Development of Algorithms for Near-shore Satellite Derived Bathymetry Using Multispectral Remote Sensing Images

(マルチスペクトル衛星画像を用いた沿岸域の水深推定のためのアルゴリズムの開発)

2017年3月

大阪市立大学大学院創造都市研究科

創造都市専攻都市情報環境研究領域

D14UD511

POLIYAPRAM Vinayaraj (ポリヤポラム　ヴィネラズ)

Table of Contents

List of Figures

Chapter Two

Chapter Three

Chapter Four

Chapter Five

List of Tables

Chapter Two

Chapter Three

Chapter Four

Chapter Five

Acknowledgements

This is time to recall many people and experiences during my PhD in Osaka City University where I lived nearly four years. First of all, I would like to express my profound and heartfelt thanks to my supervisor Prof. Venkatesh Raghavan for his constant support and motivation. I would like to express my special thanks to Prof. Venkatesh Raghavan for introducing me in the challenging field of Open Source GIS. I could not have imagined having a better advisor for my doctoral course. Therefore, he is my true mentor. I would also greatly appreciate Prof. Shinji Masumoto's comments and suggestions always amazed me and it helped me to improve my research. Prof. Shinji Masumoto was always available to shape the research in the present form.

I express my sincere thank to MEXT and Osaka City University for supporting this research work. I am also grateful to National Oceanic and Atmospheric Administration (NOAA), USA for providing high resolution LiDAR derived water depth data from Puerto Rico. National Institute of Oceanography (NIO), Goa, India is acknowledged for supporting SoNAR depth data. I express my sincere gratitude to Professor Akio Okayasu, Dept. of Ocean Sciences, Tokyo Univ. of Marine Science and Technology for providing high resolution SoNAR from parts of Tohoku region.

I would like to thanks Dr. Luca Delucchi for his great support and encouragement to develop a module in GRASS GIS. I also greatly appreciate love and encouragement of Mr. Hirofumi Hayashi and constantly helped in introducing Open Data and Open Sources GIS. I also sincerely acknowledge Dr. Markus Metz for his support and contribution in developing GRASS GIS module. I specially want to express my gratitude to Dr. Go Yonezawa and Dr. Daisuke Yoshida, for their great support in making research documents especially in Japanese during my tenure as doctoral student. I also express my sincere thankfulness to Mr. Wenlong Yu for continuous help and support in dealing with Japanese official procedure which made my life easier in Japan. I greatly appreciate the travel grant provided Department of Mathematics, Osaka City University, to attend the overseas conference. Last but not least, my sincere thanks to my family and my lab mates who helped me and supported me during my tenure as a doctoral student.

Acronyms

ANUGA	Australian National University and Geoscience Australia
DBDB	Digital Bathymetric Data Base
DEM	Digital Elevation Model
GEBCO	General Bathymetry Chart of Oceans
GIS	Geographical Information System
GRASS	Geographic Resources Analysis Support System
ICRM	Integrated Coastal Relief Model
LiDAR	Light Detection and Ranging
MLR	Multiple Linear Regression
MSI	Multispectral Instrument
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near-infrared
RBE	Radiance Based Estimation
RMSE	Root Mean Square Error
SDB	Satellite Derived Bathymetry
SLR	Simple Linear Regression
SoNAR	Sound Navigation and Ranging
SRTM	Synthetic Aperture Radar Topographic Mission
SWIR	Short Wave Infrared
TTJS	Tohoku earthquake Tsunami Joint Survey
UAV	Unmanned Aerial Vehicle
WBA	Wave-Based Approach

Abstract

There is a growing need for up-to-date and reliable data for near-shore regions to afford a better understanding of coastal processes and aid in better management and protection of coastal areas. Near-shore bathymetry is a basic and very essential requirement in coastal engineering and coastal management. Recently, mapping of near-shore bathymetry using boat based SoNAR and air-borne LiDAR have become available. These technologies can provide high resolution and accurate bathymetric measurements in shallow waters, but availability of these technologies are currently limited due to significant costs and time involved in data acquisition. Therefore, this research investigates alternative methods for near-shore Satellite Derived Bathymetry (SDB) using readily available multispectral satellite images.

As an alternative to field based approaches, a number of remote sensing methods for estimation of bathymetry have been proposed and evaluated. Among them, single spectral band based algorithms assume constant light attenuation coefficient and homogeneous bottom type. In the present research, this method was evaluated in two study areas (Ratnagiri, India and Taketomi, Japan) with three kinds of optical satellite imageries, namely Landsat-8 and ASTER. The results indicate that SDB using single spectral band are reliable for clear waters when bottom type are homogeneous. Further, single band SDB requires a good amount of field data for calibration and validation. As an alternative to single band SDB, a new Radiance Based Estimation (RBE) algorithm using single spectral band has been proposed. SDB models using single band are found to be affected by turbidity and do not adequately address issues related to heterogeneities in bottom types and water conditions.

Several researchers have proposed SDB using multispectral images in order to address heterogeneous conditions in highly dynamic and complex coastal water conditions. In general, these algorithms attempt to isolate attenuation components and estimate SDB using a combination of spectral bands. This study proposes a new spectral band combination using Near-Infrared (NIR) band as it provides some information about bottom reflection. Further, a new correction method using Short Wave Infrared (SWIR) was proposed in lieu of NIR used in previous researches. Results of using NIR in band combination provide significant improvement in SDB. Further, using SWIR provides

improved results for atmospheric and water corrections. SDB estimation using multiple linear regression does not sufficiently account for local heterogeneities. To address this issue, this research recommends the use of local coefficients rather than a single global β-coefficient. As opposed to non-weighted global models, GWR computes β-coefficients for individual pixel. The β-coefficients are determined using moving kernel and pixels close to the centroid are assigned higher weights. Two kinds of kernel settings were investigated. In Fixed-GWR, the size of the kernel remains constant and in Adaptive GWR (A-GWR) the kernel size is set to vary as function of density of calibration points. Appropriate weighting functions (e.g. bi-square and Gaussian) and optimal bandwidth (kernel radius) were determined by a double cross-validation carried out before and after estimation. Both kind of GWR model were evaluated for a test site in Puerto Rico. The results indicate that A-GWR consistently provide the best results but is computationally intensive.

Subsequently, SDB off the coast of Miyagi Prefectures, Japan were estimated using GWR models. The SDB was combined with global bathymetry (GEBCO) and terrestrial DEM (SRTM) to generate an Integrated Coastal Relief Model (ICRM). The ICRM was used as input for ANUGA hydrodynamic tsunami simulation model in order to hindcast the past Great Tohoku tsunami event of 11 March, 2011. The tsunami simulation results showing good correspondence with post-tsunami survey data in terms of extent of inundation as well as tsunami inundation height recorded and demonstrate the effectiveness of SDB.

The workflow of GWR based SDB model has been implemented using a Free and Open Source GIS framework and therefore can be easily applied in other areas without the need to invest resources for software. A new Python module, *i.image.bathymetry*, was implemented by integrating functions of R Statistical Computing package and GRASS GIS. The performance of the new module was evaluated using data for different test sites (e.g. Puerto Rico, Iwate and Miyagi) and remote sensing data (Landsat-8, RapidEye and Sentinel-2). The results obtained were demonstrated the efficacy of the module for SDB from multi-constellation satellite data. SDB estimation from microwave satellite image, multispectral UAV images and deployment of SDB workflow as Web Processing Service are proposed as future topics of research.

Chapter One
Introduction

A coastal zone is the interface between the land and water. These zones are important because around 60% of the world's population inhabit such zones. Therefore, it is essential to have comprehensive knowledge about the coastal process which can directly impact the socio-economic conditions. Incessant monitoring and data collection are essential to have a complete understanding of the coastal zone. Recent development in the field of science and technology offers potential to collect data from coastal zone. In spite of having time and cost effective technologies to collect data from the land portion of the coast, there is a few time and cost effective technologies available for data collection from coastal ocean so far. Moreover, the complexity and dynamicity of the coastal ocean make it more difficult to collect the data. Lack of coastal ocean data sets impacts significantly the research related to coastal processing and modeling.

Among various data sets, near-shore bathymetry is a key parameter which can be largely used as vital information for many coastal researches such as coastal erosion, regional sea-level rise, subsidence, hydrodynamic modeling, tsunami simulation, etc. The major ship based surveys have been carrying out to measure bathymetry over off-shore not over near-shore simply due to the inaccessibility of shallow water depth in the near-shore region. In spite of having bathymetry data sets such as DBDB5, ETOPO and GEBCO, a lacuna in the global bathymetry data sets has exists for near-shore region due to the low resolution of available bathymetry. Moreover, such low-resolution bathymetry data sets cannot be effectively used to support detailed coastal research. Therefore, the coastal researchers are compelled to survey the near-shore bathymetry either by small boats using echo-sounders or by aerial vehicles using LiDAR technology.

Even though, these technologies are highly expensive and time consuming, many researchers have been surveyed near-shore bathymetry in small areas with local scale using SoNAR and LiDAR (Costa *et al*., 2009). LiDAR can provide accurate bathymetric measurements in shallow areas, but availability of this technology is currently limited.

Apart from that, the highly dynamic nature of near-shore region leads to frequent change in bathymetry that are required to be monitored at periodic intervals, and hence the survey should be carried out repetitively. This is almost not practical to carry out the survey frequently. A suitable alternative for near-shore bathymetry estimation method is essential for complex and dynamic near-shore region. Therefore, this doctoral work aims to develop an alternative method to estimate near-shore bathymetry using multispectral satellite images generally termed as SDB. The alternative solution discussed in this thesis is a feasible, cost effective, time effective, synchronous and flexible near-shore bathymetry estimation model which can be applied in any near-shore region around the world.

1.1 Research background

The research related to SDB generation over near-shore region has history of at least two decades. Such methods have been proposed for SDB using optical remote sensing which emphasizes the principle that a portion of reflectance from near-shore region is a function of water depth. It takes advantage of shortwave radiation in the electro-magnetic spectrum that has a strong penetration capacity. As the incident solar radiation propagates through the water, it undergoes in-water scattering, absorbed by water and in-water constituents, leaving varied energy to be backscattered and recorded in remote sensing imagery. The following section sets out various methods used to estimate SDB over near-shore region. Mainly these methods can be classified as analytical models, empirical models and combined models.

1.1.1 Analytical model

Analytical modeling of near-shore bathymetry is purely based on the manner of light penetration in water. Development of this model requires a number of optical properties of water over near-shore region such as absorption coefficient suspended and dissolved materials, attenuation coefficient, scattering coefficient, backscatter coefficient, bottom reflectance, etc., (Lee, *et al.*, 1998 and Mobly, *et al.*, 2001). These kind of models generally termed as a radiative transfer model. The radiative transfer model involves the inherent assumption of a reflective bottom, an appropriate level of water quality, and/or

shallow depth (e.g. Spitzer and Dirks, 1986, Baban, 1993 Muslim and Foody, 2008). Therefore, it is not suggested to apply coastal waters that have a poor reflectance of bottom due to high turbidity.

1.1.2 Empirical model

In empirical modeling, the relationship between remotely sensed radiance of a water body and the depth at sampled locations is established empirically without regard to how light propagates in water. The correlation between the water depth and the radiance of a spectral band are used for SDB estimation. The empirical modeling rely on the assumption that total water reflectance is related primarily to water depth, and secondarily to water turbidity (Lafon *et al*., 2002). Among the models, empirical parametric regression based models are quiet popular and easy to apply such as Stumpf *et al.* (2003), Gholamalifard *et al*. (2013), etc. The other common method is nonparametric regression as explained by Ribeiro *et al*. (2008) However, this empirical model provides fast data processing, but it requires calibration depth points for estimating the coefficients.

1.1.3 Combined model

As the name denotes, it is a combination of analytical and empirical model. The combination suggested by many authors in order to overcome the demerits of both models. Among SDB models, combined model is popular and has been widely applied (Lyzenga *et al.*, 2006, Kanno and Tanaka, 2012, Su *et al*., 2014 and Vinayaraj *et al*., 2016) These models are physically based algorithm and the predictor can be analytically derived from a radiative transfer model. Calibration depth is used to calculate the attenuation coefficient of each spectral band. The empirically derived parameters/coefficients by a comparison of spectral radiance with measured depths are related to the inherent optical properties of the water and the bottom. Calibration depth is needed for such models and can estimate reliable SDB even from low quality water regions. These models are faster and needed less prior knowledge of spectral properties of the water and rapid processing can be expected. The method presented in this research is based on combined model and will shed light on a number of issues in the near-shore

bathymetry estimation procedure and propose improvements.

