data. The de-mixing matrix 'W' mentioned above is given as

$$W = \begin{bmatrix} 0.33631923 & 0.039618536 \\ -0.30437288 & 0.14431874 \end{bmatrix}$$

$$\{y_i\} = [W]\{x_j^f\} \quad (7.1)$$

where y_i corresponds to de-mixed or separated data and x_j^f corresponds to LPF (200 sec) VSE and Taurus-Z data.

The same process was followed for LPF (cutoff = 100 sec) and choosing the CWT scales 1-51 out of total 88 scales. However, the results were similar. Since the results are not significant, the relevant figures (with de-mixing matrix) are shown in Appendix B and only the comparisons are shown in the last section in this chapter.

7.5 LPF and SOBI

The application of simple frequency filters and time-frequency filters is not so fruitful. So the advanced BSS techniques are applied. The application of SOBI to the raw data without a filter is not effective and the results are not shown here. So LPF (cutoff time 100 sec or 200 sec) are applied first as pre-processing to SOBI. Several trials are done such as a) choice of selected portions of data by truncating b) choice of cut-off period (100 sec or 200 sec) or c) choice of various combinations of data by different sensors. The best performance are seen with the combination of VSE and Taurus-Z data when pre-processed by LPF with cut-off period 100 sec and truncated from time 13:30:00 to 15:22:00 i.e., time corresponding to ship motion. The application of SOBI showed improvement when compared to VSE LPF and reference data. The sudden jump in VSE LPF data corresponding to the time ship starts moving back from the sea after stopping for some time, is improved. The input to SOBI and its output are shown in figures 7.5 and 7.6

respectively. The comparison with various methods are shown in final section. The demixing matrix corresponding to figure 7.6 is

$$W = \begin{bmatrix} 0.50537654 & 1.2070863 \\ 4.946234 & -0.20708634 \end{bmatrix}$$

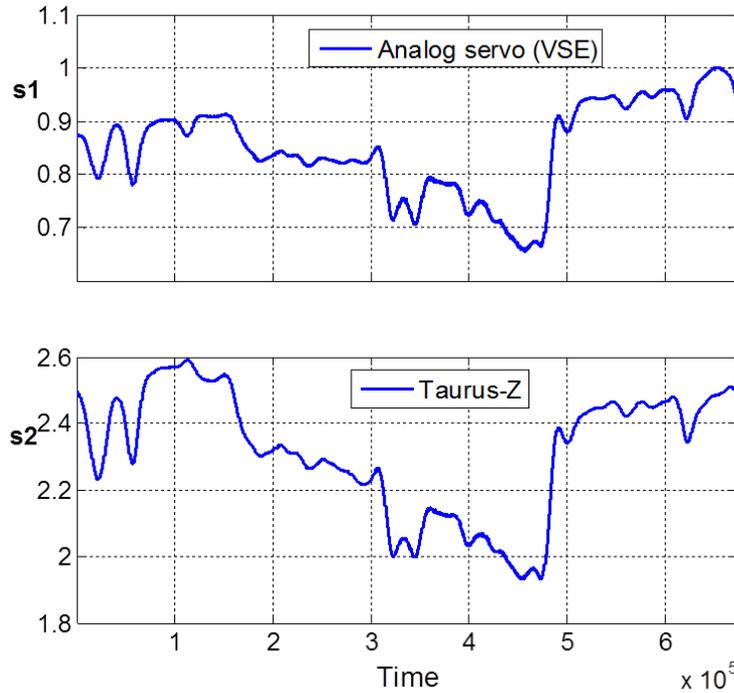


Figure 7.5: Input data to SOBI: VSE and Taurus-Z data after applying FIR LPF (cutoff = 100 sec, Time: 13:30:00 to 15:22:00).

7.6 LPF and ICA

Although the results by SOBI are improved than LPF data, it has some theoretical limitations as explained in chapter 4 and ICA is proposed as a appropriate methodology in this thesis. In order to verify the applicability of proposed ICA, its application to the data was done. The various combination of data were processed and the best combination is found to be that of VSE and Taurus-Z component. VSE data was expected to be the most reliable one and the results supported the expectation.

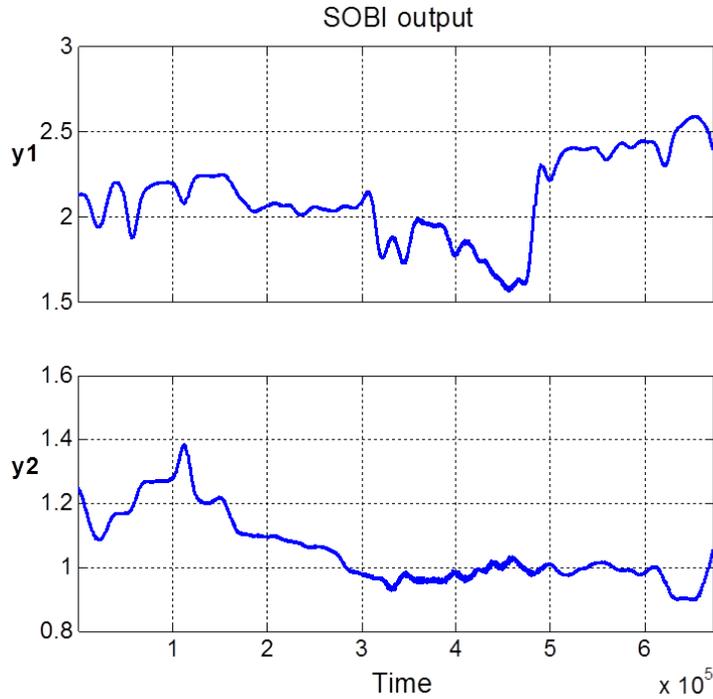


Figure 7.6: Output by SOBI for the input in previous figure 7.5.

The observed sets of data without pre-processing by a low pass filter are input to ICA, but they do not respond to ICA. The output were equivalent to the input. For an example, results by NG-FICA algorithm are shown in figures 7.7 and 7.8 . They match perfectly when input and output are compared individually (figures are not shown here).

Since the data didnt respond to ICA without filtering, FIR LPF is applied as pre-processing to ICA. The application of ICA after filtering shows improved results. Upon the choice of LPF cut-off periods, the results are better when cut-off period 100 sec is used, compared to cut-off period 200 sec. The results with cut-off 200 sec are presented in appendix B. Upon the choice of selected portions, the truncated data corresponding to ship motion time gives the better result when compared to the data without truncating. The truncation in time domain was done from 13:30:00 to 15:22:00, the time corresponding to the movement of ship away from bay. Further improvement is observed when the ship stoppage time corresponding to 14:22:00 to 14:50:30 is discarded. The input data is shown in figure 7.5 (same input to SOBI) and the output are shown in the figure 7.9. The

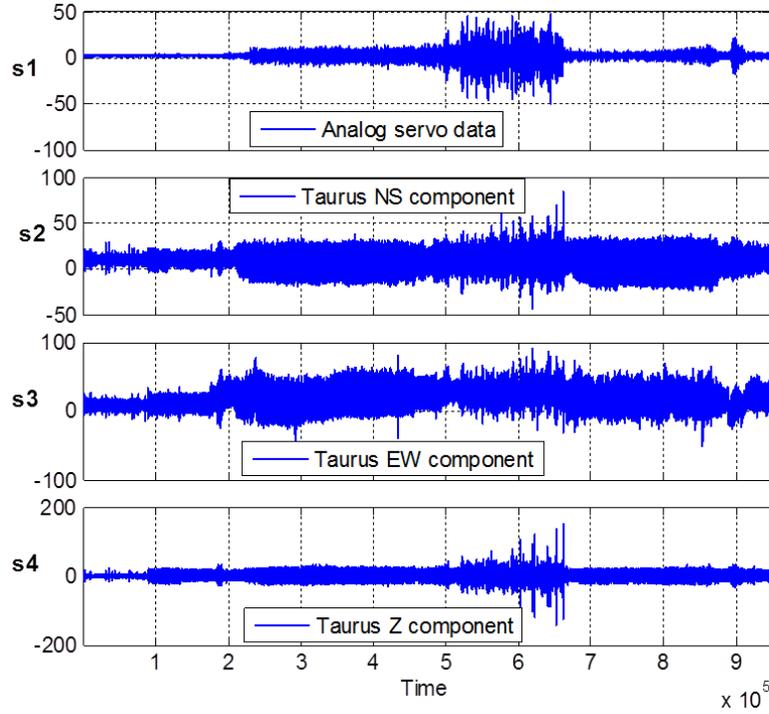


Figure 7.7: Input sets of data (VSE and Taurus-NS, EW, Z) for ICA.

demixing matrix 'W' was obtained corresponding to figure 7.9 is

$$W = \begin{bmatrix} 5.9140419 & 7.0819445 \\ -24.769004 & 11.325292 \end{bmatrix}$$

Note that the de-mixing matrix compared to SOBI is different although the output are similar to certain scalar multiples.

The other several trials are performed with the inputs with combination of VSE with other sets of Accelerometer Titan (Taurus) like Taurus-NS and Taurus-EW etc. Few of those significant trials are presented in the appendix B and best results are presented in this chapter.

7.7 Comparison of Results

The comparison of the outputs with reference data is shown in this section. For the ease in comparison, LPF is also applied to the reference data.

The LPF VSE data, Hardy wavelet (time-frequency) filtered data and its combination

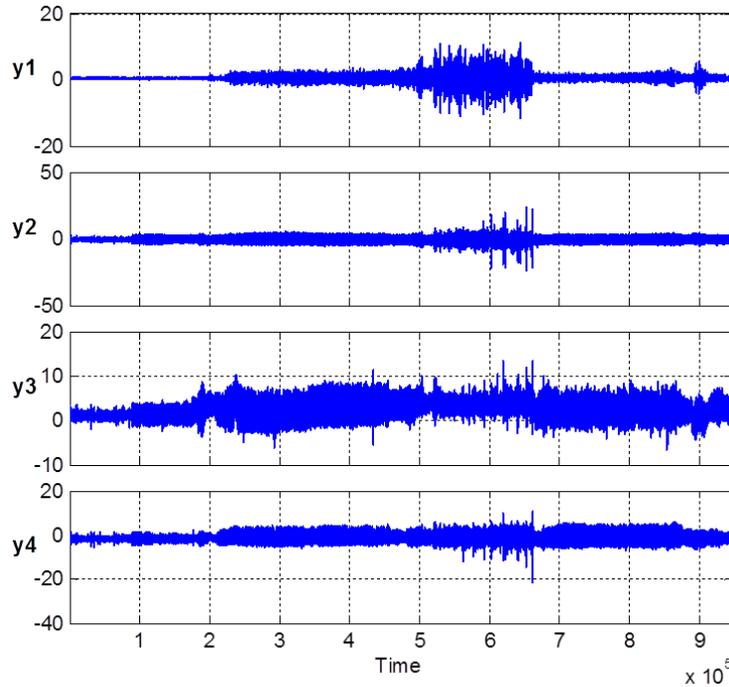


Figure 7.8: Output by NG-FICA algorithm: No difference from input.

with ICA looked almost similar. As it is also obvious in the figures 7.3 and 7.4, the comparison are not shown here.

The comparison of LPF VSE data, output of CWT combined with ThinICA and the reference data are shown in figure 7.10. Some improvement can be seen on the output data (i.e., at around time 6000 seconds in the figure, the offset has been considerable reduced in the demixed2 data compared to FIR LPGF VSE data). It should be noted that the time origin in all the figures in this chapter are not same.

Then the comparison of LPF (cut-off 100 sec), signal separated by SOBI, and the one separated by ThinICA are compared in figure 7.11. Until time around 3120 seconds in the same figure, both SOBI and ThinICA outputs are similar. From 3120 to 4800 seconds, SOBI is more stable than ThinICA output. It should be noted that this time corresponds to the ship stoppage time 14:22:00 to 14:50:30. During this period, the ship is tilting heavily and more frequently. After 4800 seconds or 14:50:30, ThinICA output seems to follow much better with the reference data than any other results. Also it looks to have improved much more from LPF VSE as seen in figure 7.12.

From time 13:30:00 to 14:22:00 corresponds to 0 to 3120 seconds in figure 7.12. During

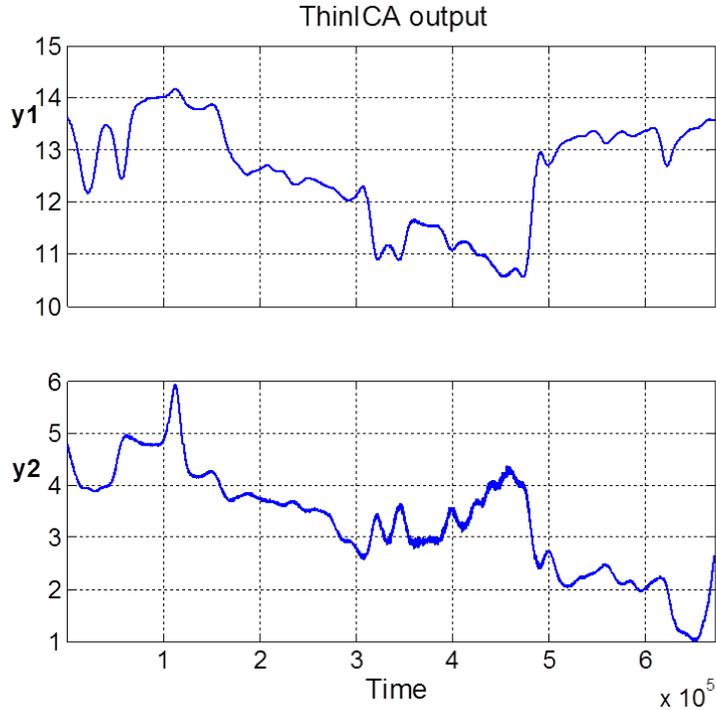


Figure 7.9: Output by ThinICA: refer figure 7.5 for input (same input given to SOBI).

this time, ship is moving away from Toyama bay as mentioned in the chapter 6. The time 14:22:00 to 14:50:30 corresponds to 3120 to 4800 seconds in the same figure 7.12. During this time, ship is stopped. From 14:50:30 to 15:22:00 or from 4800 seconds to 7200 seconds in same figure, the ship moves back to the Toyama bay through almost the same route. Since it is stopped at one location from 3120 to 4800 seconds, the data during this period are of no significance. So this duration is also truncated and the results are compared as shown in figures 7.13 to 7.16.

7.8 Conclusion and Implications

First, the application of simple frequency filter to the observed gravity data is done. The filtered data looks to follow the trend or phases of reference data at some portions. The filtered data provides the hint that gravity anomaly data is inherent in the data. However at time around 14:50:00 (when ship just started moving back after stopping for a while), the filtered data showed sudden jump and distraction from the reference data. After 14:50:00 it never followed the trend of reference data.

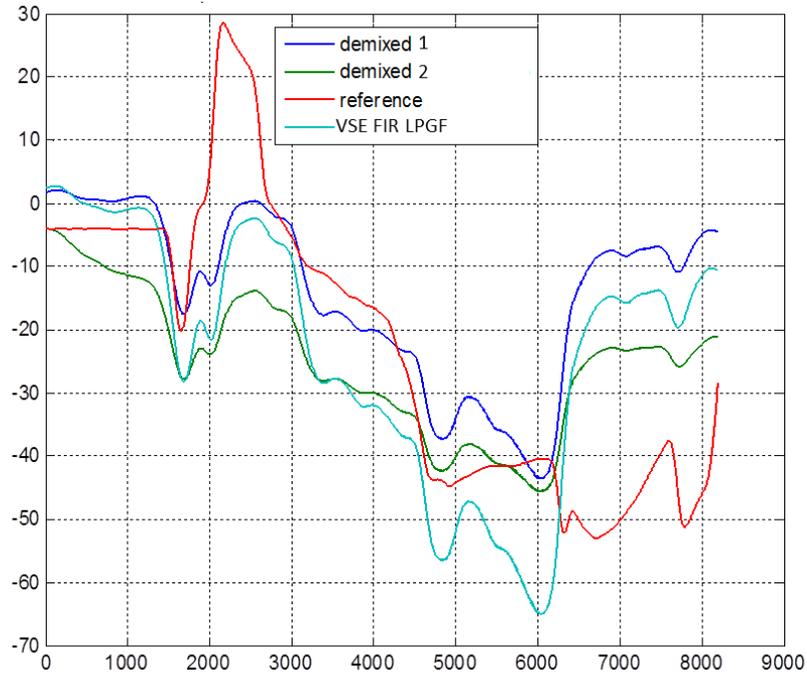


Figure 7.10: Comparison of two de-mixed (separated) data (using demixing matrix for CWT coefficients input to ThinICA) with LPF VSE data and Reference data: Demixed2 data shows some improvement as the reduction in jump at around 6000 seconds (or time 14:50:00) is observed.

The merit of application of time-frequency filter combined with ICA is also not clear in the results.

Then BSS techniques like SOBI and ICA are applied. Since the signal to noise ratio is very less, the data doesn't respond to BSS methods when applied without any pre-processing. When LPF filter is applied as a pre-processing, the data responds well to both SOBI and ICA. Since the outputs by BSS methods are always normalized to unit variances, they are multiplied by appropriate scalars in order to compare with reference data. It is observed that the troublesome jump in LPF data at time around 14:50:00 is corrected well.

The choice of appropriate filter is realized to be important. The choice of 100 sec cutoff filter shows better performance compared to cut-off 200 sec. The cut-off 100 sec corresponds to frequencies less than 0.01 Hz and cut-off 200 sec corresponds to frequencies less than 0.005 Hz. The spectrum of reference data also shows that the dominant frequencies are lesser than 0.01 Hz. So, the cut-off 200 sec LPF might incur loss of some useful infor-

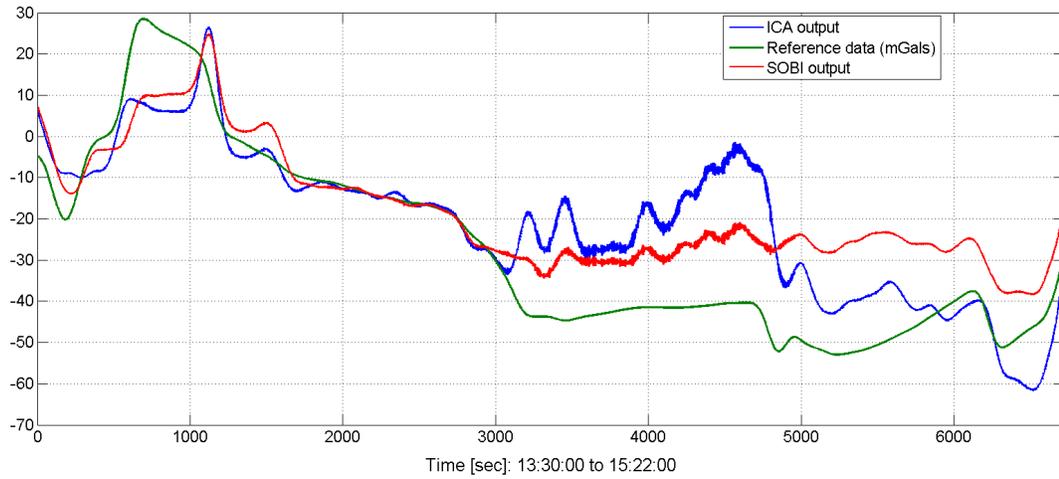


Figure 7.11: Comparison of outputs by SOBI (Red), ThinICA (Blue) with Reference data (Green): SOBI is better at ship-stoppage time (3120 to 4800 seconds or 14:22:00 to 14:50:30) and ThinICA is better after ship motion starts at 4800 seconds (i.e., at time 14:50:30).

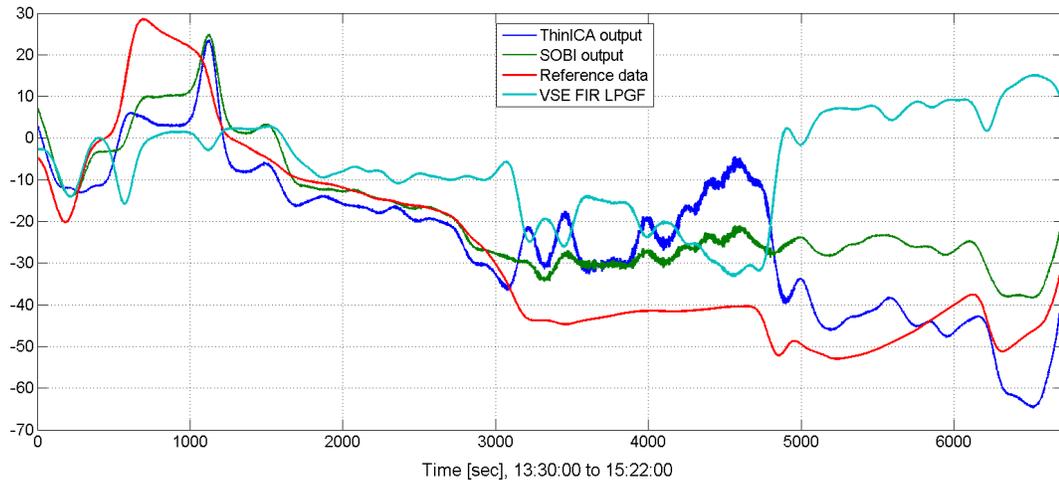


Figure 7.12: Comparison of outputs by ThinICA (Blue), SOBI (Green), VSE LPGA (light green) with Reference data (Red). VSE LPGA data is included to above figure 7.11: ICA separated data is far improved than LPGA data.

mation in the data. The choice of appropriate cut-off period is also governed by frequency ranges of all kinds of disturbances during carrier motion.

Both SOBI and ICA are showing the improved performances. The de-mixing matrices by

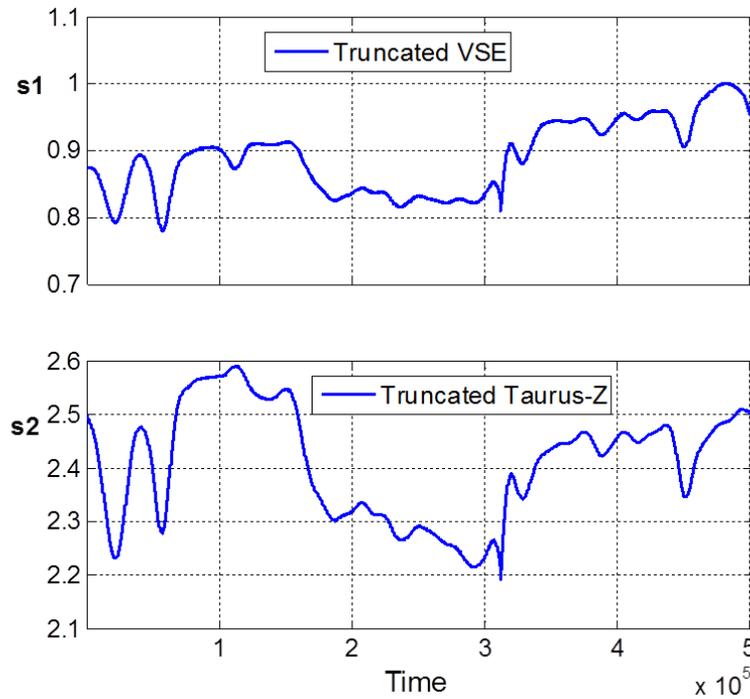


Figure 7.13: Truncated Input LPF VSE and LPF Taurus-Z data: Ship stoppage time (14:22:00 to 14:50:30) removed.

the two methods are different and their scales are also different. During the early period, when the ship moves out of the bay, it has to change its direction frequently. The change in direction incurs the changes in Eotvos effect. The frequent change in this effect might be the reason for the in-harmony between the output results with reference data.

When the LPF VSE and Taurus-Z data from 13:30:00 to 15:22:00 are input to SOBI and ICA, the output are little bit different. SOBI output seems to be more stable at ship stoppage time, but during the ship motion time, ICA output looks to be more corrected than LPF VSE data and also more harmonious with reference data as seen in figures 1.11 and 1.12. When the ship stoppage time is discarded and input to SOBI and ICA, both of them gave similar results and both were far improved than LPF VSE data with reference to reference data as seen in figure 1.16.