1.2 Research aim and scope

The overall aim or objective of the study is to develop a robust SDB model and discuss about the performance of the model in various near-shore waters around the world. Moreover, implement the model on an Open Source GIS platform for assisting the coastal researchers to generate reliable SDB from their study areas.

To accomplish the above mentioned main objective, study investigates the past researches related to SDB estimation, its implementation problems and limitation in addressing various issues. The first objective of the study was to estimate SDB using single band in order to derive a simple solution with minimum requirements of spectral properties of the near-shore water. This attempt was to investigate a combined model, and hence, the spectral properties such as attenuation coefficient and scattering coefficient can be analytically derived (Stoffle and Halmo, 1991). The second important step carried out in this research was to estimate multi-temporal bathymetry even without calibration depth. Multi-temporal bathymetry over near-shore region with minimum effort was a great desire of the coastal researchers. The developed method (RBE) was an empirical model which is not based on the manner of light propagation in water. Nonetheless, only related to the radiance of the spectral band and not considering any of the spectral properties of water.

In some previous attempts in the use of a single band the rich information in other multispectral bands was not considered. Therefore, such methods are applicable in the coastal water with uniform spectral reflectance. In complex coastal water, such uniformity cannot be expected all the time. Therefore, as second objective a multispectral band based combined model was developed which was expected to address the limitation of single band SDB models (Lyzenga *et al.*, 2006). The model constructs a function that estimate the bathymetry on the basis of the observed radiances of multiple bands. Specifically, the predictor is a linear function of image-derived variables for each visible band.

In theory, the multispectral SDB model can account for a certain level of the

heterogeneity. Nonetheless, our practical investigation shows that model suffers from spatial heterogeneity due to bottom types and water quality (Vinayaraj *et al.*, 2016). Therefore the third objective was to develop a full-fledged GWR based model and performance evaluation in order to explore the ability to address spatial heterogeneity. Further, the GWR based model was applied in several study areas with various data sets to evaluate the performance in various conditions of water quality, satellite image characteristics, distribution and quantity of calibration depth points.

Accuracy assessment of GWR based SDB was carried out with reference depth and results show significant improvements. Further, this study evaluates the applicability of the SDB in practical scenarios such as in ICRM and tsunami simulation as its fourth objective. ICRM generation and tsunami simulation were carried out using ANUGA[1] hydrodynamic model. Tsunami simulation results were evaluated with post-tsunami survey results and it shows close agreement with post-tsunami survey results. After the successful evaluation using application examples, the final objective was to implement the developed GWR based SDB model in an Open Source GIS platform. GWR based SDB model has implemented in GRASS GIS[2] as a module using python script. Further the performance of the module also evaluated in terms of memory usage, speed of the processing and accuracy of the results using three case studies.

1.3 Thesis outline

This Ph.D. dissertation is organized as follows:

- Chapter One, the present chapter, provides an introduction to the theoretical background of the SDB models, scope of the research, objective of the study and thesis outline.

- Chapter Two discusses single band based SDB models which are commonly termed as SLR model, its implementation and the theory behind it. A new radiance based empirical model (RBE) proposed to estimate multi-temporal bathymetry from satellite images. Further, this study investigates multispectral based SDB models and

[1] https://anuga.anu.edu.au/
[2] https://grass.osgeo.org/

a comparison with single band based model has been demonstrated. Apart from that, a new band combination is proposed and explained for improved SDB estimation. Lastly, a new correction method using SWIR band is introduced and effectiveness of the correction methods is demonstrated.

- Chapter Three proposes a new GWR based SDB model for improved bathymetry estimation. Detailed evaluation of GWR based SDB has been carried out in order to evaluate the performance using different weighting functions. Study elucidates effectiveness of the model in addressing heterogeneity as compared to conventional multispectral model (global model). Comparison has also been carried out between A-GWR model and Fixed-GWR model.

- Chapter Four demonstrates the application example of estimated SDB using GWR model in ICRM and tsunami simulation. Field investigations on Tsunami propogation after Great Tohoku earthquake 2011 was used to compare the result with tsunami simulation results.

- Chapter Five describes a new Open Source GIS module implemented for estimating SDB using GWR model. Here, we explain the functions and workflow of the module in detail. The performance of the module assessed using a benchmark with specific computer configuration. Speed of the processing and accuracy of the results are compared with different functions of the module.

- Chapter Six evaluates the results of each chapter, discusses the demerits of previous models illustrates the efficacy of GWR based SDB model in addressing heterogeneity. Further study discusses about application example and developed module and concludes this thesis with some future perspectives.

Chapter Two
Single and Multispectral Band SDB

2.1 Introduction

Model A number of passive remote sensing methods have been proposed by many authors (Brown *et al*., 1971, Lyzenga, 1978, Benny and Dawson, 1983, Philpot, 1989, Tripathi *et al*., 2002). Several workers (e.g. Stoffle and Halmo, 1991, Stumpf *et al*., 2003, Lyzenga *et al*., 2006 and Pacheco *et al*., 2015) have successfully demonstrated the use of satellite data for determination of depth in coastal waters. Multiple band techniques have been commonly used for SDB from shallow water (e.g. Fonstad and Marcus, 2005 and Ceyhun, Yalcyn 2010). There are conflicting opinions expressed over the suitability of spectral wavelength bands for SDB estimation. Kumar *et al*., (1997) used 0.77-0.80µm wavelength for depth measurements in an estuary. Warne, (1972) adopted 0.5-0.6µm range and Yi and Li (1988) also adopted 0.47-0.54µm band for estimating SDB. George (1997) found 0.746–0.759µm range to be more suitable compared to wavelength bands. Manessa *et al.* (2014), also observed that NIR band (0.77-0.89µm) is close to the visible spectrum and still sensitive to bottom reflectance. Ibrahim *et al.* (1990), used 0.5-0.6µm (band 4) of Landsat-3 MSS to estimate SDB by correlating the intensity of pixels and depth for Penang Island in Malaysia. The shallowest areas (approximately less than 5m) show bottom reflectance in 0.77-0.90µm range but deeper areas (greater than 15m) show bottom reflectance only in 0.45-0.52µm range (Jupp *et al*., 1988). Although several algorithms have already been discussed, main purpose of this chapter is to evaluate the efficacy of some of the algorithms to recently available high radiometric and spatial resolution remote sensing data. Further, modification and improvements have proposed for these methods to obtain better results. Algorithms dealing with SDB are mainly categorized in to single and multispectral bands in terms of number of spectral band are used. This study aims at evaluating both single band and multispectral bands approaches have proposed by Clark *et al.* (1987 and 1988) and Stoffle and Halmo, (1991). This Chapter also proposes the new RBE method to estimate depth from single band without

the need of calibration depth and also proposes improved depth estimation using Clark method with new band combination. In addition to that, a refined method for atmospheric, water surface and water column correction using SWIR was proposed. Further, these proposed methods have been validated various kinds of radiometric, spatial and temporal resolution data sets such as Landat-7, Landsat-8 and ASTER.

2.2 Study area

The study area shown in Figure 2.1 is located off Ratnagiri along the Arabian Sea coast with maximum depth of 11 m. Geographically it stretches 73º16'30"E -73º18'00"E 16º57'00"N - 16º59'30"N with 5km long coastal zone, that includes 70% of sandy beach and an estuary of Kajali River. Tides in the region are mixed and are predominately semi-diurnal and the annual rainfall of the region is around 3,000mm. The near-shore depth is highly dynamic due to sediment discharge from the Kajali River and seasonal climate changes. The water has been observed as turbid due to sediment discharge from the Kajali River.

The next study area is located at off Taketomi Island, which lies south-eastern part of Japan (Figure 2.2). Geographically it stretches from 124º3'00"E - 124º5'42"E and 24º18'54"N - 24º20'42"N. Taketomi Island belongs to the Yaeyama Island group which is one of the world's largest healthy coral reef systems. Therefore, water quality of the Taketomi Island is relatively better.

2.3 Data used

Multi-constellation and multi-resolution satellite image data were collected from both Ratnagiri and Taketomi Island regions. Landsat-7 data with 30m spatial resolution and radiometric quantization at 8 bit dynamic range was collected. Recently, Landsat-8 data are freely available which has higher radiometric resolution, quantized over a 12-bit dynamic range (this translates into 4096 potential DN value range in an image compared with only 256 DN value range in previous 8-bit instruments). Even though ASTER data has the radiometric resolution of 8 bit, it provides the spatial resolution of 15m. Two field surveys were carried out to collect depth data using single-beam echo-sounder from the

Ratnagiri and one from Taketomi. These *in-situ* depths were separated to use for calibration and evaluation of the depth estimation. Detailed description of the characteristics of data used is reported in the Table 2.1

2.4 Data pre-processing

The raw satellite images were collected over coastal waters can be impacted by many factors. Therefore adequate pre-processing has been carried out to calibrate the satellite images to make it possible to retrieve necessary information for bathymetry estimation. The subsections below discuss the pre-processing procedures for both satellite images and *in-situ* depth in detail.

2.4.1 Tide correction of *in-situ* depth

Since the satellite images and *in-situ* depth were collected at different time and tide conditions, tide correction need to be applied for *in-situ* depth in order to synchronize the tide at the time of the corresponding image capture. The data collected from Ratnagiri shows significant tide difference therefore, appropriate tide correction was applied for the Ratnagiri data. Tide data collected by tide gauges operated by National Institute of Oceanography, India were used for applying corrections. The tide at the time of satellite image capturing is unique, but in case of field surveys, start at a time with one tide and ends at a time with different tide condition. Such variation in the tide height during *in-situ* depth data collection was also addressed by tide correction. For instance, the tide of *in-situ* depth collected on 27 May, 2012 is vary from 1.53m to 2.3m, therefore the tide of the *in-situ* depth is corrected to the tide of the image (1.15m) collected on 9 May, 2012 (Landsat-7). All the *in-situ* depths collected are tide corrected according to the tide of the time of satellite image is used to estimate depth. In case of Taketomi Island the *in-situ* depth data were collected as with zero tide. And hence, 1.37m (tide at the time of Landsat-8 image capture) was added to *in-situ* depth.

2.4.2 Radiance conversion

All the satellite data were already geometrically and radiometrically calibrated. Digital Numbers (*DN*) were converted to physical units of band averaged spectral units

(Watts/(m^2×srad×μm)). Equations for radiance ($L(\lambda)$) conversion for Landsat-7, Landsat-8 and ASTER are explained below. For Landsat-7 image,

$$L(\lambda) = (L_{max} - L_{min}/255) \times DN + L_{min} \quad (1)$$

Where, L_{min} is the minimum radiance, L_{max} is the maximum radiance and *DN* is the digital number of a particular band. For Landsat-8 the equation used to convert to radiance is:

$$L(\lambda) = ML \times Qcal + AL \quad (2)$$

Where, *ML* is band-specific multiplicative rescaling factor (RADIANCE_MULT_BAND_x, where x is the band number), *AL* is Band-specific additive rescaling factor (RADIANCE_ADD_BAND_x, where x is the band number), *Qcal* is quantized and calibrated standard product pixel values (*DN*). In case of ASTER, radiance is calculated as shown in the equation (Lillesand *et al*., 2004) below:

$$L(\lambda) = (DN - 1) \times Unit\ conversion\ coefficient \quad (3)$$

2.4.3 Water delineation and filtering

There are two very important procedures that must be undertaken prior to atmospheric and water corrections. The first step is to distinguish water from the land. The ratio of green and NIR bands was computed and ratio of greater than or equal to 1 was classified as water and less than 1 as land. In addition, a visual inspection has been carried out to well distinguish the water and land. Water region separated by masking 1 and area has been used for further analysis. The second procedure involves correction of the imagery to remove random noise and stripping. Since the methods are sensitive to random noise and striping, image smoothing with a low-pass 3×3 filter has been carried out (Lyzenga, 1981).

2.5 Satellite image correction

The radiance observed by a satellite sensor on shallow water basically consist of four components, namely, atmospheric scattering component, surface reflection component,

in-water volume scattering component, and bottom reflection component as shown in Figure 2.3 (Kanno and Tanaka, 2012). Many authors (Baban, 1993, Muslim and Foody, 2008) have also suggested a component of bottom reflectance in shallow water images. Bottom reflectance can be transformed to depth values after removing other three components (atmospheric scattering component, surface reflection component, in-water volume scattering component) successfully.