Further, when ship cruises away from the bay, its speed is moderate. During this period, output data is perfectly matching with reference data and performance of ICA is good. When ship cruises back to bay, speed is higher. During this period, output signal is away from reference data although it is following some trend. When the ship stability

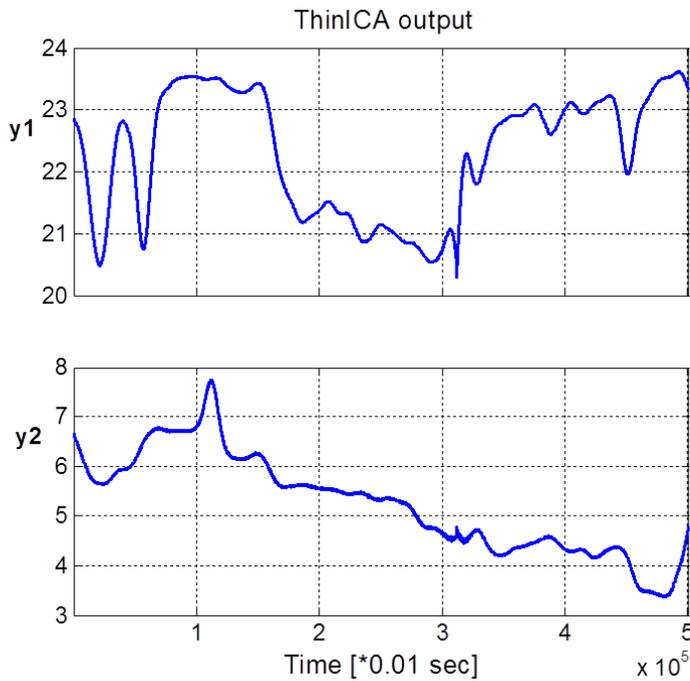


Figure 7.14: Output by ThinICA for truncated input with Ship stoppage time (14:22:00 to 14:50:30) removed.

was heavily affected during its stopping in the middle of sea, the ICA performance is not good. The sensitivity of data acquiring environment is realized to affect the performance of ICA. The LPF VSE and LPF Taurus-Z data look very much similar to each other. Since they both are recording the same gravity data this is quite natural and as expected. They are also most closer to reference data compared to other components and they serve to be the best combination to be applied for BSS. Based on this experience it may be implicated that the two data showing the similar trends may offer the best combination for further processing by BSS or they may be closer to the actual gravity anomaly trend. The output given by ICA has satisfied that the gravity anomaly and other noises are statistically independent. Since the output data followed the trend of reference data the applicability of ICA is verified.

ThinICA and SOBI results both look quite close in this case. Both of them are far more improved than LPF VSE data as can be easily observed in the figure 7.16.

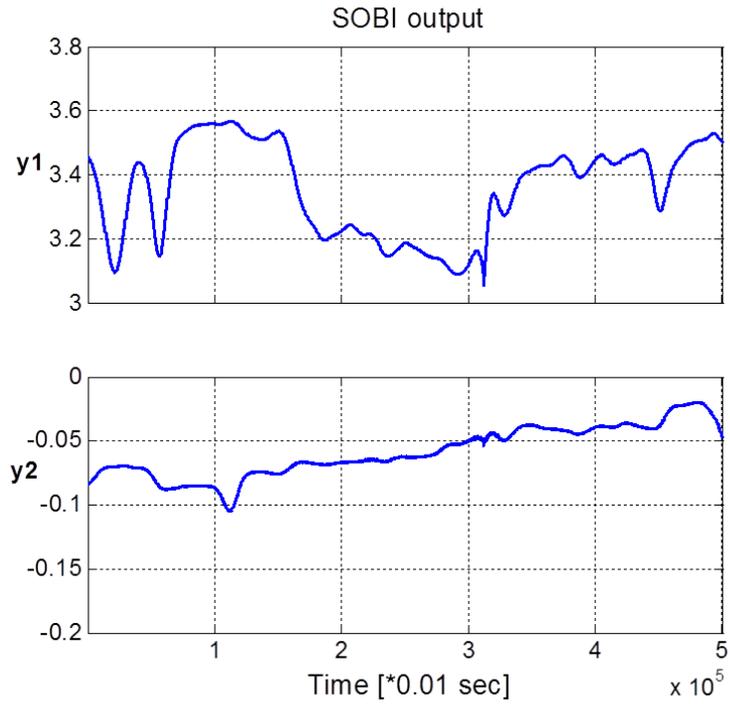


Figure 7.15: Output by SOBI for truncated input with Ship stoppage time (14:22:00 to 14:50:30) removed.

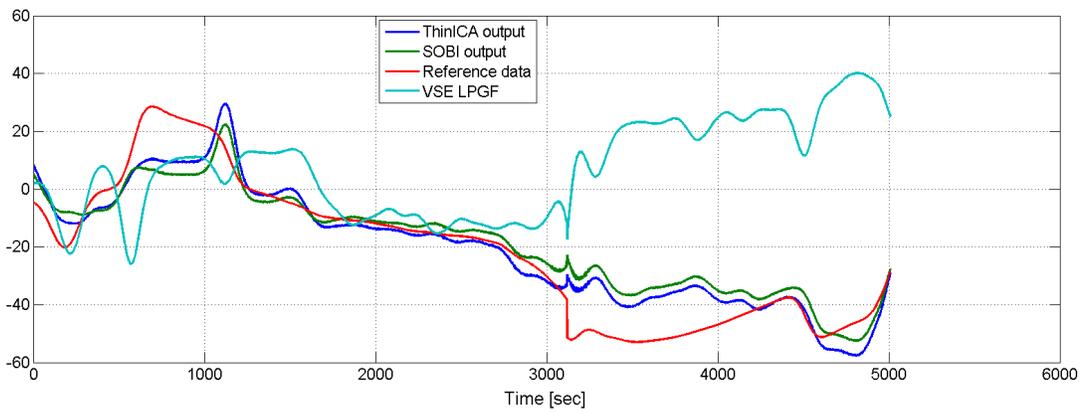


Figure 7.16: Comparison of output by ThinICA and SOBI with VSE LPF and Reference data (mGals), for truncated input with ship stoppage time removed: Both SOBI and ThinICA show improvements than LPF data with similar performance.

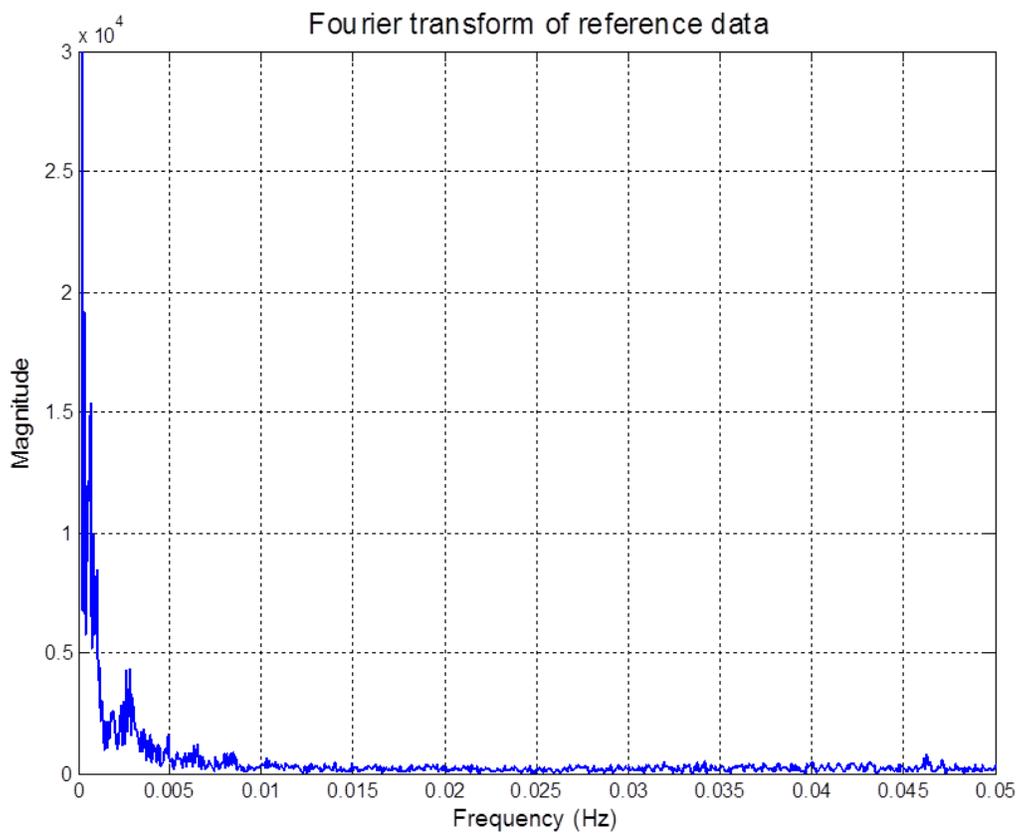


Figure 7.17: Reference data in frequency domain: Frequencies less than 0.01 Hz are observed to be dominant.

Chapter 8

Conclusions

The recently developed prototype mobile gravimeter (Team Morikawa) uses FB accelerometer sensor. FB sensor is highly sensitive to high frequency noises. Since it is developed to be used in a small carrier, they are associated with higher instability. The carrier motion, vibration, frequent tilting etc. yields to severe contamination on recorded data. In order to extract the gravity anomaly from such data, an appropriate data processing methodology is needed.

The conventional gravimeters were less sensitive to high frequency noise and the data observation was done in the stable environment using a large ship or carrier. Since the observed data were less affected by noise the conventional filtering technique was often enough for de-noising. The second order statistics method using correlation were also used. However, the noises associated in the data by FB sensor are far more severe on both amplitudes and ranges of frequencies. The conventional filtering techniques works well but not perfectly. So it is necessary to develop an appropriate methodology for the purpose of gravity anomaly data separation. Since the gravity data and other noises are originating from different physical processes, we can make an (weak) assumption that the gravity anomaly data and the noises are statistically independent. The conventional second order statistics method (such as SOBI) separates the sources from their mixture by considering that the source components are mutually un-correlated at several time-lags. SOBI is a good method, but it uses un-correlation, which is a weaker form of independence. Under certain conditions, the mutually un-correlated data are not mutually independent. Thus we propose to use more strict condition, i.e. independence of signals. Recently an advanced BSS technique known as Independent Component Analysis (ICA) has been widely used as a method of considering independence in several fields.

By knowing only the mixed signals, ICA estimates the original sources by maximizing independence of separated signals. Independence is measured by higher order moments (like fourth order moments in Kurtosis) in contrast to un-correlation that is measured by second order moments. Since the mutual independence of signals is not affected by their amplitudes, ICA is suitable for our purpose where we are dealing with the very low Signal/Noise ratio (up to 0.00001). ICA is unable to identify the amplitude of original signal and its output are normalized but it does not create a problem for our purpose since it is only about determining an appropriate scalar multiple.

We observe the gravity data at Toyama bay, Japan. The high quality data produced by AIST, Japan is available for reference. With the help of this reference data we are able to verify the performance of proposed method. The prototype consists of multiple sensors. The Analog servo (VSE) is the major sensor and it is accompanied by other sensors. Since ICA requires minimum two sets of data, the performance of all sensors are compared in an association with VSE data.

Assuming that gravity anomaly data is dominant at lower frequencies, high frequency components are filtered out using LPF. Such frequency based filtering is realized to be important. Until the high frequency components are filtered out using a LPF, the data do not respond to any BSS method. The presence of high frequency noises is understood to be hazardous to data separation. ICA or SOBI works only after low pass filtering. The choice of an appropriate cut-off period for LPF should also be given importance. Results clearly show the effects of difference in cut-off frequency. The effectiveness of ICA at various conditions are analysed. The ICA results are good for the combination of VSE data with vertical component of Taurus (Acc. Titan). When more than two combinations are input to ICA, the results are not good. The other horizontal components of Titan also do not make a favourable association with VSE data. Further, the importance of data acquisition environment is realized. At the durations when ship stability is severely hampered, such as during the ship stopping time on the sea, ICA results are eccentric to the trend of reference data. While going away from bay, the average ship speed is around 11 km/hr. At most of that duration, ICA result is perfectly matching with reference data. While returning back to bay, the ship velocity is roughly around 18 km/hr. During this period the result is not as harmonious as the former section when ship speed was lesser. At the moment, ship is changing the direction rapidly, ICA output signal is deviated from reference data. So the three factors carrier stability, carrier speed and frequency in change of carrier direction are realized to play important role to describe the effectiveness of ICA.

The very good separation of components satisfies the assumption that gravity anomaly and other sources are mutually independent. The harmony of ICA separated data with reference data verifies the applicability of the proposed method, at certain conditions of data observation environment. The further improvement in data processing methodology is considered to be the part of future works.

Based on the current performance, it can be concluded that the results are encouraging and they are at acceptable limit for the purpose of subsurface modelling. The consistency in the results in future and further improvements, if possible, will certainly lead to the improvement in mobility of gravity method. The mobility of gravity method will not only facilitate the economic combination of multiple subsurface survey methods but also will facilitate with the continuous sets of data leading to abundance of information on subsurface strata. The improved accuracy in subsurface modelling will contribute to improve the quality of GM simulation and seismic design.

Appendices

Appendix A

Discrete Analytic Signal Wavelet Decomposition

This appendix is derived from the paper published by this author jointly with Prof. Riki Honda for 15WCEE and 9CUEE [26] and [25]. The title of the paper is "Discrete Analytic Signal Wavelet Decomposition for Phase Localized in Time-Frequency Domain for Generation of Stochastic Signal with Phase Uncertainty".

Summary

Fourier transform has been widely used in dynamic analysis so far. The stochastic ground motions can be generated by altering the Fourier phase in original ground motion. However, the time-frequency information is only obtained by wavelet transform. But usually, when the phase of a wavelet basis function is altered, the total power of real part is affected and thus the generation of artificial stochastic ground motions is not possible. In this context, we propose to use analytic signal hardy wavelet analysis that describes the time-frequency characteristics and allows consideration of phase uncertainty for generation of stochastic ground motions. This paper verifies the performance of presented scheme with numerical simulations showing more localized disturbance in generated artificial ground motions compared to the conventional scheme. It is also verified that the same transmitting function as in Fourier transform can be used for evaluating the response of a linear structural system.