2.5.1 Correction using averaged deep water radiance

In the correction method introduced by Lyzenga, (1981), sea-surface scattering or atmospheric scattering are implicitly assumed to be homogeneous over the target area. In case of deep water, the observed spectral radiance ($L(\lambda)_i$) at infinite depth, $L_\infty(\lambda)_i$ is assumed not to include bottom reflectance, such that the water depth only consists of information related to external reflection from the water surface and atmospheric scattering. Subsequently, the effects of atmospheric scattering, surface reflection and in-water volume scattering can be eliminated by subtracting the average radiance of the deep water $L_\infty(\lambda)_i$

$$X(\lambda)_i = log(L(\lambda)_i \ - mean(L_\infty(\lambda)_i)) \quad (4)$$

Where, $X(\lambda)_i$ is log transformed radiance values of i^{th} band, *mean* ($L_\infty(\lambda)_i$) represents the averaged radiance of deep water and subtracted standard deviation of radiance for i^{th} band. Whereas pixels corresponding to shallow waters are of current interest, deep water pixels can be considered to corresponding to an infinite depth and discarded. The deep water pixels have a low overall reflectance than the shallow water pixels, and hence easy to separate. Subsequently the log transformed radiance values were calculated from blue, green, red and NIR bands separately as mentioned in equation (4). The steps for used in this correction have illustrated in Figure 2.4.

2.5.2 Correction using NIR band

In order to retrieve bottom reflectance from the spectral radiance some authors (e.g. Lyzenga, 1981 and Gholamalifard *et al*., 2013) have proposed direct subtraction of deep water spectral radiance from the shallow water spectral radiance (as explained in section

2.5.1). In this study we adopted a more refined way of retrieving bottom reflectance originally proposed by Lyzenga *et al.* (2006) and further modified by Kanno and Tanaka (2012). This method assumes that spectral radiance of deep water region is mainly a contribution of atmospheric scattering, surface reflectance and in-water volume scattering. Therefore, one can expect correlation between visible band and NIR band. A simple linear regression between each visible band and NIR band in deep water region to estimate regression coefficient as shown in equation below (Kanno and Tanaka 2012).

$$L_{\infty}(\lambda)_i = \alpha_{0i} + \alpha_{1i}\,(L_{\infty}(\lambda)_{NIR}) \tag{5}$$

According to Lyzenga (1981), Stoffle and Halmo (1991) and Gholamalifard *et al.* (2013), in order to extract bottom reflectance from shallow water region, a logarithmically transformed subtraction of mean of spectral radiance of deep water pixels from shallow water (equation 4) has to be carried out. The radiance of deep water pixel ($L_{\infty}(\lambda)_i$) from equation (5) can be substituted in equation (4) to derive a transformed radiance as shown below;

$$X(\lambda)_i = log\,(L(\lambda)_i - (\alpha_{0i} + \alpha_{1i}\,(L(\lambda)_{NIR}))) \tag{6}$$

Thus, the estimated values of α_{0i} and α_{1i} for the deep-water pixels can be used to calculate the transformed ($X(\lambda)_i$) of shallow water pixels for respective band.

2.5.3 Correction using SWIR band

Many authors (Lyzenga *et al.*, 2006, Kanno and Tanaka, 2012) have been established correction method by NIR band from the optical remote sensing to estimate better bathymetry from shallow water region. In contrast, we utilize SWIR band to correct the atmospheric and water surface components from the image and assume that the corrected bands are linearly related to the water depth. Wang and Shi (2007) have also established a correction algorithm utilizing SWIR band from MODIS data. Subsequently, deep water pixels were determined for further processing. The reflectance in visible bands ($L_{\infty}(\lambda)_i$) are proportional to those in SWIR ($L_{\infty}(\lambda)_{SWIR}$) in deep water. As explained in section 2.5.2 with the correction using NIR band, we can also expect a correlation between $L_{\infty}(\lambda)_i$ and $L_{\infty}(\lambda)_{SWIR}$ for an arbitrary visible wavelength. Further a linear regression of visible bands

($L_\infty(\lambda)_i$) bands against SWIR band ($L_\infty(\lambda)_{SWIR}$) over deep water pixels were carried out and the coefficients of regression α_{0i} and α_{1i} were computed. The equation (6) will transform the form as shown below.

$$X(\lambda)_i = log\ (L(\lambda)_i - (\alpha_{0i} + \alpha_{1i}\ (L(\lambda)_{SWIR}))) \quad (7)$$

2.6 Bathymetry estimation

2.6.1 Simple Linear Regression (SLR) model

The Stoffle and Halmo, (1991) have proposed algorithm that have been used to estimate depth from single band. The depth is estimated based on equation below:

$$\widehat{D} = -1/2kX(\lambda)_i + 1/2k(logV_0) \quad (8)$$

Where, $\widehat{D}$ is estimated depth; $X(\lambda)_i$ is corrected transformed radiance which is linearly related to the depth, k is water attenuation coefficient. V_0 is sensitivity factor related to solar irradiance at the surface, the bottom reflectance, atmospheric transmission, and sensor equipment (Gholamalifard *et al*., 2013). The equation (8) takes the form of depth = slope $X(\lambda)_i$ + *constant* (shown in equation 9). The slope of this line is related to the water attenuation coefficient such that slope is *-1/2k*, and the constant value is given by *1/2k* $(logV_0)$. The SLR model assumes a constant attenuation coefficient and homogeneous bottom type in the overall study area and also assumes the transformed radiance of single band corresponds directly to water depth. The limitation of this method is that it needs calibration depth to compute linear relationship and estimate depth.

A single band was selected from particular satellite imagery to apply SLR. In empirical modeling, the relationship between the remotely sensed radiance of a water body and the depth at sampled locations is established without considering how light is transmitted in water (Gao, 2009). To select a single band for applying algorithms, correlation coefficient is calculated by simple linear regression between the transformed bands and *in-situ* depth measurement value. The transformed band is considered as independent variable ($X(\lambda)_i$) and depth ($\widehat{D}$) as a dependent variable as shown in equation below:

$$\widehat{D} = a + bX(\lambda)_i \quad (9)$$

Any single band that provides better correlation coefficient as compared to other bands is selected for depth estimation. It has been observed that in all the multi-temporal data, NIR (0.77-0.90µm) band provides better correlation compare to other bands in case of Ratnagiri. In lighted the above observations, NIR band was chosen to apply SLR method for Ratnagiri study area.

2.6.2 Multi-temporal bathymetry (RBE)

The maximum depth value of the estimated SDB using SLR model over the study area is arbitrarily chosen as an input value for the RBE method to estimate depth when there is no corresponding calibration depth available. The workflow of RBE method is explained in detail below with an example of Landsat-7 data. First of all, RBE is applied to Landsat-7 image 30 March, 2003 and, subsequently, reliability of the method was tested with other data set too.

The RBE method directly applied to the transformed single band without need to consider about the radiative transfer of light in the sea water. The attenuation coefficient in a particular single band is calculated using SLR algorithm to understand the radiative transfer of light for a particular band. RBE is suggested to be used as an extension to SLR model to estimate depth from multi-temporal satellite imageries. In RBE, a single band is selected in order to derive better depth estimation and this particular band of other satellite imageries is used to estimate depth. Only one reference data (maximum depth value) which is estimated by SLR method has to be used as input for the RBE. The highest value for depth derived by SLR model is considered as constant for all multi-temporal images. Apart from that, the constant maximum value can also be assigned arbitrarily in case *in-situ* depth is unavailable to apply RBE. A conversion factor is calculated by the following equation.

$$Conversion\ factor = \widehat{D}_{max}/N \tag{10}$$

Where, $\widehat{D}_{max}$ is the maximum depth value derived by SLR method, which is considered as constant for other satellite imageries, N is unique cell values in a single transformed band. The conversion factor is applied to the transformed band and depth is

estimated.

To apply RBE method to an imagery only one *in-situ* depth value as maximum value is necessary. Here, maximum depth value is derived from 9 May, 2012 image (11m) by applying SLR is assigned as constant for all images of other dates in order to apply RBE. Tide correction is applied to the single maximum depth value with respect to the tide during image capture. The tide height during Landsat-7 image acquisition (30 March, 2003) was 2.01m. Since the tide height of image collected from 9 May 2012 depth is 1.15m, the maximum depth value will be changed to 11.86m after adding 0.86m to the constant maximum depth (11m) in order to correct the tide. So the maximum depth value in case of 30 March, 2003 image is changed to 11.86m. The transformed radiance (Figure 2.5b) has 68 unique cell values; hence the maximum depth 11.86m is divided by 68 to derive conversion factor (0.174). The maximum value in the transformed radiance (Figure 2.5b) is 2.36.

The cells having unique value as 2.36 in the transformed radiance were changed to 0.1745m depth. Subsequently all the cell values are transformed to depth by adding conversion factor to the transformed radiance. Eventually, the cells having unique value as 0.54 in the transformed radiance are assigned to 11.86m depth. Hence, 68 classes of unique depth values are derived by the RBE method (Figure 2.5c). The same way, depth was estimated from other data sets such as Landsat-8 and ASTER. The count of different cell values (denoted by *N* in equation 10) will vary according to the *DN* value ranges of the data.

2.6.3 Multiple Linear Regression model (Global model)

Several authors (Clark *et al.*, 1988, Hamilton *et al.*, 1993, Kanno and Tanaka, 2012) recognized Multiple Linear Regression (MLR) model would provide good depth estimation using multiple bands over shallow water region. Originally, Clark *et al.* (1987) proposed a multiple regression analysis between *in-situ* depth and multispectral bands to estimate regression coefficients. Further, these coefficients have been used to estimate SDB. This method addresses the heterogeneity of the bottom type by utilizing all the

visible bands. This algorithm attempts to isolate water attenuation and hence depth using a combinations of spectral bands. MLR utilizes multispectral band, so that it accounts for varying attenuation coefficients for different bottom types as it calculates water depth (Van Hengel and Spitzer, 1991). The equation to estimate SDB using multispectral bands is following:

$$\widehat{D} = \beta_0 + \beta_1 X(\lambda)_1 + \beta_2 X(\lambda)_2 + \cdots + \beta_n X(\lambda)_i \quad (11)$$

Where $\widehat{D}$ is the estimated depth, $\beta_0, \beta_1, \beta_2, \ldots, \beta_n$ are derived constants and $X(\lambda)_1, X(\lambda)_2, \ldots, X(\lambda)_i$ are the log transformed atmospheric, water surface and water column corrected radiance values of multispectral bands. Where, log transformed bands ($X(\lambda)_1, X(\lambda)_2, \ldots X(\lambda)_i$) are taken as independent variable and *in-situ* depth has taken as dependant variable. The coefficients ($\beta_0, \beta_1, \beta_2 \ldots \beta_n$) derived from multiple linear regression have been used to estimate the depth for all pixels. The workflow is explained in Figure 2.4.

2.6.4 Multiple Linear Regression model using new band combination

There is a lack of agreement amongst researchers about the right band combination for better SDB estimation. Therefore this study has investigated different combination of multispectral bands in order to evaluate the performance. Further, a new band combination that includes visible and NIR band for improved bathymetry estimation is proposed. As described in the equation (11) multiple linear regressions were carried out between the log transformed visible+NIR bands and calibration depth. The new combination of bands was tested with all available data sets with corresponding *in-situ* depth. If the corresponding *in-situ* depth was not available, *in-situ* depth collected close to the image capture date was used for multiple regression and calibration. The details of the *in-situ* depth and satellite imageries used for SDB estimation are shown in Table 2.1. From 2000 to 2014, various multi-temporal and multi-source data sets were used to estimate the SDB along Ratnagiri coast and Taketomi Island. Landsat-7, Landsat-8 and ASTER data were used to estimate depth and used to compare the behavior of the algorithms in terms of various spatial and radiometric resolutions.

2.7 Results

In this section, results derived from each model explained in section 2.6 are described. SDB results derived from existing models such as SLR and MLR are presented. The comparative analysis of the results derived from methods which originally proposed in this Chapter is also discussed.

In Ratanagiri study area, SLR model was applied to 9 May, 2012 (Landsat-7) image, 12 November, 2013 image (Landsat-8) and 10 December, 2012 (ASTER) are calibrated and evaluated using *in-situ* depth collected on 27 May, 2012, 17 November, 2013 and 17 November, 2013 respectively. The evaluation of the results derived by RBE was carried out with *in-situ* depth collected closer to the image acquisition date. In the case of 8 May, 2000 and 30 March, 2003 (Landsat-7) and 16 January, 2003 (ASTER) images, NHO chart data surveyed on 2003 was used, In case of 31 January, 2014, 4 March, 2014 (Landsat-8) and 10 December, 2012 (ASTER) *in-situ* depth collected on 17 November, 2013 was used. The comparison of estimation results of all the satellite imageries used in this Chapter carried out by correlation coefficient (R), coefficient of determination (R^2) and RMSE. These parameters used to evaluate the accuracy of results shows that SDB from Landsat-8 data provide better accuracy than other satellite imageries (Table 2.2). Landsat-8 data was collected on 31 January, 2014 and 04 March, 2014 were providing correlation coefficient of 0.89, 0.80, coefficient of determination of 0.81, 0.66 and RMSE of 1.39m, 2.09 m respectively. Figure 2.6a and Figure 2.6b are showing the bivariate scatter plot between SDB estimated by RBE model and reference depth from Landsat-8 data. High radiometric resolution of Landsat-8 data provides better depth estimation as compared to other data sets. Even though ASTER data has high spatial resolution compared to Landsat-8, it was not effective in providing comparable results of Landsat-8 data.

Several authors (Clark *et al.,* 1987, Clark *et al.,* 1988, Kanno *et al.,* 2011 and Pacheco *et al.,* 2015) have reported that reliable SDB estimation could be made using MLR model. Therefore, the research has been carried out to estimate depth from MLR model with different combination of spectral bands. A combination of visible and NIR (MLR-NIR) was used to estimate SDB from all available satellite imageries. In case of Landsat-

7, four bands (0.45-0.90µm), in case of Landsat-8 five bands (0.43-0.88µm) and in case of ASTER three bands (0.52-0.86µm) were used. The results derived by MLR-NIR model were compared and evaluated with the results derived by MLR model. The results show that SDB by MLR-NIR model has provided better accuracy than the MLR Table 2.3. Especially, in Landsat-8 data (12 November, 2013, 31 January, 2014 and 04 March, 2014) provides better accuracy results in terms of correlation coefficient (0.95, 0.95, 0.90), coefficient of determination (0.91, 0.91, 0.83) and RMSE (0.83m, 1.23m) respectively. The Figure 2.6a shows the bivariate scatter plot SDB and reference depth derived from Landsat-8. Figure 2.6c, Figure 2.6e and Figure 2.6g show the bivariate scatter plot of SDB from MLR and Figure 2.6d, Figure 2.6f and Figure 2.6h are from MLR-NIR. Bivariate scatter plots clearly demonstrate that the proposed new band combination (MLR-NIR) is better than the previous results. The new band combination shows significant increase in the accuracy irrespective the satellite imageries used. Observed higher accuracy in Landsat-8 was due to the availability of new coastal aerosol (0.43-0.45µm) band.

In case of Ratnagiri, turbidity which occurs due to the sedimentation from Kajali River was a factor which could impact the accuracy of the results. Therefore, evaluation tests were also carried out to estimate accuracy and error obtained due to turbidity in depth estimate derived by both single band and multispectral band models from Landsat-8 data. The study area was divided into two sections based on NDWI. The NDWI originally introduced by Rogers and Kearney (2004) to demarcate coastal marsh, where, lower NDWI value denote high turbidity and higher NDWI values indicate relatively low turbidity. The evaluation tests confirm that the accuracy of the depth estimates at low turbid area is higher than the high turbid region. Table 2.4 summarizes the evaluation results carried out by using R, R^2 and RMSE. The RMSE values indicate that error associated the depth estimates increases in high turbid region. The estimated SDB results at high turbid region and low turbid region are compared using bivariate scatter plot in Figure 2. Low turbid region shows stronger linear relationship between estimated depth and reference depth than high turbid region. Figure 2. demonstrates the scatter plot of

SDB from Landsat-8 data (31 January, 2014) by MLR-NIR method and reference depth. Other data summarized in the Table 2.4 also show the similar trend. However, the turbidity due to wave action along the coast or dynamic sediment discharge near estuary is potentially responsible for reducing the accuracy of the depth estimation.

Study has been further focused to develop an improved SDB estimation method using established MLR-NIR method by integrating a better correction method. The correction method is basically proposed by Lyzenga *et al.* (2006) (details of the correction method explained in section 2.5.2). Many studies demonstrated (e.g. Kanno and Tanaka, 2012) that improved correction algorithm proposed by Lyzenga, (2006) has potential to produce good quality SDB. On the other side these authors used NIR band for atmospheric and water corrections. However, this leads to significant errors in the shallow depth region due to reflectance from the bottom in NIR band (Manessa *et al.*, 2014). Since, NIR band (0.77-0.89µm) is sensitive towards the low depth region and variety of bottom types like very bright sand bottom. Therefore, an approach has been carried out as correction method using SWIR band for as explained in section 2.5.3.

As mentioned before, many authors were observed bottom reflectance components in NIR region of the electromagnetic spectrum. Therefore this study utilizes available four visible bands and one NIR (MLR-NIR) band from Landsat-8 images. The key contribution of this study is that the introduction of new correction method using SWIR (1.57-1.65µm) band. We utilize the fact that the longer wave length bands attenuate very soon in the water region. Hence, we expect no reflectance components from the bottom even at shallow depth. The studies have carried out with Landsat-8 data over Taketomi and Ratnagiri regions. Results derived by the SWIR correction method also show good correlation with reference depth data in terms of R, R^2 and RMSE are at Taketomi, 0.90, 0.81, 2.19m and at Ratnagiri 0.90, 0.81, 1.24m respectively. In order to comprehend the outcome of the SWIR band correction method, a comparative analysis was carried out with NIR band correction method and SWIR correction method provides better SDB estimation than NIR band correction method. Details of the comparison assessment are reported in Table 2.5 and bi-variate scatter plot shown in Figure 2.7.

Chapter Three
Geographically Weighted Regression Based SDB Model

3.1 Introduction

The common practice adopted in previous attempts on SDB has been to calibrate a single set of coefficients using MLR models (hereafter referred to as global model). Theoretically, global model were purported to address the spatial heterogeneity introduced by different bottom classes (Van Hengel and Spitzer, 1991). In practice, such global models do not fully address the heterogeneity as they only reflect the relation between the spatial average of spectral radiance and water depth. Therefore, we consider that depth estimate derived by global model can result in significant errors. The previous Chapter has discussed about the benefit of using global model in order to overcome the problems related to SLR. Nonetheless, previous attempts have not discussed about the limitations of the global model in addressing the issues generated by heterogeneity due to variations in bottom types and water quality. The derived coefficients in global model are used to estimate depth from the transformed radiance (e.g. Clark *et al*., 1988, Pacheco *et al*., 2015). However, single set of coefficient is inadequate to address variation in coastal waters due to local conditions. In order to address this lacuna in conventional global model, this study proposes and evaluates the GWR model to address spatial heterogeneity issues and improve the SDB results. The performance of global model, motivation to the GWR model and improvements made by the GWR model in addressing heterogeneity will be presented in this Chapter.

3.2 Study area

The near-shore coastal waters of the south-western Puerto Rico, in north-eastern Caribbean Sea were chosen as the test case to derive water depth. Study area shown in Figure 3.1, geographically stretches from 17°49' N - 18°5' N, and 67°00' E - 67°18' E and covers 160km^2 along 30km coastal stretch off Puerto Rico. The Puerto Rico was selected as the study area due to two reasons; one of them is availability of open high resolution

LiDAR depth (16-bit Geotiff image with 4 × 4m spatial resolution)[3] provided by NOAA. Secondly, clear water conditions observed in the region could afford better estimation.

3.3 Data used

LiDAR bathymetric data provided by NOAA were acquired by using a Laser Airborne Depth Sounder (LADS) Mk II Airborne System which is surveyed from 07 April, 2006 to 15 May, 2006. This airborne LiDAR is a 900 Hz Nd: YAG (neodymium-doped yttrium aluminium garnet) laser, which is split by an optical coupler into an infrared (1.064µm) beam and a green (0.532µm) beam. Final product image (16 bit Geotiff image) is produced at 4 × 4m bathymetry surface (Hernandez and Armstrong, 2016). The horizontal accuracy of dataset is better than ±5m + 5% of the depth, and the vertical accuracy or maximum total vertical uncertainty of dataset is better than ±0.82m (Costa *et al.*, 2009).

Landsat-8 and RapidEye data were collected in order to test the algorithm at moderate and high spatial resolution. The heterogeneous bottom type and clear water condition have been noted by visual examination of satellite images collected. Both Landsat-8 and RapidEye have higher radiometric resolution quantized over a 12-bit dynamic range can be rescaled to radiance using radiometric rescaling coefficients provided in the product metadata file. In addition, RapidEye provides higher spatial resolution of 5m as compared to 30m spatial resolution of Landsat-8. Five spectral bands (four visible and one NIR) of Landsat-8 were used for SDB estimation and SWIR (1.57-1.65µm) band was used for correction. In case of RapidEye data four visible bands were used for estimation and NIR (0.76-0.85µm) was used for correction. The characteristics of the data used for this study is given in Table 3.1. In this Chapter LiDAR derived depth is used as *in-situ* depth, and hence in this Chapter *in-situ* depth will be referred as LiDAR.

3.4 SDB models

The study first of all estimated SDB using global model by applying equation (11). As explained in the equation (11), transformed radiance ($X(\lambda)_i$) have been considered as independent variable and randomly distributed LiDAR derived depth points as dependant

[3] https://data.noaa.gov/dataset/noaa-geotiff-4-meter-lidar-bathymetry- u-s-caribbean-puerto-rico-southwest-projects-opr-i305-

variable in a global model. The global model was tested using 30m spatial resolution Landsat-8 image of 25 November, 2013. Firstly, correction algorithm has been applied to each visible spectral band such as coastal aerosol (0.43-0.45µm), blue (0.45-0.51µm), green (0.53-0.59µm), red (0.64-0.67µm) and NIR (0.85-0.88µm) bands. The large attenuation coefficient of SWIR band (1.57-1.65µm) is used to correct the spectral radiance of each band effectively by using equation (7). Further, transformed radiance are used for depth estimation as shown in equation (11). A total of 10,000 randomly distributed LiDAR derived depth points were used to estimate SDB and 2,000 independent LiDAR depth points used for validation.

The global model was also tested on 5 m spatial resolution RapidEye image collected on 01 May, 2010. Firstly, the correction has been applied to each visible spectral bands, namely blue (0.44-0.51µm), green (0.52-0.59µm), red (0.63-0.69µm) and red-edge (0.69-0.73µm). Since there is no SWIR band available for RapidEye, NIR band (0.76-0.85µm) was used for correction (equation 6). Considering that a single pixel of Landsat-8 image corresponds to about 36 RapidEye pixels, 360,000 randomly distributed LiDAR derived depth points were used to estimate SDB and 2,000 independent LiDAR depth points used for validation.

3.4.1 Class based model

As opposed to global model, class based models is also a linear regression model that computes β-coefficients for each class. These class based coefficients will be further used to estimate depth for each class and merged together for SDB. A residual map was generated by subtracting LiDAR depth from estimated depth to assess the effects of spatial heterogeneity in the global model. Six classes were generated by maximum likelihood classification; signatures of the classes were determined by visual examination of the residual map. Therefore, the equation (11) is transformed to equation (12) as shown below.

$$\widehat{D}_{(j)} = \beta_{0(j)} + \beta_{1(j)} X(\lambda)_{1(j)} + \beta_{2(j)} X(\lambda)_{2(j)} + \cdots + \beta_{n(j)} X(\lambda)_{i(j)} \quad (12)$$

Where, Where $\widehat{D}_{(j)}$ is the depth at j class, β_0 is representing the y-intercept and $\beta_{1(j)}$,

$\beta_{2(j)}$,.., $\beta_{n(j)}$ are the slope of the multiple linear regression at j class respectively. $X(\lambda)_{1(j)}$, $X(\lambda)_{2(j)}$, ..., $X(\lambda)_{i(j)}$ are the transformed radiance at *j* class.

3.4.2 Geographically Weighted Regression (GWR) Model

In order to address local heterogeneity, GWR could be a feasible and effective method as compared to global model since it allows local coefficients to be applied. Several authors (Brundston *et al.*, 1996, Fotheringham, *et al.*, 1998 and Yrigoyen *et al.*, 2008) have suggested the efficacy of GWR as a spatial predictor model.

As opposed to non-weighted global and class based models described in the earlier section, the GWR model is a weighted regression model that computes β-coefficients for each pixel. In GWR, β-coefficients are determined using moving kernel and pixels close to the centroid are assigned higher weights than the pixels away from the centroid. Distance decay of the weights can be estimated using different functions such as exponential, bi-square and Gaussian. In general, weighting schemes can be broadly classified as continuous and discontinuous depending on the functions used. The discontinuous bi-square function gives fractional decaying weights according to the proximity of the current pixel to centroid (Lu, *et al.*, 2014). In GWR model, the term bandwidth denotes the size of the kernel window. As shown in equation (13), bi-square function decay of weighting would be applied only if distance from the centroid is less than bandwidth else the assigned weight will be zero. On the other hand, the continuous Gaussian function (equation 14) gives fractional decay of weights according to the proximity of the current pixel to centroid even if distance from the centroid is greater than bandwidth.

$$w_p = (1 - (d/bw)^2)^2 \tag{13}$$

$$w_p = exp(-0.5 * (d/bw)^2) \tag{14}$$

Where, *bw* is the bandwidth, *d* is the distance from current pixel to the centroid and w_p is the weight assigned to the current pixel.