A.1 Introduction

The dynamic analysis of structures is essential for modern seismic design, such as performance-based design. Since the powerful computing environment is available these days, we are able to perform complicated numerical computations for seismic design and the importance of design input ground motion is emphasized.

Fourier transform has been widely used in dynamic analysis. It analyses an input signal defining both phase and amplitude at various frequencies. It uses sinusoidal base function ($e^{i\omega t}$) that has shift-invariance and orthogonal properties. The shift-invariance property assures the conservation of total power of the base function after the phase is altered, while the orthogonal property assures the unique decomposition and perfect reconstruction of signal. And because of these two properties, the generation of stochastic ground motions considering phase uncertainties is possible using Fourier transform.

Fourier transform also has inconvenient property. Firstly, the temporal variation cannot be explained explicitly since the sinusoidal base function extends from $-\infty$ to $+\infty$ and the exact time cannot be estimated. When the behaviour of a non-linear system is discussed, the time-frequency information of the input signal is very important because once the non-linearity is reached during the vibration, the behaviour of structure might be different after every cycle of loading and unloading depending upon the variation of natural frequency of structure with time. Secondly, the power is distributed throughout the time history of the generated ground motion and the desired localized disturbance is not possible.

Recently, wavelet transform has been adopted for time-frequency analysis. In particular, orthogonal wavelet transforms are widely used since they allow inverse transforms and therefore are suitable for wave synthesis. Since the wavelet transform provides explicit time-frequency representation, concept of phase is not important for the sole purpose of temporal representation. Phase can be manifested in wavelet coefficients, but in usual cases, as the phase is altered, the total power of real part is also changed and thus the uncertainties in phase cannot be considered for the purpose of generation of stochastic ground motions.

In this paper, we present an analytic signal hardy wavelet as a base function (as shown in Figures A.1 and A.2). Since it is an analytic signal, the shift-invariance property is attained and hardy wavelet inherently has an orthogonal property. These properties in presented scheme enable the explicit definition of both 'phase' and 'amplitude' with the

time-frequency representation. The analytic wavelet coefficients can be expressed into the phase and amplitude like the Fourier coefficients. Then the artificial ground motions can be generated by considering uncertainties in phase spectrum and also the localized disturbance can be achieved in the generated ground motion unlike in Fourier transform.

This paper is organized as follows. The analytic signal is defined and the analytic wavelet transform is explained in section A.2, signifying the importance of shift-invariance and orthogonal properties. It is highlighted that analytic Hardy wavelet signal enables phase localized in time-frequency domain in the same section. Section A.3 explains the representation of phase uncertainty with the comparison among Fourier transform and analytic wavelet transform with numerical simulations. Section A.4 shows the derivation of relation between Fourier coefficients of wavelet coefficients of an input signal and its response. The relation is also verified by the numerical simulations and the procedure is explained in the same section. Section A.4 also highlights the convenience maintained in wavelet transform due to the conservation of transmitting function. Finally, section A.5 concludes the paper.

A.2 Analytic Wavelet Transform

In this paper, Wavelet transform that uses analytic Hardy wavelet signal as a base function, is proposed as a tool for localized time-frequency analysis that allows the consideration of uncertainties due to phase changes.

A.2.1 Analytic Signal

An analytic signal is a complex signal [15] that has no negative frequency components. A real signal $\psi(t)$ is converted into an analytic signal $\psi_A(t)$ by integrating twice its Fourier transform over the positive frequency range as

$$\psi_A(t) = 2 \int_0^{\infty} \psi(\omega) e^{i\omega t} \quad (\text{A.1})$$

so that the total power of real signal is maintained in the analytic signal. Here, $\psi(\omega)$ denotes discrete Fourier transform of a real signal. For discrete analysis, $\omega = k\Delta\omega$ where $\Delta\omega = 2\pi/N\Delta t$ denotes interval of angular frequency, $k = 0, 1, 2, \dots, (N - 1)$ being the frequency index, while δt denotes time interval, and, time $t = n\delta t$, where

$n = 0, 1, 2, \dots, (N - 1)$ being the time index, and N being the total number of discrete data.

A.2.2 Discrete Analytic Wavelet Transform

Analytic Hardy wavelet is a complex function that has sinc function as its real and imaginary parts as shown in Figure A.1. For numerical simulation, discretized analytic Hardy wavelet signal is reconstructed by inverse wavelet transform using the filter banks. In case of Fourier transform, the sinusoidal base function extends uniformly from $-\infty$ to $+\infty$ in time domain, but a wavelet function has its amplitude localized in certain time duration and rest of the period have almost zero amplitude. Figure A.2 shows the same wavelet function in frequency domain and it has no negative frequency components since it is an analytic signal. This feature helps this tool to retain the shift-invariance property, i.e., the phase change doesn't affect the total power of real part. It is also observed that the amplitude of the Hardy wavelet signal is bounded in a certain frequency range. It equipped the Hardy wavelet with an orthogonal property, i.e., the inner product of wavelet functions with different scale or shift generate zero. Orthogonality assures the unique decomposition of an original input signal and perfect reconstruction of the signal by inverse transform.

The discrete analytic wavelet transform of the time series $x(t)$ is given as

$$\tilde{X}(\tau, s) = \sum_{n=0}^{(N-1)} x(n\Delta t) \psi_A\left(\frac{n\Delta t - \tau}{s}\right) \quad (\text{A.2})$$

where s denotes scale and τ denotes time-shift. It decomposes a time series signal into wavelet coefficients at different scales (or frequencies) and at various time-shifts as shown in Fig. A.3. And, the inverse wavelet transform is given as

$$x(t) = \sum_{\tau} \sum_s \tilde{X}(\tau, s) \psi_A\left(\frac{n\Delta t - \tau}{s}\right) \quad (\text{A.3})$$

A.2.3 Phase and Amplitude localized in time-frequency domain

In Fourier transform, phase and amplitude are defined in frequency domain. In case of wavelet transform, amplitude is defined in time-frequency domain but in usual cases phase is not explicitly defined and we cannot utilize the concept of phase uncertainty in ground motion simulation. So our aim is to present a tool that enables phase localized in

time-frequency domain. As is shown above, an analytic Hardy wavelet function is able to define phase in wavelet function as: $\sum \psi_A(t)e^{i\theta}$. The formulation is almost identical with that of Fourier transform and it also retains shift-invariance and orthogonal properties. Because of the localized property of wavelets, the localized disturbance is possible and ground motion simulation is more efficient compared to that by Fourier transform. It is observed in Section A.3.

In Fourier transform, the response of a linear structure can be evaluated conveniently by the product of input signal and transmitting function in frequency domain. Section A.4 verifies that the convenience is maintained in wavelet analysis since the same transmitting function can be used for the evaluation of response in time-frequency domain.

In Fourier transform, the response of a linear structure can be evaluated conveniently by the product of input signal and transmitting function in frequency domain. Section A.4 verifies that the convenience is maintained in wavelet analysis since the same transmitting function can be used for the evaluation of response in time-frequency domain.

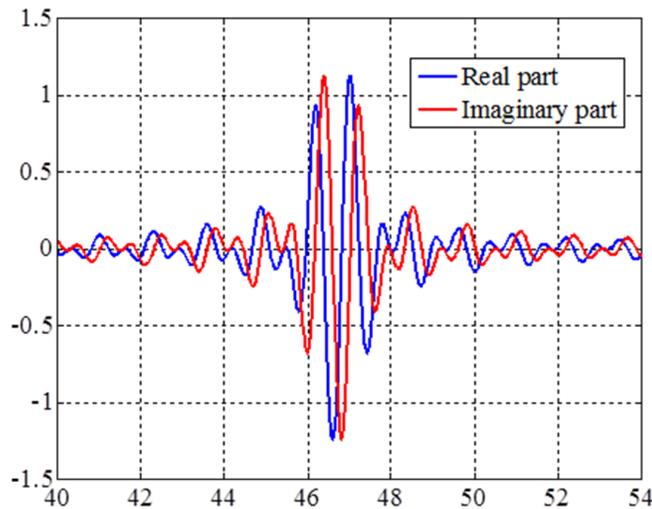


Figure A.1: Time history of an analytic Hardy wavelet signal.

A.3 Uncertainty in Phase

A simple and powerful method for simulating ground motions is to combine parametric or functional descriptions of the ground motion's amplitude spectrum with a random phase spectrum modified such that the motion is distributed over a duration related to

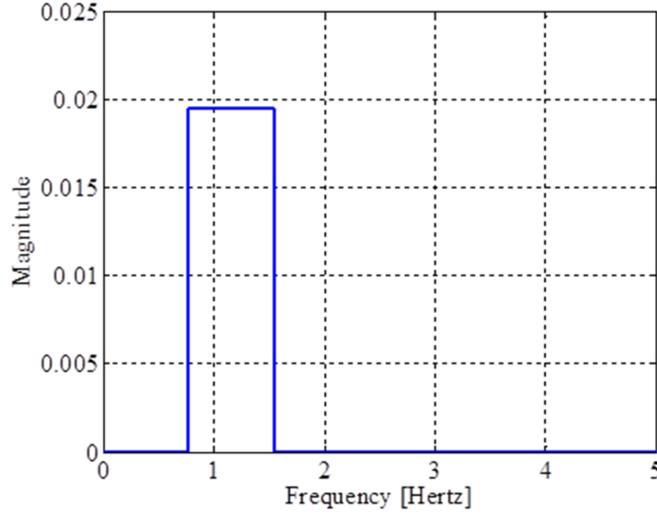


Figure A.2: Fourier transform of analytic Hardy wavelet signal.

the earthquake magnitude and to the distance from the source [10]. This is known as the stochastic method in which a number of stochastic ground motions are generated from the original ground motion by giving random phase perturbation and keeping the total power of signal constant. Phase is nothing but the relative position of signal and is expressed as an angle from 0 to 2π radian. The response of a structure is likely to be different to each of these generated stochastic ground motions. The random phase changes can represent the uncertainty of ground motion. The representation of phase uncertainty for simulation of input ground motion is discussed in this section with comparison among Fourier transform, usual wavelet transform and proposed analytic Hardy wavelet transform.

A.3.1 Phase in Wavelet Transform

When a real wavelet signal is used, the wavelet coefficient $\tilde{X}(\tau, s)$ of Equations (A.2) or (A.3) has no phase. In analogy to the Fourier transform, it is possible to introduce phase θ in the form such as: $|\tilde{X}(\tau, s)|e^{i\theta}$. The definition of this phase is almost identical with the phase of Fourier transform and therefore it is expected that we can make use of conventional methods based on Fourier transform without significant difficulty. However, phase defined in this manner does not satisfy the property required for the manipulation of time series signal when real wavelet function is utilized. For example, if the phase value θ is changed to $\pi/2$, the real part disappears. It means that fluctuation of the phase can cause change in the total power of the signal and it is not suitable for our purpose of

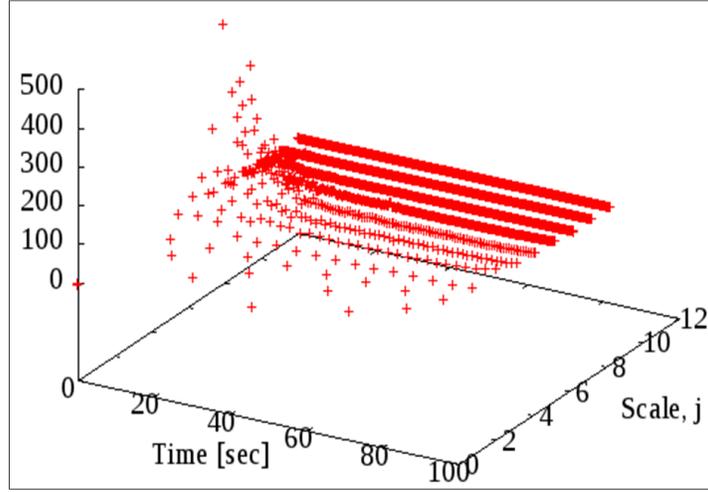


Figure A.3: Discrete Analytic Hardy wavelet coefficients representing time-frequency characteristics of a time series signal.

generating stochastic ground motions.

We use a complex analytic Hardy wavelet signal that is accompanied by the imaginary part as shown in Figure A.1. This analytic Hardy wavelet function defines phase, retains shift-invariance and orthogonal properties like in Fourier transform and thus it makes a tool that is suitable for our purpose.