Bandwidth determines the spatial coverage of the local kernel and assigning

appropriate bandwidth is critical. There are two ways of assigning bandwidth, one is fixed bandwidth (Fixed-GWR) and other is adaptive bandwidth (A-GWR). In A-GWR model, size of the kernel is set by considering density of calibration points in the local neighborhood. As illustrated as A-GWR model in Figure 3.2, the size of the kernel is smaller when calibration points are denser, and size of the kernel increases when calibration points are sparse. Previous researchers (e.g. Yrigoyen *et al.*, 2008) have demonstrated that, adaptive bandwidth performs relatively better, particularly in case of randomly distributed calibration points.

Large bandwidth will produce results similar to global model, and spatial heterogeneity cannot be resolved. The problem is, therefore to select the optimal bandwidth. Optimal bandwidth can be determined using a cross-validation approach wherein the validation scores are minimized (Harris *et al.*, 2010 and Lu *et al.*, 2014). This minimization of cross-validation score can be carried out by estimating the Root Mean Square Predicting Error (RMSPE) for a range of bandwidth. The bandwidth which provides minimum RMSPE (equation 15) is considered as optimal (Harris *et al.*, 2010 and Harris *et al.*, 2011) and used in subsequent estimation.

$$RMSPE = \sqrt{\left(\frac{1}{N}\right)\sum_{i}^{N}\left(D_{(x,y)} - \widehat{D}_{(x,y)}\right)^2} \quad (15)$$

Where, N is the number of neighborhood points of an adaptive bandwidth, $D_{(x,y)}$ is the actual depth and $\widehat{D}_{(x,y)}$ is predicted depth.

Appropriate weighting function and bandwidth can be selected to estimate β-coefficients for each pixel of independent variable using A-GWR model. The derived β-coefficients are further used to estimate SDB as shown in equation (16).

$$\widehat{D}_{(x,y)} = \beta_{0(x,y)} + \beta_{1(x,y)} X(\lambda)_{1(x,y)} + \beta_{2(x,y)} X(\lambda)_{2(x,y)} + \cdots + \beta_{n(x,y)} X(\lambda)_{i(x,y)} \quad (16)$$

Where $\widehat{D}_{(x,y)}$ is the depth at a particular location, β_0 is representing the y-intercept and β_1, β_2, β_n are the slope of the weighted multiple linear regression respectively.

The workflow of the A-GWR model has shown step by step in Figure 3.2. The first

step is to convert georeferenced multispectral bands (DN values) to spectral radiance using the radiometric rescaling coefficients available as metadata provided along with satellite image. In the second step, deep water pixels are selected by visual interpretation to estimate linear regression coefficients as explained in section 3.2). The third step involves the atmospheric, water column and water surface reflectance corrections. In fourth step, log transformation of the corrected spectral radiance was carried out. Further, in the fifth step corrected or transformed spectral radiance applied to GWR model. Determination of β-coefficients by calibrating LiDAR derived depth, and water depth estimation were carried out as the final step. Since, in GWR model same LiDAR derived depth is used for calibration and validation the optimal bandwidth determined by cross-validation procedure may not be appropriate. Guo *et al*. (2008) also observed that the GWR cross-validation may not reflect an optimal bandwidth. Therefore investigations were undertaken to evaluate the depth estimates obtained over a range of bandwidth using both bi-square and Gaussian weighting functions.

3.5 Results

Various SDB methods described in the previous section were tested using Landsat-8 and RapidEye images over dynamic and complex coastal water of Puerto Rico. A total of 62 and 2,250 randomly distributed calibration points per square kilometer were used for Landsat-8 and RapidEye images which comprise of a total of 10,000 and 360,000 respectively. Further, 2,000 randomly distributed LiDAR derived depth points were used to validate the performance and behavior of the models in heterogeneous coastal water region.

Initially, the performance and behavior of the global model were evaluated for Landsat-8 image. Depth estimated by global model appears to be consistent in terms of correlation coefficient (0.88) and coefficient of determination (0.78), but shows high RMSE of 2.63m. Further, a class based model was generated as explained in section 3.4.1 to evaluate the performance of the model in addressing heterogeneity. Bivariate scatter plot between LiDAR depth and estimated depth by global model over each class is shown in Figure 3.3a. The multiple trend lines in Figure 3.3a suggested the necessity to

determine coefficients based on individual classes. Therefore, in the proposed class based model, coefficients were calculated for 6 classes and were used to estimate depth for each class. Accuracy of the class based estimation and global model is compared using a density plot (Figure 3.3b) and statistical analysis. The results shown in Table 3.2 clearly demonstrate an improved depth estimation in class based model.

In case of RapidEye, the global model results are showing correlation coefficient of 0.88, coefficient of determination (0.78) and RMSE of 2.48m. Further, class based model was applied for improved depth estimation. In case of class based model, procedures of generation of classes are same as in case of Landsat-8. The improvement in estimation using class based model is reflected in bivariate plot shown in Figure 3.4a. The results (Table 3.2 and Figure 3.4b) indicate an improvement in depth estimation and clearly demonstrate that heterogeneity can be better addressed by using class based coefficients. Significant improvements have been observed in depth estimates derived from class based model indicating the robustness of using a local coefficient to address heterogeneity rather than using a global coefficient.

Furthermore, A-GWR model was applied as it can be expected to perform better than class model by weighting regression coefficients locally. In A-GWR model, an appropriate weighting function was evaluated by cross-validating the depth estimation results computed by using bi-square and Gaussian functions. The weighting function which provides minimum RMSE and appropriate range of minimum and maximum value was selected as the suitable weighting function. Gaussian function provides relatively better estimates for both Landsat-8 and RapidEye in terms of RMSE. Range of estimated depth shows close correspondence with the LiDAR derived depth. Considering above observations, Gaussian weighting function was used for SDB.

Optimal bandwidth selection was carried out by cross-validation over a bandwidth range of 5 to 200 neighborhood points. Consequently, 5 and 15 neighborhood points were selected as optimal bandwidth for Landsat-8 and RapidEye as they show minimum RMSE of 1.13m and 0.41m respectively. As a result of the optimal bandwidth selection, the range of estimated depth show close correspondence with the range of LiDAR

derived depth. The SDB result derived from Landsat-8 has shown in Figure 3.5.

The accuracy of the SDB estimates of A-GWR model were compared and evaluated for Fixed-GWR and global model (Table 3.3). A-GWR model results show significant improvement for both Landsat-8 and RapidEye and show strong linear relationship (Figure 3.6a and Figure 3.6c). Errors associated with underestimation due to high reflectance and overestimations due to low reflectance in the spectral bands are evident from the RMSE values. The RMSE for Landsat-8 is reduced from 2.68m to 1.14m and for RapidEye image from 2.48m to 0.41m. Density curve of residuals for Landsat-8 and RapidEye calculated by subtracting estimated depth from LiDAR derived depth are shown Figure 3.6b and Figure 3.6d respectively. In Figure 3.6b peak of the curve extends up to 0.8 and around 90% of area under the curve is covered by the values very close to zero. In Figure 3.6d peak of the curve extend up to 2.0, almost 95% of area under the curve is covered by the values very close to zero and these results reveal a significant improvement in SDB estimation.

The results were verified by several cross-sectional profiles as shown in Figure 3.7a. The cross shore profile (A-B in Figure 3.7a) across continuous surface for Landsat-8 (Figure 3.7b) and RapidEye (Figure 3.7c) indicate excellent match with the LiDAR derived depth for both. Two profiles cross shore profile (C-D and E-F in Figure 3.7a) across discontinuous surface with islands for Landsat-8 (Figure 3.7d, Figure 3.7f) and RapidEye (Figure 3.7e, Figure 3.7g) are also showed good agreement. In case of medium resolution Landsat-8 relatively fewer LiDAR derived depth points were used to estimate depth, consequently cross profiles show slight discrepancies in depicting crest and trough of the profile. The statistics of difference of profiles were calculated Table 3.4. Generally, the statistics of profiles are also showing good agreement with LiDAR depth. However, the profile across continuous surface (A-B) is showing better estimation than other profiles across discontinuous surface (C-D and E-F).

The estimated depth data were separated to 5 depth categories at 5m interval to evaluate the performance of the A-GWR model in each depth category and to determine

the maximum depth that could be reliably estimated by the model. Details of the accuracy assessment in each depth category are given in the Table 3.5. Depth derived from Landsat-8 and RapidEye data show better results in 1-5m depth category compared to the other depth categories as reflected by RMSE and up to 20m all depth categories show no significant difference in terms of correlation coefficient and coefficient of determination. The RMSE indicates that there are only very few pixels are over or under estimated in 1-5m depth category. Depth beyond 20m is considered as limit for reliably estimating depth in this study area as see from decrease in correlation coefficient, coefficient of determination and increase in RMSE. Lower accuracy observed in depth estimate of Landsat-8 compared to RapidEye for each depth category is due to the fewer LiDAR depth points used for estimation.

The study further evaluates the efficacy of A-GWR model in two practical scenarios where high resolution reference depth data may not be available. Two scenarios for available reference depth data could be considered, namely, low resolution gridded data extracted from bathymetric charts or freely available ocean data (e.g. GEBCO) or field echsounder survey data. Considering the first scenario, (Scenario-1) water depth is estimated using sparse grid of LiDAR derived depths. Figure 3.8, demonstrates the distribution of LiDAR depth points as in Scenario-1. Figure 3.9, demonstrates LiDAR derived depth along arbitrary tracks, as in field based data collection used in case of the second scenario (Scenario-2).

In order to test the A-GWR model for Scenario-1, LiDAR derived depth at an interval of 200 meter were prepared. A total of 5,200 LiDAR derived depth were used for both Landsat-8 and RapidEye to estimate the depth using A-GWR model. Optimal bandwidth selection was carried out by cross-validation over a bandwidth range of 5 to 200 neighborhood points. As a result, 15 neighborhood points were selected as optimal bandwidth for both Landsat-8 and RapidEye as they show minimum RMSE of 1.38m and 1.55 m respectively. Results shown in Table 3.6 suggest that, even with sparse gridded calibration data, A-GWR model provides excellent results. Figure 3.10a (Landsat-8) and Figure 3.10b (RapidEye) show the bivariate scatter plot between estimated depth by A-

GWR model and LiDAR derived depth and reveals the excellent agreement between them. Correlation coefficient is as high as 0.96 and 0.97 and coefficient of determination (0.94 and 0.95) and RMSE (1.38m and 1.55m) are also showing reliable estimates for Landsat-8 and RapidEye respectively. Capability A-GWR model in resolving features in terms of calibration data density have been attempted by many authors (e.g. Chen *et al.*, 2011). Visual examination of depth estimate derived in case of Scenario-1 reveals that larger bandwidth not effective in resolving finer details in bathymetry.

In Scenario-2, the traditional way of field data collection with echo-sounder survey is considered for applying the A-GWR model. LiDAR derived depth along arbitrary track shown in Figure 3.9 is used to calibrate the A-GWR model. Same tracks are used for both Landsat-8 and RapidEye images; a total of 5,440 for Landsat-8 and 32,640 LiDAR derived points for RapidEye along the track were used. Further, cross-validation of estimated depths was carried out at ranges of bandwidth from 5 to 200. Consequently, 20 and 25 neighborhood points were selected as optimal bandwidth for both Landsat-8 and RapidEye image with minimum RMSE of 1.40m and 1.44m for Landsat-8 and RapidEye respectively.

Bivariate scatter plot between LiDAR depth and estimated depth are shown in Figure 3.11a and Figure 3.11b for Landsat-8 and RapidEye image respectively. The results indicate that Scenario-2 also provides reliable estimate as evident from the high correlation coefficient (0.96 and 0.96), coefficient of determination (0.93 and 0.92) and RMSE (1.40m and 1.44m) for Landsat-8 image and RapidEye image respectively.

In this study we were also evaluated the performance of bi-square and Gaussian weighting function over different surface and different distribution pattern of calibration points. Two sub-areas (AoI-1 and AoI-2 are shown in Figure 3.9) in the study are used for the evaluation test. AoI-1 is a continuous surface, while AoI-2 is discontinuous surface with many islands. Two types of distribution of the LiDAR derived points (randomly distributed points and arbitrary tracks) were used over both AoI-1 and AoI-2.

In case of continuous surface (AoI-1), good estimates are reported for both bi-square

and Gaussian weighting functions. However, bi-square with RMSE of 0.43m and range of estimated depth show close correspondence with the range of LiDAR derived depth with minimum value as 0.76m and maximum as 21.5m provides slightly better estimate than Gaussian. Arbitrary ship track LiDAR points over AoI-1 provides almost similar results for bi-square and Gaussian with RMSE (0.63m and 0.64m), range of estimated depth with minimum (-12m and -13m) and maximum (22m and 21m) respectively. In case of discontinuous surface (AoI-2), accuracy of estimation is poor compared to results of AoI-1. However, randomly distributed LiDAR points over AoI-2 provides similar results for both bi-square and Gaussian with RMSE (1.45m and 1.42m), range of estimated depth with minimum (-22m and -22m) and maximum (21m and 20.78m). Arbitrary ship track LiDAR points over AoI-2 provides better estimates by Gaussian with RMSE of 1.64m and range of estimated depth show close correspondence with the range of LiDAR derived depth with minimum value as -15m and maximum as 21m.

The above results indicate that depth estimates with A-GWR model are sensitive to the bandwidth and the kernel weighting function used. Distribution, density of calibration points and spatial continuity of the surface to be estimated also influence the results. In case of continuous surface, when calibration points are randomly distributed, bi-square weighting function could provide better estimates. In case of LiDAR points distributed along arbitrary track, using a Gaussian weighting function or tuning the bandwidth for bi-square by cross-validating the results may be needed. In case of discontinuous surface, Gaussian weighting function could work better.

Chapter Four

Application for Integrated Coastal Relief Model and Tsunami Simulation

4.1 Introduction

GWR models have shown significant improvements as predictive weighted linear regression model in various fields (Brundston *et al.,* 1996, Fotheringham, *et al.,* 1998, Yrigoyen *et al.,* 2008 and Lu, *et al.,* 2014). The previous Chapter illustrated that the newly proposed GWR based SDB model is able to produce high quality SDB. The results derived from GWR based model suggests that it can be applied in various conditions of coastal water, density of calibration depth points, spatial and radiometric resolution of the satellite data. Therefore in this Chapter, the application of derived SDB as input for practical scenarios is attempted. Study aims to generate an ICRM over parts of Miyagi prefecture by combining derived SDB with various resolutions topographic and bathymetry data. Developed ICRM could be used to demonstrate a practical application scenario of the SDB in tsunami simulation. Therefore, this Chapter evaluates the application of generated coastal bathymetry in tsunami simulation to hindcast the real event. Coarse accurate coastal bathymetry data were used for many previous tsunami simulation application researches and indicated the drawbacks due to lack of reliable accuracy coastal bathymetry data in tsunami simulation (Griffin *et al*., 2015). Hence, research aims to use reliable accuracy ICRM developed over Miyagi region. Details of the data and study area are discussed in the following sections.

4.2 Study area and data usage

Coastal area of Tohoku, Japan was significantly affected by earthquake driven Tsunami occurred off Japan, on March 11, 2011 (Okayasu *et al*., 2013). This tsunami was the third mega earthquake generated tsunami in this decade. Miyagi, Japan was the one of the most affected prefectures in the region (Mori *et al*., 2011). The epicenter of the Tohoku tsunami was reported as 38°19' N, 142°22' E and Miyagi prefecture was the closest location. Therefore, a tsunami simulation has been carried out in parts of Miyagi prefecture

(Figure 4.1) along 110km coastal stretch.

Medium spatial resolution (15m) ASTER Open Data collected on 10 September, 2010 were used to estimate SDB. Medium radiometric resolution (8-bit quantized) green band (0.52 - 0.60µm) and red band (0.63 - 0.69 µm) were used for estimation and NIR band (0.76 - 0.86µm) used for correction (Table 4.2). J-DOS bathymetry[4] data of 500m grid resolution (JODC-Expert Grid data for Geography) was used as calibration depth points to estimate 15m reliable SDB approximately over an area of 219km^2 in a part of Miyagi coastal area. A total of 649 evenly distributed *in-situ* depth points were used for calibration and for evaluation of accuracy. Even though depth points were sparse, the distribution of the data was at equal interval. Therefore, Fixed-GWR model was used to estimate the SDB and which was expected to perform better than A-GWR model in this situation. The results of SDB generated from Fixed-GWR shows better agreement with reference depth in terms of R (0.93), R^2 (0.87) and RMSE 1.65m than A-GWR model R (0.91), R^2 (0.84) and RMSE (1.95m). Both ASTER satellite image and J-DOS data were collected before the tsunami occurred, therefore the estimated SDB can be considered to be reliable.

4.3. Integrated Coastal Relief Model

ICRM is a regional bathymetric-topographic model for the area of a part of Miyagi prefecture. This model incorporates the most recent, publicly available topographic data and bathymetry data to provide a comprehensive view of Miyagi coastal zone (Figure 4.1). Publicly available low resolution (900m) offshore bathymetry collected from GEBCO[5], medium resolution (30m) terrestrial topographic data collected from SRTM[6] were combined with SDB estimated using ASTER (15m) to develop the seamless ICRM. The ICRM was generated using ANUGA Open Source Hydrodynamic model (Nielsen *et al.*, 2005). ANUGA is an Open Source Software package with most components being implemented in Python[7] and is capable of modeling the impact of hydrological disasters such as dam breaks, riverine flooding, storm-surge or tsunamis.

[4] http://jdoss1.jodc.go.jp/vpage/depth500_file_j.html
[5] http://www.gebco.net/
[6] http://srtm.csi.cgiar.org/
[7] https://www.python.org/

Firstly, the raster GEBCO, SRTM and SDB data were individually converted to ASCII format using *r.out.ascii* module of GRASS GIS. Further this ASCII data was converted into ANUGA DEM format using the script *anuga.asc2dem*. Further, the DEM file was converted to point format using anuga.dem2pts. Further, anuga.geospatial_data.geospatial_data (ANUGA script) was used to combine the point file data of multi-resolution topographic data to generate a single seamless ICRM.

4.4 Tsunami simulation

In case of tsunami simulation, 2D non-linear shallow water equations are commonly implemented and solved numerically on a mesh or grid, and there are many software solutions available for tsunami simulation, such as TUNAMI (Goto *et al.*, 1997), ANUGA, and TsunAWI (Rakowsky *et al.*, 2013). In ANUGA, the conservative finite volume scheme allows discontinuities at the edges of all mesh triangles and therefore can simulate wetting and drying of mesh elements (Jakeman *et al.*, 2010). Tsunami inundation is simulated through numerical solution of the non-linear shallow water equations over a model of bathymetry and topography with appropriate extensions to model wetting and drying processes.ANUGA uses unstructured triangular meshes, with internal polygons used to define the maximum allowable size of an individual mesh. The ICRM represents unstructured triangular meshes as explained in section 4.3. Boundary conditions such as rainfall, tide, wind stress, surface roughness, etc., were not considered in the simulation process. In addition, mean sea level was assumed as the initial water level and fixed Manning's coefficient was used.

4.5 Results

Tsunami simulation was carried out with duration as 50 minutes and results were evaluated with post-tsunami survey results (Mori *et al.*, 2011). In this study, we compared the inundation extent and heights of the tsunami simulation with post-tsunami survey data. The simulation experiment is showing that around 115km^2 of area was inundated. Simulation tsunami about 14m high traveled inland up to 5km and Ishinomaki and Higashimatsushima region were affected the most. These regions are low lying land and Kitakami River is passing through Ishinomaki region, these factors were amplifying the

tsunami inundation (Figure 4.2 and Figure 4.3). Figure 4.2 and Figure 4.3 illustrates the comparison between the simulation results and the post-tsunami survey results. Figure 4.2a shows the actual extent of the inundation and Figure 4.2b shows the simulated inundation extent in 50 minutes duration. The surveyed tsunami inundation heights at 243 points were overlaid on simulated tsunami inundation height map (Figure 4.3a). Figure 4.3b demonstrates the correlation between surveyed and simulated tsunami inundation heights. The Figure 4.3a and Figure 4.3b depict that the maximum inundation observed was about 10m and 14m in surveyed and simulated events respectively. The evaluation of the simulated tsunami inundation heights shows a reasonable agreement with post-tsunami survey data in terms of R (0.75), R^2 (0.57) and RMSE (2.98m).

Chapter Five
Implementation of GWR Based SDB Model as Open Source GIS Module

5.1 Introduction

The previous Chapters discussed the several SDB models for improved SDB estimation. Among them, GWR based SDB model proposed in Chapter Three has been shown significant improvements for SDB. Further, evaluation of performance with application scenarios in Chapter Four also has shown robustness of the model in practical use. In addition, authors like Su *et al.*, 2014 also has successfully used GWR model as an improved SDB method. However, until now, no software module has been proposed to automate the procedure of GWR based SDB models. Therefore, this Chapter is proposing a new Open Source GRASS GIS module called *i.image.bathymetry* to automate the SDB estimation using multispectral bands. Since *i.image.bathymetry* has been entirely implemented using a Free and Open Source GIS framework, it can be easily applied in other areas without the need to invest resources for software.

5.2 Validation of implemented SDB algorithm

Several case studies have been carried out with various satellite imageries to evaluate the performance of the *i.image.bathymetry* in estimating SDB. Here, we present three case studies carried out in Puerto Rico, North-eastern Caribbean Sea, parts of Iwate and Miyagi prefecture, Japan. The first case study is carried out in the area demarcated by AoI-1 illustrated in Chapter Three and Figure 3.9, satellite data collected from Sentinel-2 (MSI) was used for SDB estimation and calibration depth points is same as explained in section 3.3. The characteristics of Sentinel-2 image used will be discussed in section 5.2.1. The second case study is carried out in parts of Miyagi prefecture, the study area and data used are exactly same as explained in Chapter Four and shown in Figure 4.1. The third study was carried out in parts of Iwate prefecture characteristics of the data and study area will be explained in the section 5.2.2.

5.2.1 Puerto Rico, North-eastern Caribbean Sea

Sentinel-2 high spatial and radiometric resolution data were collected on 25 December, 2015. As compared to other satellite imageries, Sentinel-2 have a spatial resolution that varies from 10m to 60m for multispectral bands and radiometric resolution quantized over a 12-bit dynamic range and rescaled to 16-bit. In addition to that, Sentinel-2 provides several red edge bands that also can be used as additional bands for SDB estimation. The availability of SWIR band (1.610µm) is another significant feature of Sentinel-2 which can be used for atmospheric and water corrections. A bi-linear interpolation has been carried out to resample the pixels of SWIR band from 20m to 10m. An interpolation method has been used to keep the original pixel value of SWIR as it is and interpolate only the redundant pixel by using the equation (17). One 20m pixel can be divided in to four 10m pixels, the value of first pixel was assigned as the original value of 20m pixel and values of other three pixels were interpolated using the equation (17).

$$Z = Z_1(1-t)(1-u) + Z_2u(1-t) + Z_3t(1-u) + Z_4ut \qquad (17)$$

Where, Z is the value of the pixel and x,y is the coordinates of Z in a matrix, respectively Z_1, Z_2, Z_3 and Z_4 are the neighboring pixels which are used for the calculation and (x_1,y_1), (x_1,y_2), (x_2,y_1), (x_2,y_2) are the coordinates respectively. t and u are the slope between these coordinates, can be written as $(x-x_1)/(x_2-x_1)$ and $(y-y_1)/(y_2-y_1)$, respectively.

5.2.2 Iwate, Japan

The third study area demonstrates SDB estimation in a coastal area of the Iwate prefecture, Japan (Figure 5.1). This study area is comparatively much smaller than the previous study area, covering only $4km^2$ and stretching along 6km of coastal line. Freely available Landsat-8 with medium spatial resolution (30m) and high radiometric resolution was collected on 31 October, 2014. All the available visible spectral bands and NIR band were used for estimation and SWIR band (1.57-1.65µm) band were used for correction.

5.3 System environment

The *i.image.bathymetry* module to estimate SDB has been developed in GRASS GIS Version 7. GRASS GIS is a robust Open Source GIS widely used in academia,

commercial settings and governmental agencies. A powerful feature of GRASS is the availability of Python scripting library which is used to implement several customizable modules for geoprocessing. In addition, GRASS GIS also provides access to R Statistical Computing through numerous packages for geospatial and geostatistical analysis.

5.3. GRASS Python scripting library

Python, a widely used programming language provides a powerful scripting environment in GRASS GIS. It enables users to efficiently exploit the capabilities of GRASS GIS software for developing new modules and extensions. In this study, GRASS Python scripting library has been used to combine multiple modules, functions from GRASS and R to implement the SDB model.