A.3.2 Numerical Simulations

When a phase uncertainty ξ is introduced in the Fourier coefficient of a signal, the total power of real part is unchanged although the configuration of signal may be totally changed, which can be understood mathematically as follows:

$$x_{\xi}(t) = \sum_{k=0}^{(N-1)} |\hat{X}(k\Delta\omega)| e^{i(\theta+\xi)} e^{i(k\Delta\omega)t} \quad (\text{A.4})$$

where $|\hat{X}(k\Delta\omega)|$ denotes Fourier amplitude of $x(t)$ and θ denotes Fourier phase. In case of wavelet transform, we can assume almost identical equation as

$$x_{\xi}(t) = \sum_{k=0}^{(N-1)} |\tilde{X}(k\Delta\omega)| \psi\left(\frac{t-\tau}{s}\right) e^{i\xi} \quad (\text{A.5})$$

The NS component of a strong ground motion data observed at Kobe observatory during the 1995 Hyogoken Nanbu Earthquake is chosen for numerical simulation. The

total number of Fourier coefficients is 2048 and out of them, around 200 coefficients are dominant that contribute to most of the total power. Then, 50 out of these 200 coefficients are chosen randomly to be disturbed by the phase noise ξ . The ξ value is fluctuated from 0 to π as represented in equation (A.4). Figure A.4 shows the comparison of time series of original wave and the Fourier phase-disturbed wave.

In the similar manner, we apply the perturbation to wavelet phase of the same input signal. The total number of analytic wavelet coefficients is 2048. Then, 40 out of around 200 dominant wavelet coefficients are chosen randomly and disturbed by the phase change uncertainties ξ varying from 0 to π as shown in equation (A.5). Figure A.5 shows the comparison of time series of original wave and the wavelet phase-disturbed wave.

In both cases of Figures A.4 and A.5 it is verified that the total power (root mean square) of the wave is unchanged after the phase disturbance. The normalized total power of original signal is 6397.2096 and it is 6397.1740 and 6397.2042 after Fourier phase changes and analytic wavelet phase changes respectively. It verifies that the shift-invariance property is retained in case of analytic Hardy wavelet analysis like in Fourier analysis.

Let us compare the difference of the time history of the original and generated waves in Figures A.4 and A.5. In case of Figure A.4, the unwanted ripples appearing towards the latter part of time history implies that the power is distributed throughout the time history and disturbance is created throughout. On the other hand, the ripples can be hardly seen in those areas in case of wavelet function in Figure A.5. This is because the wavelet function is localized in time domain and the influence of the phase disturbance. This special feature allows us to consider the uncertainty in phase property without losing the time characteristics of the original wave. Since we want to focus mainly on dominant part of the ground motion and we do not want to reduce the severity in dominant part in the generated waves, this feature in wavelet case proves to be very useful.

A.4 Conservation of Transmitting function

The generation of artificial ground motions is followed by the computation of response of a structural system. Let us consider a linear structural system. The response is given as the convolution of impulse response function of the linear system and the input motion. It is well known that Fourier transform of the output response is given as the product of Fourier transform of impulse response function (transmitting function) and that of the

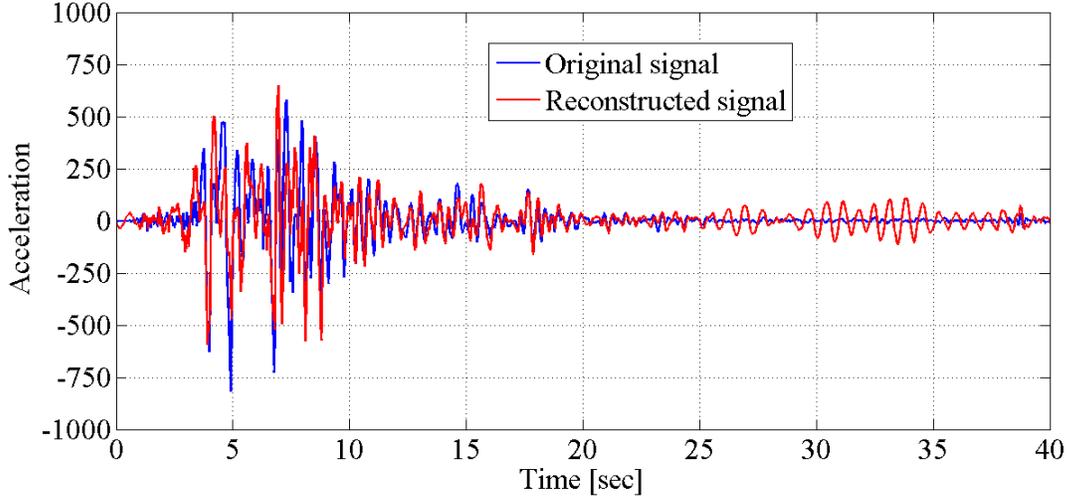


Figure A.4: Comparison of time series of an original and reconstructed signal after the Fourier phase of 2 percent of the total or 25 percent of the dominant Fourier coefficients are contaminated by noise. Reconstructed wave is accompanied by the ripple which did not exist in the original signal.

input motion as shown in equation (A.10). It allows us to evaluate the response of the structural system easily and therefore very useful for practical purpose.

In this paper, we show that similar relationship can be presented with wavelet coefficients and that the response of a linear structural system can be calculated from wavelet coefficients of input motion and the same transmitting function used in Fourier transform.

A.4.1 Calculation of Response Using Wavelet Coefficients and Transmitting Function

The analytical relation between response of a linear single degree of freedom system, the transmitting function and the input signal is determined as follows. Let us consider $x(t)$: An input signal in time domain $y(t)$: Response of linear system in time domain for the input signal $x(t)$ $h(t)$: Impulse response of the linear system in time domain $\hat{X}(\omega_k)$, $\hat{Y}(\omega_k)$ and $\hat{H}(\omega_k)$: Discrete Fourier transform of $x(t)$, $y(t)$ and $h(t)$ respectively, where $\hat{H}(\omega_k)$ is also known as a transmitting function, the $\hat{\square}$ sign representing discrete Fourier transform $\tilde{X}(\tau, s_j)$: Discrete Analytic Wavelet transform of $x(t)$, the $\tilde{\square}$ sign representing wavelet transform $\hat{\tilde{X}}(\omega_k, s_j)$, $\hat{\tilde{Y}}(\omega_k, s_j)$: Discrete Fourier transform of discrete wavelet transform of $x(t)$, $y(t)$, respectively

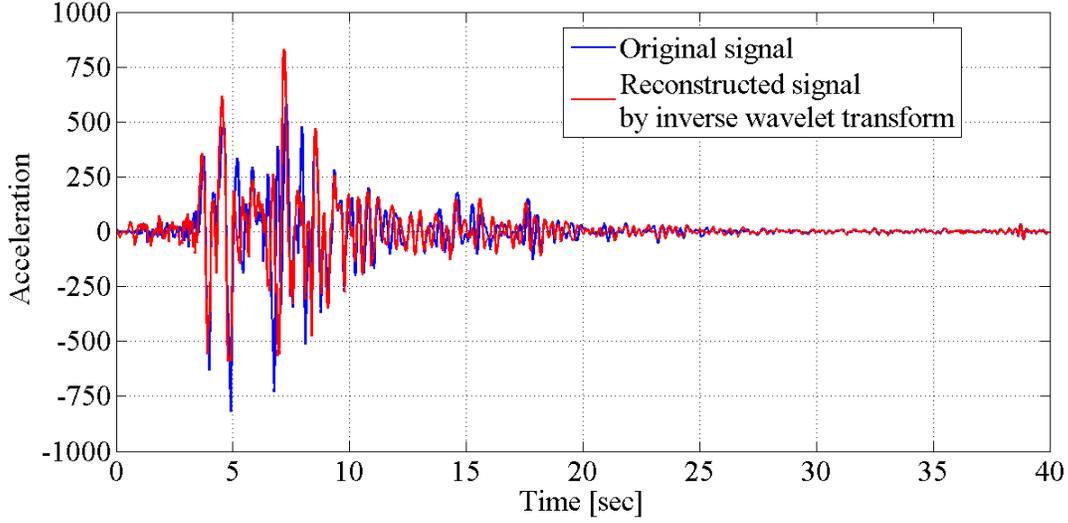


Figure A.5: Comparison of time series of the original and reconstructed signal after the phase of 2 percent of total or 20 percent of the dominant analytic Hardy wavelet coefficients are contaminated by noise. Reconstructed signal is more affected locally compared to the reconstructed signal in Figure A.4.

The discrete analytic wavelet transform of $x(t)$ is given by its convolution with analytic Hardy wavelet function and is represented as

$$\tilde{X}(\tau, s_j) = \sum_{n=0}^{N-1} x(n\Delta t) \psi_A\left(\frac{n\Delta t - \tau}{s_j}\right) \quad (\text{A.6})$$

Let us define the wavelet function corresponding to the j -th scale and n -th shift as

$$\psi_A^{j,n}(\tau) = \psi_A\left(\frac{n\Delta t - \tau}{s_j}\right) \quad (\text{A.7})$$

The wavelet coefficients are distributed in time-scale (time-frequency) domain. It should be noted that the number of coefficients ($N_j = 2^{j-1}$) is different at each scale s_j . So the discrete Fourier transform of discrete wavelet coefficients $\tilde{X}(\tau, s_j)$ in terms of τ is calculated separately for each scale and is represented as

$$\hat{X}(\omega_k, s_j) = \hat{X}(\omega_k) \hat{\psi}_A^{j,n}(\omega_k) \quad (\text{A.8})$$

Similarly, the discrete analytic wavelet transform of response $y(t)$ is represented as

$$\hat{Y}(\omega_k, s_j) = \hat{Y}(\omega_k) \hat{\psi}_A^{j,n}(\omega_k) \quad (\text{A.9})$$

Utilizing the conventional relationship among $\hat{X}(\omega_k)$, $\hat{Y}(\omega_k)$ and $\hat{H}(\omega_k)$:

$$\hat{Y}(\omega_k) = \hat{H}(\omega_k)\hat{X}(\omega_k) \quad (\text{A.10})$$

where $\omega_k = 2\pi k/N\Delta t$ denotes angular frequency, $k = 0, 1, 2, \dots, (N-1)$ is the frequency index, $j = 1, 2, 3, \dots, m$ denotes scale index, $N = 2^m = \sum_{j=1}^m N_j$ represents total number of discrete time series data, $N_j = 2^{j-1}$ represents total number of coefficients at each scale index j , and, Δt denotes time interval.

From Equations (A.9) and (A.10) we get

$$\hat{\tilde{Y}}(\omega_k, s_j) = \hat{H}(\omega_k)\hat{X}(\omega_k)\hat{\psi}_A^{j,n}(\omega_k) \quad (\text{A.11})$$

From Equations (A.8) and (A.11), it leads to

$$\hat{\tilde{Y}}(\omega_k, s_j) = \hat{H}(\omega_k)\hat{\tilde{X}}(\omega_k, s_j) \quad (\text{A.12})$$

This equation (A.12) shows that the discrete Fourier coefficients of discrete wavelet coefficients of response can be evaluated by the product of discrete Fourier coefficients of discrete wavelet coefficients of input signal and transmitting function of the structural system. During the computation, some constant terms might be introduced in this equation (A.12) depending upon the definition of Fourier transform of impulse response function, such as

$$\hat{H}(\omega_k) = \frac{1}{N} \sum_{n=0}^{N-1} h(n\Delta t)e^{-i\omega_k n\Delta t} \quad (\text{A.13})$$

From equations (A.12) and (A.13) we get

$$\hat{\tilde{Y}}(\omega_k, s_j) = N\Delta t\hat{H}(\omega)\hat{\tilde{X}}(\omega_k, s_j) \quad (\text{A.14})$$

In order to reconstruct the signal from the coefficients of response obtained in equation (A.14) two steps are followed. Firstly, the inverse Fourier transform are taken separately for each scale:

$$\tilde{Y}(\tau, s_j) = \sum_{n=0}^{N-1} \hat{\tilde{Y}}(\omega_k, s_j)e^{i\omega_k n\Delta t} \quad (\text{A.15})$$

Secondly, the inverse wavelet transform is applied to discrete wavelet coefficients obtained in equation (A.15) to obtain the response in time series:

$$y(t) = \sum_{\tau} \sum_{s_j} \tilde{Y}(\tau, s_j)\psi_A\left(\frac{n\Delta t - \tau}{s_j}\right) \quad (\text{A.16})$$

A.4.2 Numerical Simulation

The relation in equation (4.14) is verified by the numerical computation by using strong motion record data as an input signal. We consider a single-degree-of-freedom system exposed to an earthquake motion. The mass and stiffness are set so that natural time period is given as $\tau_n = 1$ second. Damping factor ($\zeta = 0.02$) is also added. The strong motion record (NS component) obtained at the Kobe observatory in Kobe during the 1995 Hyogoken Nanbu Earthquake is used as an input signal. The same input signal is used for numerical simulation in section A.3 and is shown as original signal in Figures A.4 or Figure A.5. Impulse response of the linear single degree of freedom system to the input signal is calculated in time-domain by Newmark- β method [5]. Both the input signal and impulse response are converted into analytic signals, following the definition of analytic signal in equation (A.1). Then the response of structure is obtained by following the procedure explained in section A.4.1.

As a reference, the response of the linear structural system is also obtained by using time domain Newmark- β method. Time histories of the response evaluated from wavelet transform and the one obtained by time integration are compared in Figure A.6. The two time histories show good agreement with each other, which verifies the relation obtained in equation (A.14).

Two figures here.

A.5 Conclusion (Appendix A)

The representation of uncertainties is important for dynamic analysis of structures and synthesis of design input ground motion. A number of uncertain stochastic ground motions can be generated by disturbing the Fourier phase spectrum randomly and maintaining the frequency characteristics and total power of an original ground motion. In short, Fourier transform considers the temporal change in ground motion characteristics by using the concept of 'phase'. However, the Fourier phase changes disturb the time-frequency characteristics and localized disturbances at desired time intervals are not possible. For consideration of localized time-frequency characteristics, wavelet analysis is more widely accepted. But unlike in Fourier phase, usually we cannot utilize various conventional methodologies to generate artificial uncertain ground motions considering the wavelet phase. Considering such problems, this paper proposes a scheme to use discrete

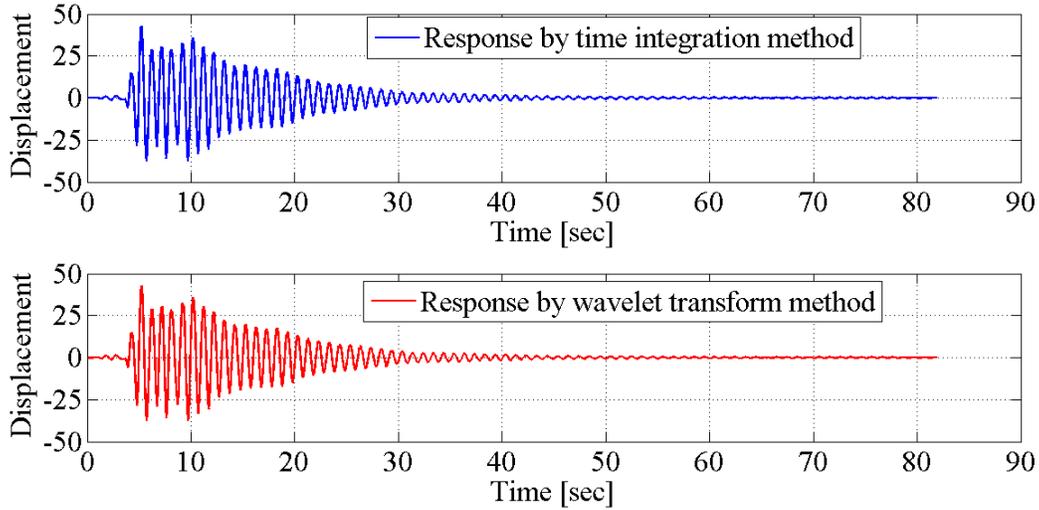


Figure A.6: Comparison of time histories of responses of a linear system calculated by the presented wavelet scheme and the time-integration scheme. They show a very good agreement.

wavelet transform using analytic Hardy wavelet function. This tool enables phase localized in time-frequency domain by retaining shift-invariance and orthogonal properties like in Fourier analysis and allows performing ground motion simulation considering uncertainties in wavelet phase. Thus the presented scheme incorporates the advantages of both Fourier and wavelet transform by expressing the signal using the phase and amplitude in the similar manner as in Fourier transform and representing localized time-frequency characteristics of a signal as in wavelet transform. Numerical simulation shows that the presented scheme is not only able to generate ground motions with noise in phase but also able to produce localized disturbance in time history of the synthesized signal. Also it is verified with the numerical simulation that the dynamic response of a linear structural system can be evaluated by wavelet transform method using the same transmitting function defined by Fourier transform.

Appendix B

Computation Results (Contd.)

The computation results that were not included in chapter 7 are presented in this appendix B.

B.1 Low Pass Filter

The FIR LPF (200 sec cut-off time) of Taurus-Z, NS and EW component are as follows.

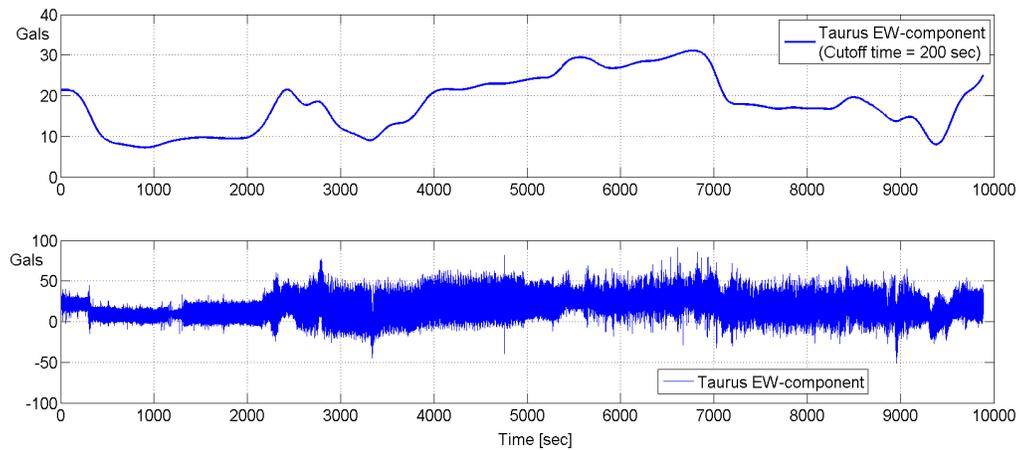


Figure B.1: a) Top: FIR LPF Taurus EW data with cut-off period 200 sec and b) Bottom: Accelerometer Titan (Taurus EW-component).

The FIR LPF (cut-off period = 100 sec) of digital servo is shown below.

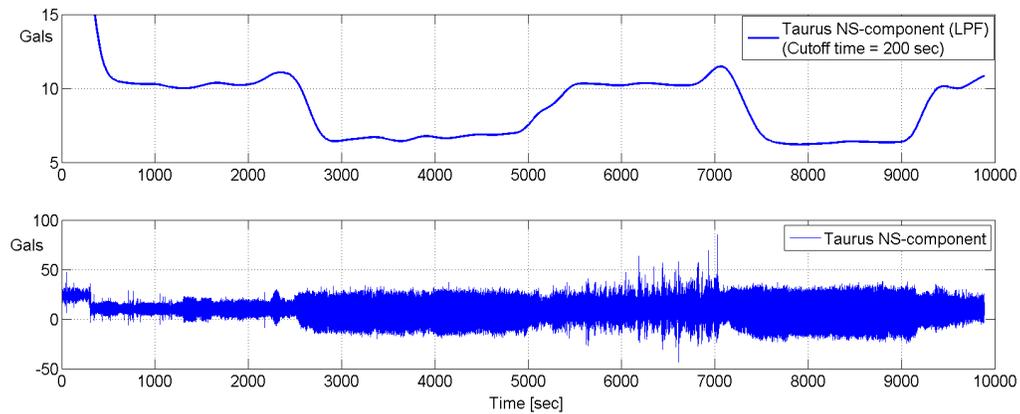


Figure B.2: a) Top: FIR LPF Taurus NS data with cut-off period 200 sec and b) Bottom: Accelerometer Titan (Taurus NS-component).

B.2 Continuous Wavelet Transform

The input to ICA: CWT coefficients (real part) corresponding to first 43 scales ($j0.005$ Hz) with coefficients reduced to zero at ship-stoppage time, ICA output and the de-mixing matrix are shown in figure below.

B.3 Low pass filter and SOBI

The results for four sets of data applied to SOBI are presented below.

B.4 Low pass filter and ICA

B.5 Comparison of Results

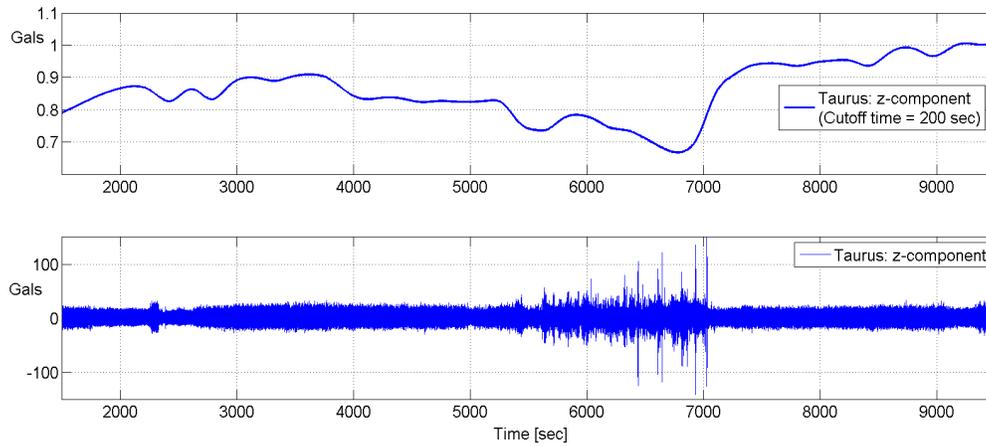


Figure B.3: a) Top: FIR LPF Taurus-Z with cut-off period 200 sec b) Bottom: Accelerometer Titan (Taurus-Z component).

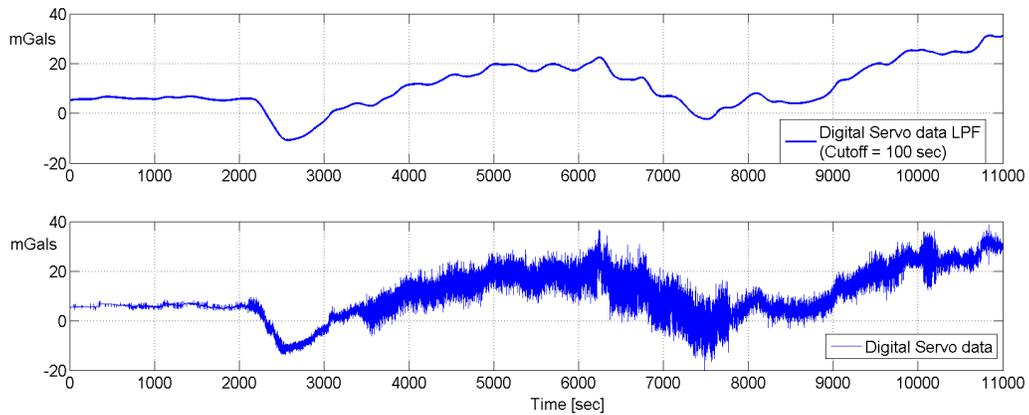


Figure B.4: a) Lower: Digital Servo data and b) Upper: FIR LPF digital servo data with cut-off period 100 sec.

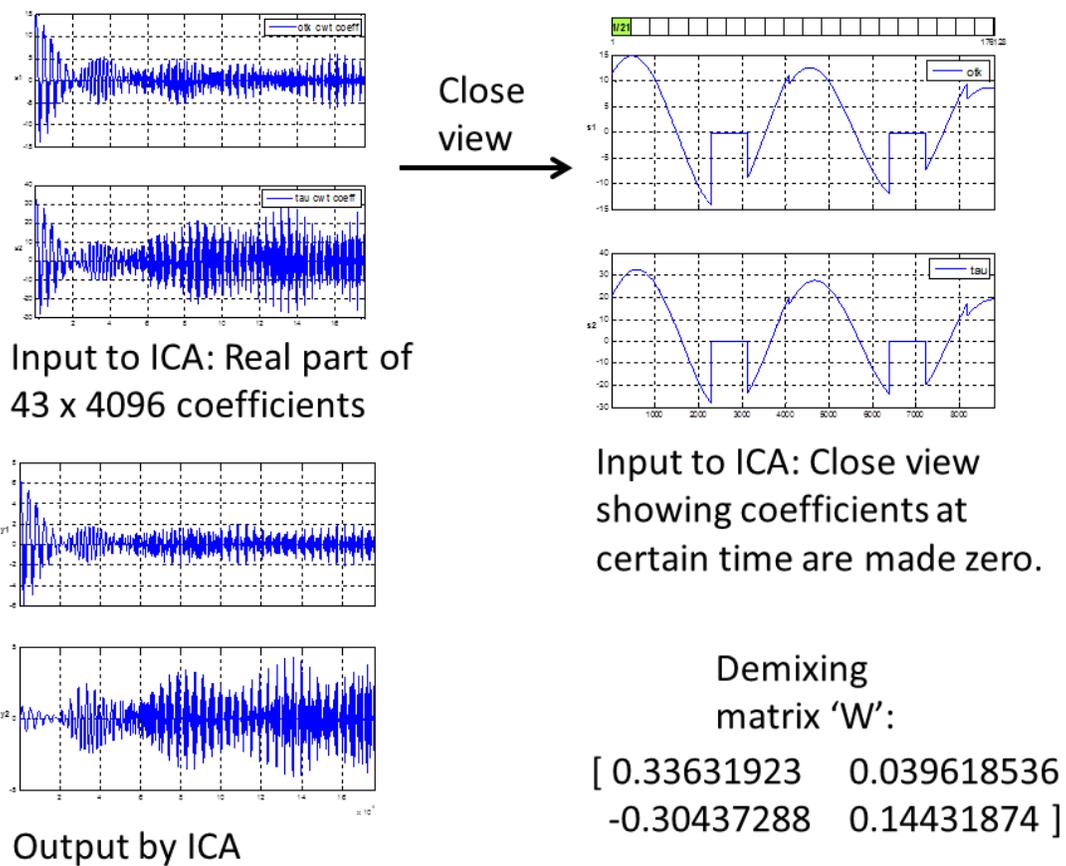


Figure B.5: Input CWT coefficients (upper left), close view of input (upper right), Output by ICA (lower left) and the de-mixing matrix (lower right).