5.3.2 R packages

R is an Open Source statistical computing environment that provides several spatial analysis packages and functions. *rgrass7* work as an interface between GRASS GIS 7 and R. The package provides access to all GRASS commands from the R command line. The *rgrass7* package can be used to import/export data from GRASS to R and vice versa. The *rgrass7* package is only available beyond R Version 3.1, therefore, that should be installed in order to use the *i.image.bathymetry* module. Another library *data.table* is used to manage the spatial data frame of the raster data in R. *GWmodel* (Gollini *et al*., 2013) is collection of function which is considered as particular branch of spatial statistics called geographical weighted models and is used to deploy the GWR functionality.

5.4 Workflow of the module

The *i.image.bathymetry* is a collection of many existing GRASS GIS and R modules and new functions. Main functionalities of the module are 1) delineating water region, 2) atmospheric and water corrections, 3) GWR. Geometrically and radiometrically corrected spectral bands of any optical multispectral remote sensing data can be used to estimate SDB from the suitable coastal region. In the flowchart (Figure 5.2), the dotted box shows the required spectral bands, optional spectral bands and other input data such as calibration depth points (as vector point type) and tide height. The spectral bands in the green, red, NIR wavelength are mandatory inputs. Other spectral bands available in the

visible domain can be supplemented as additional bands depending upon the satellite sensor.

The first step in the processing chain of *i.image.bathymetry* is the delineation of water region. A rule based combination of NDVI and band ratio between green and infrared band were used for effective delineation. Band ratio of greater than or equal to 1 was classified as water, less than 1 as land. Masked water region has been used for further processing. NDVI has been used to remove the cloud, ice, etc., from the water region. GRASS GIS module *r.mapcalc* is used for these calculations to delineate water pixels effectively (Step 1 in Figure 5.2). In the practical experience it is observed that delineating water region without a visual interpretation is potential to produce error. Therefore, in addition to the band ratio (explained section 2.4.3) a NDVI is also used for water delineation.

In spite of the availability of numerous satellite images, very rarely satellite images are acquired at the time of zero tide. Therefore, *i.image.bathymetry* module has an option to provide tide height specific to the time of satellite image capture and to carry out tide correction (step 2 in Figure 5.2). Tide height of the time of image acquisition can be provided and used to correct the calibration depth. If the tide of the time of satellite image acquisition is lower than zero, negative value can be provided as tide height. This option can also be used when the same calibration depth is used in multi-temporal SDB estimation.

Theoretical background and implementation of atmospheric, water corrections (step3 in Figure 5.2) and GWR (step4 in Figure 5.2) algorithm are explained in section 2.5.3. Step 4a and 4b are illustrating the Fixed and A-GWR respectively. In case of Fixed-GWR, the size of the kernel is the same all over the working domain. Therefore, Fixed-GWR treats the entire region uniformly irrespective of the distribution of calibration depth points. Fixed-GWR model is less computationally intensive and less memory consuming as compared to an adaptive GWR (A-GWR) model. Consequently, Fixed-GWR model is available in many software packages and therefore it is easy to apply. In

GRASS GIS, the module *r.gwr* computes the Fixed-GWR and A-GWR with bi-square, and Gaussian kernel. The *i.image.bathymetry* module has adopted *r.gwr* module to process Fixed-GWR by selecting a system generated optimal bandwidth. The corrected spectral bands are used as independent variables and calibration depth as dependent variable to calculate Fixed-GWR (step 4a in Figure 5.2). A flag (-f) has been implemented in order to carry out depth estimation using Fixed-GWR.

In A-GWR model, size of the kernel is set by considering density of calibration points in a local neighborhood. The module called *GWmodel* available in R has been used to calculate A-GWR. Corrected spectral bands (step 3 in Figure 5.2) are imported to R using the library *rgrass7* and these spectral bands are further converted to a spatial data frame using the library *data.table* (step 4b in Figure 5.2). A spatial data frame of spectral bands is used as independent variable and reference depth is used as dependent variable to compute A-GWR using the library *GWmodel*. In *GWmodel*, the function called *bw.gwr* is used to calculate the optimal bandwidth. In *GWmodel*, the optimal bandwidth is determined using a cross-validation approach where the validation scores are minimized (Harris *et al*., 2010, Harris *et al*., 2011 and Lu *et al*., 2014). An A-GWR model is slower and needs more memory than a Fixed-GWR model. Therefore, *i.image.bathymetry* uses as default an A-GWR model. If the particular system not able to run the A-GWR model due to low memory, a Fixed-GWR will be used to estimate SDB (step 4b in Figure 5.2).

5.5 Results

Three case studies were carried out in order to assess the behavior and performance of the module related to different factors such as size of the data, spatial/spectral/radiometric resolution, and cloud coverage and water quality of the study area. Table 5.2 evaluates the performance of the module in order to comprehend impact of above mentioned factors in SDB estimation. Table 5.2 mainly compares the accuracy of the results in terms of kernel weighting functions (bi-square or Gaussian) and the mode of the GWR estimation (Fixed-GWR or A-GWR). The machine used for the benchmark was a laptop with an Intel Core i5-3320M CPU @2.60 GHz. The system has 16Gb of RAM and a solid state disk (SSD) of 512 Gb. The installed operating system (OS) is GNU/Linux

(Ubuntu 14.04 LTS x64-bit). GRASS 7.0.4 version and R version 3.3.1 were used for the benchmark. The screenshots shown in Figure 5.3 demonstrate options and workflow of *i.image.bathymetry*. Figure 5.3 used an example of case study 1 to demonstrate the processing using bi-square kernel. Figure 5.3a shows the required input used in the processing and Figure 5.3b demonstrates the optional input and optional flags used for Fixed-GWR based SDB estimation. The -f flag used in Figure 5.3b can be removed in order to process A-GWR based SDB estimation. Figure 5.3c demonstrates the optimal bandwidth estimation procedure for Fixed-GWR based computation. The optimal bandwidth estimation begins from a minimum value. The module performs GWR computation once the optimal bandwidth is determined. In Figure 5.3c, 47 was determined as optimal bandwidth for Fixed-GWR estimation. Figure 5.3d shows the A-GWR optimal bandwidth estimation in R using a cross-validation score. The number of points (15) selected as optimal adaptive bandwidth is shown in Figure 5.3d. Figure 5.3e and Figure 5.3f show the SDB estimated using Fixed-GWR and A-GWR in GRASS GIS monitor respectively.

The area covered (40 km^2) in case study 1 (AoI-1 in Figure 3.9) was relatively larger and spatial/spectral/ radiometric resolutions of the image used was higher. Apart from that, there were more depth points used for calibration and water quality of the area was also relatively better. Therefore, case study 1 produced high accuracy SDB from both Fixed and A-GWR models (Table 5.2). Processing time was evaluated using Fixed-GWR and A-GWR with bi-square and Gaussian kernel. In case of Fixed-GWR model, around 2 minutes were needed to finish processing and in case of A-GWR model, around 6 minutes were needed to finish the processing. In case study 1 (AOI-1), the study area was a continuous surface with dense calibration points distributed randomly over relatively clear waters, thereby, facilitating SDB generation with good accuracy SDB for both Fixed and A-GWR models. All the available bands in the visible domain and NIR band were used for SDB estimation as shown in Table 5.2. Tide height during satellite image acquisition was nearly zero; therefore no tide correction was applied. The estimation has been carried out using 1,260 depth points as dependent variable and another 2,000 depth

points were used to validate the results. Various weighting functions in Fixed-GWR and A-GWR were tested. Among different options attempted, A-GWR using bi-square kernel produced high accuracy SDB. Results were evaluated in terms of R (0.99), R^2 (0.98) and RMSE (0.61m).

Figure 4.1 shows the second study area with the parts of Miyagi prefecture which were used for ICRM generation and tsunami simulation. In case of Fixed-GWR model, around 3 minutes were taken to finish processing and in case of A-GWR model, around 260 minutes were taken to finish the processing (Table 5.2). The calibration depth points were sparse and distributed in unit interval (500m), and hence, Fixed-GWR model was expected to provide better or almost similar estimates like A-GWR model (Gollini, *et al*, 2015). Fixed-GWR with bi-square kernel provides better SDB. The results of SDB generated from Fixed-GWR shows better agreement with reference depth in terms of R (0.93), R^2 (0.87) and RMSE 1.65m than A-GWR model R (0.91), R^2 (0.84) and RMSE (1.95m)

The third study area covered ($4km^2$) a part of Iwate prefecture (Figure 5.1), was relatively smaller and SDB estimation was based on medium spatial resolution of 30m. However, spectral/radiometric resolution and number of depth points used for calibration of SDB were also relatively good. A total of 3,360 depth points (10 June, 2012) were collected by echo-sounder and 2,342 depth points were used for estimation and remaining depth points were used to evaluate the accuracy of the result. The characteristics of echo-sounder data are shown in Table 5.1. In case of Fixed-GWR model, around 2.5 minutes were taken to finish processing and in case of A-GWR model, around 180 minutes were taken to finish the processing (Table 5.2). A-GWR with bi-square kernel provides better SDB estimate compared to the other modes. In case study 3, a continuous surface of the study area and randomly distributed denser calibration depth points were considered as two favorable factors for better estimation using A-GWR model. The Fixed-GWR does not perform well in this area and could perhaps be the result of non-clear water in study area. Tide height during acquisition of satellite imagery was nearly 1.35m and was provided as tide height value in the module option to apply correction. Various weighting

functions in Fixed-GWR and A-GWR were tested. Among that A-GWR using bi-square kernel produced high accuracy SDB in terms of R (0.98), R^2 (0.97) and RMSE (1.50m). Detailed results of the SDB are shown in Table 5.2.

Chapter Six
Discussion and Conclusions

In this thesis available single spectral and multispectral band based SDB models were investigated and several improvements to the existing models in terms of estimation algorithms, and atmospheric and water corrections algorithms have been suggested. In addition, a GWR based SDB model have been developed and evaluated with application examples. Further, the developed GWR based SDB model has implemented in to an Open Source GIS environment therefore it can be used in any study areas without the need of additional software resource.

Firstly, this study investigated SDB model using single spectral band and further improved the available models by developing a new model for multi-temporal bathymetry. Further, multispectral approaches were carried out to estimate improved SDB. Three approaches such as RBE, MLR-NIR and new correction method using SWIR band were presented in the Second Chapter is originally proposed by this research in order to improve the accuracy of SDB. The RBE single band method is an extension to the SLR method. In the SLR method, attenuation coefficient assumed to be constant in the study area, to consider otherwise would require more ground truth knowledge about bottom type that are often not available. The attenuation of light in a particular band is estimated by using SLR method and multi-temporal images of same band was used for investigating bathymetric changes. RBE method depends only on radiance value of transformed band. More significantly, RBE shows a unique estimation capability for all the data sets. Since RBE directly converts transformed radiance to depth, it is very sensitive to correction methods. If the correction method is able to remove all the noises obtained from the atmosphere, water, then the RBE could provide far better results than occurred. The advantage of this RBE is that it is feasible even in situations where calibration data is lacking. Results derived by SLR and RBE models from various data sets with different spatial and radiometric resolution were evaluated using reference depth and observed good accuracy in quantitative terms. Landsat-8 data with high radiometric

resolution provides better results using SLR and RBE. Further, an attempt has also been carried out to comprehend the reduced accuracy of the SLR and RBE due to turbidity. Results show that turbidity affects the attenuation of the light and affects accuracy of derived depth but does not significantly degrade the estimation. The limitation of single band method is that, it does not address the issues of heterogeneity occurred due to difference in bottom type and water quality.

The MLR-NIR is a modification of multispectral band combination. The multispectral band models account for variations in attenuation coefficients depending on wavelength used, sea bottom types, and water column properties. Since NIR wavelength is capable of capturing some bottom reflectance in shallow water, a new multispectral band combination including NIR band has been proposed and evaluated. A comparative study was carried out between MLR and MLR-NIR. The results show that the MLR-NIR provides better SDB depth estimates as evident from improved correlation coefficient, coefficient of determination and RMSE.

The new correction method has been proposed using SWIR band to overcome the problems of NIR band correction. Study was observed as like many other previous authors, NIR band carries components of bottom reflectance in shallow water. Therefore, NIR band is not recommended for use in correction. However, this study is relying on the fact that there is no bottom reflectance component is contained in SWIR band (1.57-1.65µm). Therefore, SWIR band was used for atmospheric water correction. The comparative study has indicated that NIR band correction method provides null values due to high bottom reflectance in NIR region. The alternative correction method using SWIR provides depth information where NIR correction method was not able to provide any depth information. Thus, study recommends use SWIR band if available instead of NIR band for atmospheric and water correction for SDB estimation.