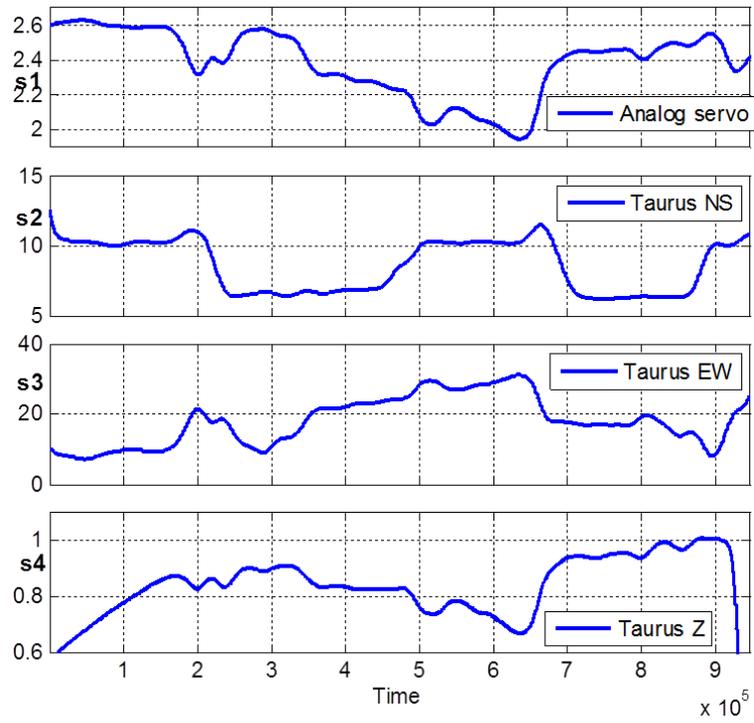


Figure B.6: Input for SOBI after filtering 200sec lpgf.

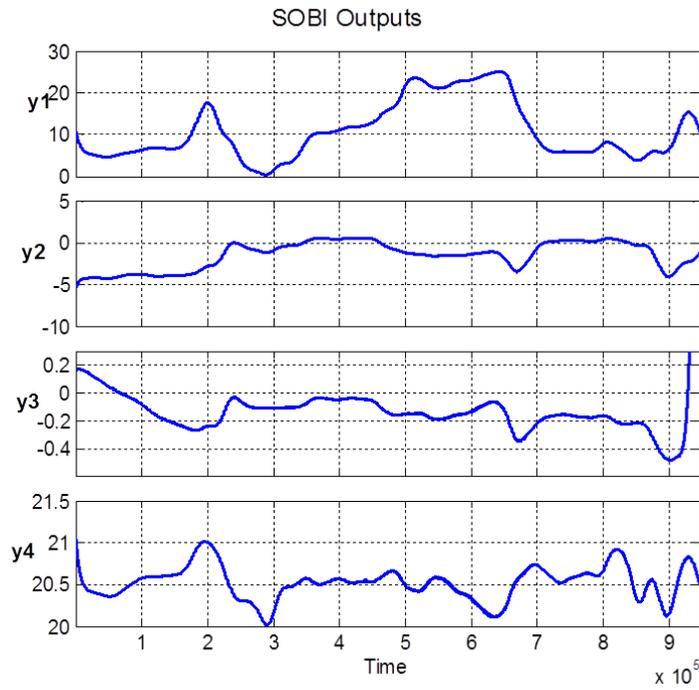


Figure B.7: SOBI outputs after filtering 200sec lpgf.

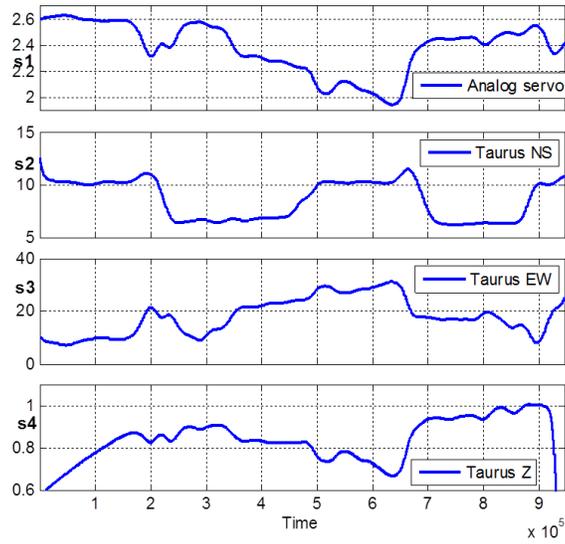


Figure B.8: Input for ICA after applying 200sec lpgf.

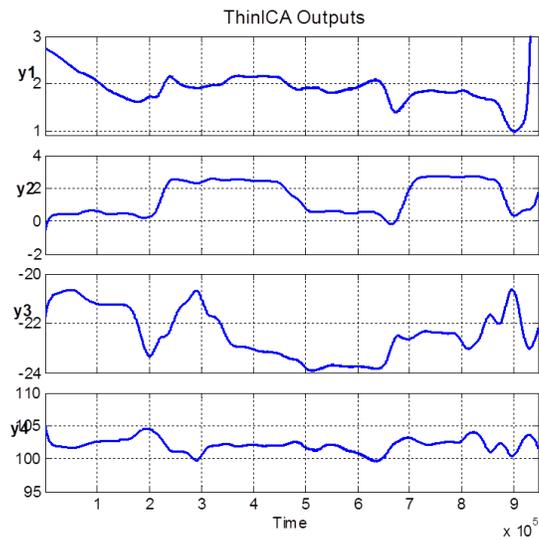


Figure B.9: Output by ThinICA (for four inputs (200 sec lpf): Neither component is very good.

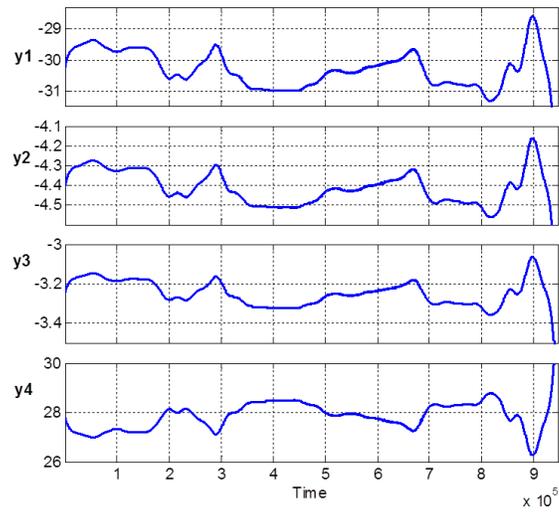


Figure B.10: Output by Power ICA(for four inputs (200 sec lpf): All outputs are same: comparison shown later: Not very good.

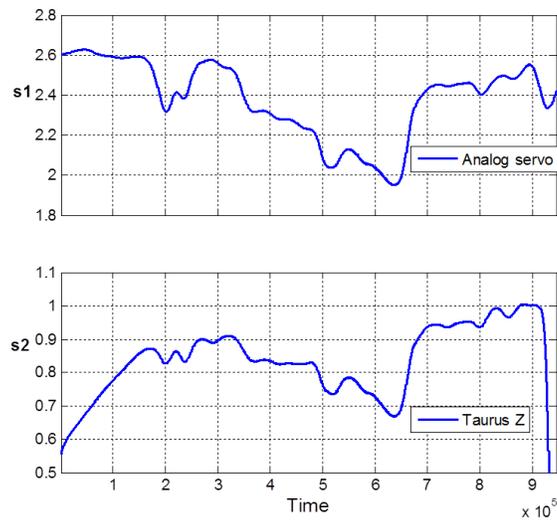


Figure B.11: Input to ThinICA vse and tauZ data (lpf 200sec): Time 13:00 till end.

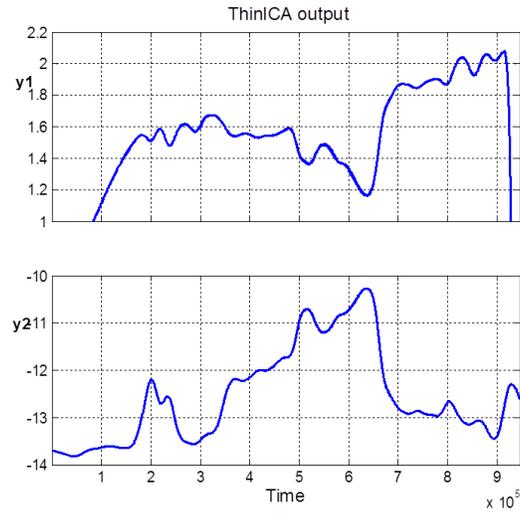


Figure B.12: Output by ThinICA for vse and tauZ data (lpf 200sec): 13:00 till end: Performance is not good.

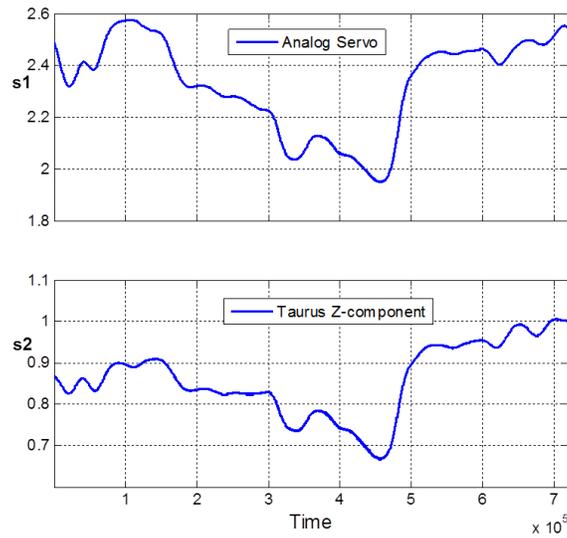


Figure B.13: Input to ThinICA(VSE and Taurus Z) LPF(200 sec). Time: 13:30:00 to 15:30:00 truncated.

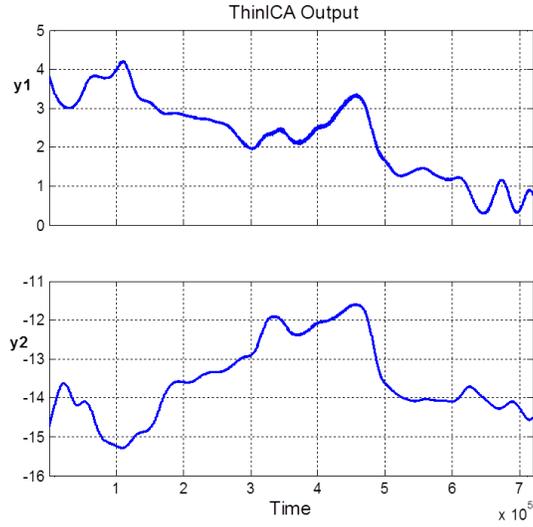


Figure B.14: Output by ThinICA for 13:30:00 to 15:30:00 truncated lpf(200 sec) data (input vse and tauZ: One output (top) is good (same output for 100 sec cutoff is better than this).

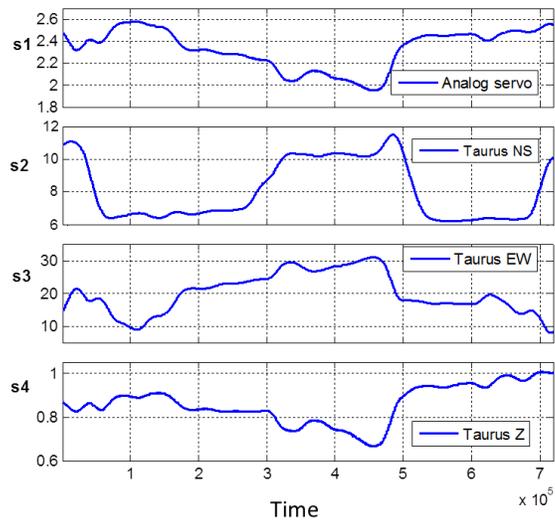


Figure B.15: Input for ICA (200sec lpf) and truncated four sets of data 13:30 to 15:30).

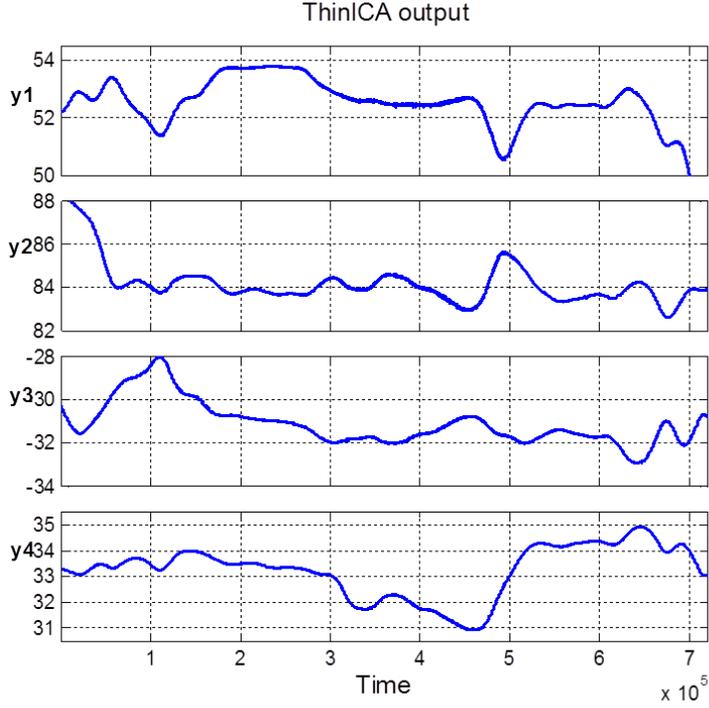


Figure B.16: Output by ICA for input (200sec lpf) and truncated four data 13:30 to 15:30: Third output is good.

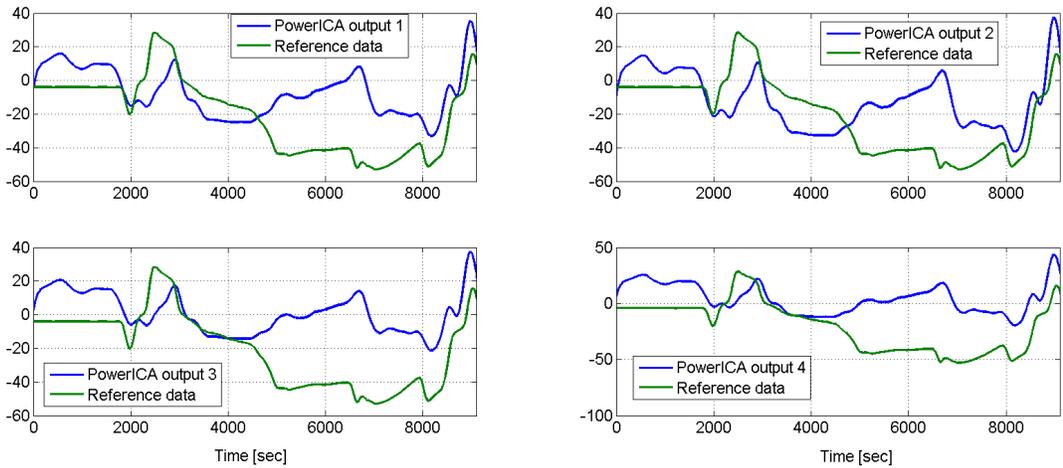


Figure B.17: Four Outputs by Power ICA (input: four 200sec lpf data) compared with reference data.

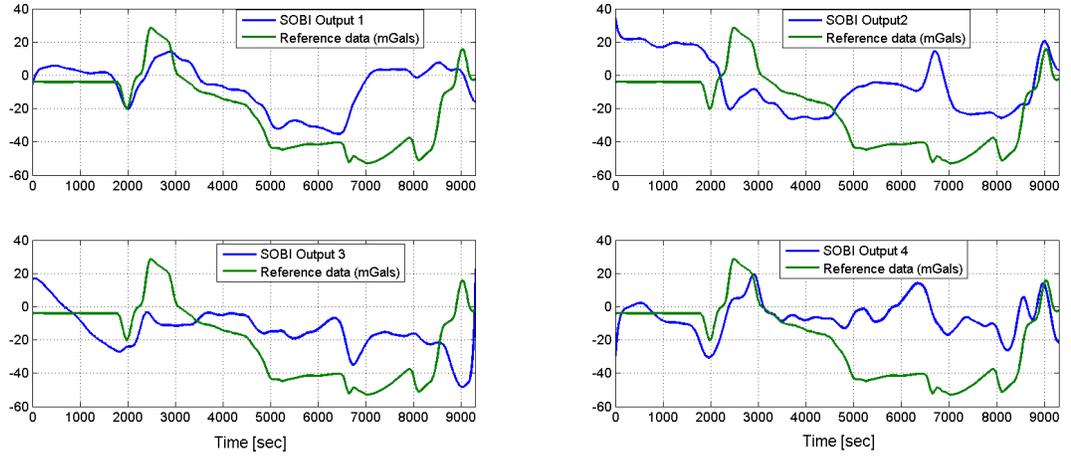


Figure B.18: Four Outputs by SOBI (input 200sec lpf) compared with reference data 13:00:00: Fourth output follows some trend of reference but not very good.

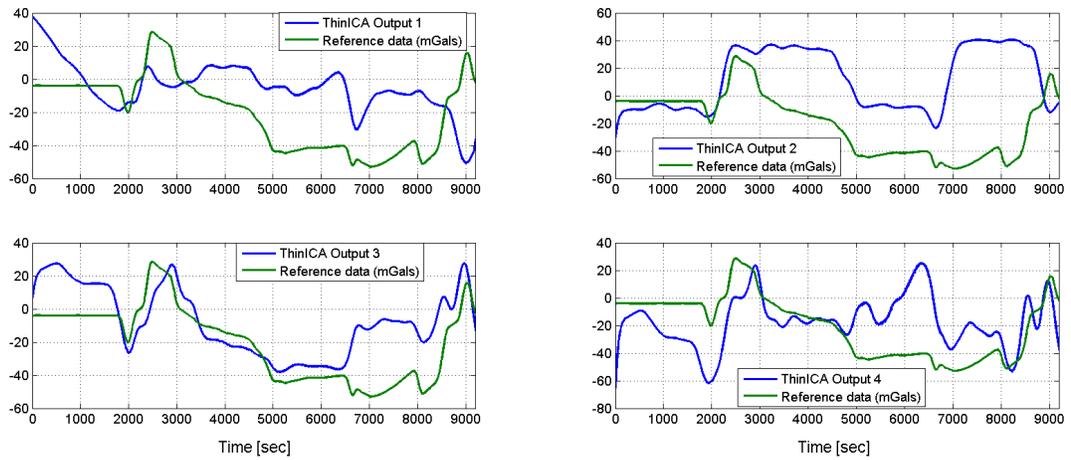


Figure B.19: Four Outputs by Thin ICA (input 200sec lpgf) compared with reference data 1300 onwards: Fourth output is good.

Bibliography

- [1] J.F.Cardoso A.Belouchrani, K.Abed-Meraim and E.Moulines. A blind source separation technique using second-order statistics. *IEEE Transactions on Signal Processing*, 45:434–444, 1997.
- [2] A.Cichocki and S.Amari. *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. Wiley, USA, first edition, 2003.
- [3] K.Siwiek T.Tanaka A.Cichocki, S.Amari and A.H.Phan. Icalab toolboxes version 3, 2007.
- [4] J.Karhunen A.Hyvarinen and E.Oja. *Independent Component Analysis*. John Wiley and Sons Inc, USA, first edition, 2001.
- [5] A.K.Chopra. *Dynamics of Structures: Theory and Applications to Earthquake Engineering*. Prentice-Hall, Inc., USA, third edition, 2007.
- [6] A.Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, USA, third edition, 1991.
- [7] M. Komazawa H. Sekiguchi J. Akamatsu C. Takahashi, H. Morikawa and S.Sawada. Modeling a 3-d subsurface structure around damaged area due to the recent earthquakes in niigata, japan using gravity survey. In *The 14th World Conference on Earthquake Engineering*, 2008.
- [8] M. Komazawa H. Sekiguchi J. Akamatsu C. Takahashi, H. Morikawa and S.Sawada. Modeling a 3-d subsurface structure around damaged area due to the recent earthquakes in niigata, japan using gravity survey. In *The 14th World Conference on Earthquake Engineering*, 2008.

- [9] C.Torrence and G.P.Compo. A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society.*, 79:61–78, 1998.
- [10] D.M.Boore. Simulation of ground motion using the stochastic method. *Pure and Applied Geophysics*, 160:635–676, 2003.
- [11] D.Weinstein. *Method and apparatus for seismic geophysical exploration*. 1976.
- [12] J.Broome G.M.Ross and W.Miles. Potential fields and basement structure: Western canada sedimentary basin, in g. d. mossop and i. shetsen, comps. *Geological atlas of the Western Canada Sedimentary Basin*, pages 41–46, 1994.
- [13] H.Morikawa H.Kiku H.Goto, S.Sawada and H.Ozalaybey. Modeling of 3d subsurface structure and numerical simulation of strong ground motion in the adazapari basin during 1999 kocaeli earthquake, turkey. *The Bulletin of Seismological Society of America*, 95:6:2197–2215, 2005.
- [14] H.Okada. *Micrometer Survey Method Monograph series*. The Society of exploration geophysicists, USA, geophysical monograph series no. 12 edition, 2003.
- [15] L.Cohen. *Time-Frequency Analysis*. Prentice-Hall, Inc., USA, first edition, 1995.
- [16] L.L.Nettleton. *Elementary Gravity and Magnetism for Geologists and Seismologists*. The Society of exploration geophysicists, Oklahoma, first edition, 1971.
- [17] R.Nishida I.Ohata T.Yamashita M. Adachi, T.Noguchi and K.Omura. Determination of subsurface structure of izumo plain, southwest japan using microtremors and gravity anomalies. In *The 14th World Conference on Earthquake Engineering*, 2008.
- [18] M.C.Jones and R.Sibson. What is projection pursuit? *Journal of the Royal Statistical Society, Ser. A*, 150:1–36, 1987.
- [19] M.Farge. Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech.*, 24:395–457, 1992.
- [20] M.Mirzaoglu and U.Dykmen. Application of microtremors to seismic microzoning procedure. *Journal of the Balkan Geophysical Society*, 6:143–156, 2003.

- [21] V.J.S.Grauch R.O.Hansen T.R.LaFEhr Y.Li W.C.Pearson J.W.Peirce J.D.Phillips M.N.Nabighian, M.E.Ander and M.E.Ruder. Historical development of the gravity method in exploration. *Geophysics*, 70:63ND–89ND, 2005.
- [22] M.N.Toksoz and R.T.Lacoss. Microseisms: Mode structure and sources. *Science*, 159:872–873, 1968.
- [23] P.J.Huber. Projection pursuit. *The Annals of Statistics*, 13(2):4351–475, 1985.
- [24] R.G.Lyons. *Understanding Digital Singal Processing*. Pearson Education, Inc, India, second edition, 2004.
- [25] R.Honda and P.P.Khatri. Discrete analytic signal wavelet decomposition for phase localized in time-frequency domain for generation of stochastic signal with phase uncertainty. In *Proceedings of the 15th World Conference on Earthquake Engineering (15 WCEE)*, 2012.
- [26] R.Honda and P.P.Khatri. Discrete analytic signal wavelet decomposition for phase localized in time-frequency domain for simulation of ground motions with stochastic property in phase spectrum. In *Proceedings of the Joint 9th International Conference on Urban Earthquake Engineering (9CUEE) and 4th Asia Conference on Earthquake Engineering(4ACEE)*, 2012.
- [27] S.Cruces and A.Cichocki. Combining blind source extraction with joint approximate diagonalization: Thin algorithms for ica. 2004.
- [28] S.Qian and D.Chen. *Joint Time-Frequency Analysis*. Prentice-Hall Inc., USA, first edition, 1996.
- [29] T.M.Cover and J.A.Thomas. *Elements of Information Theory*. Wiley, USA, first edition, 1991.
- [30] W.C.Campbell. *Introduction to geomagnetic fields*. Cambridge University Press, London, geophysical monograph series no. 12 edition, 1997.