This study also proposed and evaluated the potential of GWR based SDB model in estimating near-shore bathymetry. Detailed investigation has been carried out in Puerto Rico (Figure 3.1) using Landsat-8 image and RapidEye image. In fact the proposed

GWR based SDB model is an extension of combined SDB model: the predictor can be analytically derived from a radiative transfer model. The results clearly demonstrate that GWR model is more suitable for SDB with enhanced reliability as compared to SLR and MLR (global model). It is also observed that conventional global models provide reasonable depth estimates, but do not sufficiently address the issues of heterogeneities that are normally observed in most areas where data is required. In addition, the global model suffers from over or under estimation of depth in many parts of the study area. The problem of over or under estimation are mainly attributed to difference in bottom type and water quality. On the other hand, GWR model can effectively address heterogeneity and able to significantly minimize such errors.

In GWR model, bandwidth is the crucial parameter and two mode of considering bandwidth is available one is Fixed-GWR and other is A-GWR. Many authors (e.g. Guo, *et al.* 2008) have suggested different mode of bandwidth (fixed vs adaptive) for better estimates. Therefore, apart from A-GWR model, we also investigated and evaluated the estimates generated from Fixed-GWR model in order to demonstrate the efficacy of A-GWR model. The comparative results shown in Table 3.3 clearly indicate that A-GWR model with Gaussian weighting function is the best option for estimating depth in this study area (Figure 3.1).

The main problem observed in A-GWR model is overshooting and undershooting when the bandwidth is too low. Larger bandwidth will produce a flat surface with little spatial variation (Yrigoyen *et al.,* 2008). Overshooting and undershooting at a few locations can be overcome by selecting optimal bandwidth with an appropriate weighting function. Many authors (Yrigoyen *et al.,* 2008, Su *et al.* 2014 and Monteys *et al.,* 2015) have suggested different weighting function in order to get better estimate in different situations. Some authors (e.g. Guo, *et al.,* 2008) have tried to comprehend the behavior of weighting functions and bandwidth in case of denser and sparser calibration points.

As demonstrated in this research, bandwidth is the crucial parameter in A-GWR model. Therefore, density of the calibration points impacts the results significantly. Two

cross-validation procedures were carried out prior and subsequent to the estimation, in order determine optimal weighting function and bandwidth. Since the study area is discontinuous surface with many islands, Gaussian weighting function was used to estimate optimal bandwidth. Based on the experiment, it was found that, optimal bandwidth with suitable weighting function could improve the estimation accuracy and address the heterogeneity effectively. It can also be concluded that the density of calibration points used for estimation is an influential factor for any kind of regression model. However, A-GWR model still produces reliable estimates, even with relatively few calibration points. The results indicate significant improvements in SDB estimation using GWR based model.

The importance of GWR based SDB model was demonstrated clearly and forms the main contribution of this thesis. GWR based SDB model is proposed as an effective solution to estimate SDB over near-shore where the conventional models perform poorly. Conventional SDB models provide good estimation when calibration depth points are dense. Unlike conventional model, GWR based SDB models provide reliable depth estimates even when calibration depth points are sparse. Therefore, the proposed model can be used to apply various situations such as low water quality and fewer calibration depth points.

As observed from this study SDB model was able to produce reliable results in various conditions in terms of spatial, radiometric resolution of satellite imagery and characteristics of calibration depth points. Even though the reliability of a GWR based model evaluated adequately, it is necessary to verify the estimated SDB to assist practical coastal modeling. The efficacy GWR based SDB model in practical scenario of tsunami simulation. GWR based SDB applied as one of the input for generation of ICRM and tsunami simulation. The seamless ICRM was developed by combining SRTM, SDB and GEBCO over parts of Miyagi prefecture in order to provide geometry polygon for ANUGA hydrodynamic model.

The tsunami simulation was carried out to hindcast the real tsunami occurred in

Tohoku region during 2011. The main earthquake shocks lasted for 3 to 4 minutes and, owing to the proximity of the epicenter to shore, the first significant waves reached Japan only 10 minutes after the event started (Grilli *et al.*, 2012). Coarse accuracy coastal bathymetry data were used for many previous tsunami simulation application researches. Several studies (e.g. Griffin *et al.*, 2016) have been reported issues pertaining to tsunami simulation due to lack of reliable accuracy coastal bathymetry data. Therefore, this study used SDB as an additional high resolution data set to reproduce the tsunami inundation extent and heights. Here, the inundation extent and inundation heights of the tsunami were compared with real data collected from the post-tsunami survey. These results were showing good agreement with post-tsunami survey of inundation extent and tsunami inundation height.

This research also presents a user friendly Open Source GIS module to estimate bathymetry from optical satellite imageries. The efficacy of the *i.image.bathymetry* module has been evaluated using three case studies with multi-resolution and multi-constellation data. The module incorporates pre-processing of satellite data and an option for performing tide correction of calibration depth data with tide at the time of satellite image capture. A limited number of depth points for calibration may be collected by LiDAR and SoNAR or extracted from the hydrographic charts. The SDB estimation region is automatically determined by the distribution of calibration depth points. Therefore it is suggested to be mindful while preparing the calibration depth points in order to cover the area of interest. In addition, an optional parameter has been developed to supply a polygon vector file which defines the area to be estimated. Deep water pixels are used for atmosphere and water corrections; therefore, users are suggested to supply the multispectral data which contains deep water region. Satellite image which does not have deep water region can also be used for SDB estimation; hence, the estimation will be carried out without using deep water coefficient for correction. In addition, Vinayaraj *et al.* (2016) has suggested the use SWIR band for correction instead of NIR band. NIR band is suggested to use for correction only when multispectral imagery do not include SWIR band such as in case of ASTER. The key functionality of the *i.image.bathymetry*

module is GWR based estimation. It has been demonstrated that GWR model can effectively address the issues of heterogeneity due to different bottom types and water quality and hence provide improved SDB.

The *i.image.bathymetry* module allows user to choose both Fixed and A-GWR model with kernel weighting function either bi-square or Gaussian, this option can be effectively used if the user has prior knowledge about the study area. This study also evaluated the computational cost for estimation Fixed and A-GWR models comparing with accuracy of the SDB generated. Such evaluation illustrates that A-GWR model is computationally intensive especially for large dataset but offers best results. However both Fixed and A-GWR model estimated SDB with acceptable accuracy and hence, it is suggested to use either Fixed or A-GWR according to the distribution of calibration depth points, size of the image and available RAM on computer.

The module can be downloaded from GRASS GIS svn repository[8] and installed using *g.extension* inside GRASS GIS. Accuracy of SDB depends upon characteristics of satellite imagery, number and distribution of calibration depth points and water quality of the study area. All the case studies carried out in the study were aimed to evaluate the performance of the module in addressing characteristics of the data which varied depending on data used and area investigated.

6.1 General conclusions

Concluding this chapter, it can be said that this study has highlighted and addressed a number of key problems in existing SDB models. Even though the problem of estimating bathymetry from remote sensing data has a relatively long history, last five to ten years the reliability of the SDB models have been significantly improved. Moreover, SDB has becoming popular in coastal researches. Study has investigated various problems in the SDB models mentioned by the previous authors and an attempt have been made to provide feasible solutions. The results were more favorable towards combined SDB model rather than simple empirical model. The aim of such combined models was to

[8] https://svn.osgeo.org/grass/grass-addons/grass72/imagery/i.image.bathymetry

provide better accuracy estimates and at the same time to be able justify the predictor using the radiative transfer model.

Two single band based SDB models and various multispectral band based SDB models were proposed in the thesis. Most of the models investigated in this study including single and multi-spectral bands were based on a linear regression approach. In case of single band, a simple linear regression (SLR) was used. Meanwhile, in case of multispectral bands, a multiple linear regression (MLR) has been used. RBE, the single band model was an extension of SLR model in order to estimate temporal SDB from multi-temporal satellite images without the need of multi-temporal calibration depth points. GWR based models were an extension of MLR model which was essentially a weighted linear regression based approach.

The SDB models described in this thesis was mainly intended to retrieve maximum bottom reflectance from a satellite image by addressing the issues of low water quality, different water types and bottom types. These were the three crucial problems that should be adequately investigated and addressed. The research was carried out step by step from simple models to advanced models in order to elucidate of these three problems. Among the SDB models, single band based model rely on the assumptions that attenuation coefficient is constant and bottom types are homogenous. This ideal situation rarely meets in complex coastal water region. Since the single band based models are only suitable for unique coastal waters with homogenous bottom type, a multispectral band based models investigated for more complex coastal waters. The correction method has improved upon using SWIR band to overcome the above of mentioned problems. MLR-NIR models have shown an improvement in accuracy of the estimates but do not satisfactorily address the issues related to variations in bottom type and water quality. Thus, study proposed a local linear regression using GWR based models. Consequently, the derived SDB showi better estimates and satisfactorily addresses the spatial heterogeneity. Study evaluated in detail about the different approach of GWR based model and can be used in various coastal waters with suitable GWR approach.

Since GWR based SDB was found to be addressing low quality water and heterogeneity issues, it is concluded that GWR based SDB model as a reasonable solution. Research further examined the efficacy of the estimated SDB in ICRM and tsunami simulation as a key application scenario. An Open source GIS module has been developed and evaluated in order to apply the proposed SDB model easily in other study areas around the world. Improved SDB demonstrated in this study has potential to supplement conventional charting and bathymetric surveys. Further, if a generalized predictor can be developed and applied to multi-temporal imagery for multi-temporal bathymetry changes.

6.2 Future perspectives

The results of this study suggest a number of new avenues for future research. Passive remote sensing based SDB techniques cannot be applied in regions with cloud cover or highly turbid waters. Therefore, in such situation Wave Based Approach (WBA) models to retrieve bathymetry from swell waves shoaling and refraction phenomena can be used. Moreover, the work presented here was using only satellite images as a source of bathymetry estimation. In addition to use only satellite images, research can use airborne UAV derived high resolution multi-spectral images. UAV-based approach can be effectively used to estimate multi-temporal SDB. Therefore the future perspective is to improve all aspect of the bathymetry estimation such as spatial resolution, spatial coverage and temporal resolution by using an integrated approach with passive satellite images, UAV images and SAR images. Either optical remote sensing or SAR remote sensing based methodology can be applied with respect to the coastal environment, and hence, we expect to produce coastal bathymetry even in a very complex coastal environment with wide coverage, which can really support further coastal research/application.

Another concern of the future study is to deploy *i.image.bathymetry* as a Web Processing Service (WPS) to facilitate on-demand SDB data to the user without requiring them to have software and data locally. In case of Japan, low resolution (500 m) depth data (JDOS) are available as Open Data. JDOS depth data coupled with satellite images

like Landsat-8, ASTER and Sentinel-2 could be used to SDB as a service to users with limited expertise in Remote Sensing and geospatial analysis. GRASS GIS also offers a framework for parallel computing which may be useful when considering future implementation of SDB algorithm as a Web Service.

References

Okyaso, A., Takenori, S., Hidekatsu Y., Takeyoshi, N. and Shinji, S. (2013) Severe erosion of sandbar at Unosumai River mouth, Iwate, due to 2011 Tohoku tsunami, *Coastal Dynamics*, pp.1311-1320.

Baban, S.M.J. (1993) The evaluation of different algorithms for bathymetric charting of lakes using landsat imagery. *International Journal of Remote Sensing* 14, pp.2263-73.

Benny, A.H., and Dawson. G. J. (1983) Satellite imagery as an aid to bathymetric charting in the Red Sea. *The Cartographic Journal* 20(1), pp.5-16.

Brunsdon, A.C., Fotheringham, S. and. Charlton, E.M. (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Analysis* 28 (2), pp.283-297.

Ceyhun, O., and Yalçın, A. (2010) Remote sensing of water depths in shallow waters via artificial neural networks. *Estuar. Coast. Shelf Sci* 89,pp.89-96.

Chen, G., Zhao, K., Mcdermid, G. J. and Hay. G. J. (2012) The influence of sampling density on geographically weighted regression: a case study using forest canopy height and optical data. *International Journal of Remote Sensing*, 33(9), pp.2909-2924

Clark, R. K., Fay, T. H. and Walker, C. L. (1987) Bathymetry calculations with Landsat-4-TM Imagery under a generalized ratio assumption. *Appl. Opt* 26(19), pp.4036-4038

Clark, R. K., Faye, T. H. and Walker, C. L. (1988) Bathymetry using thematic mapper imagery. *Proc. SPIE, Ocean Optics IX* 925, pp.229-231.

Costa, B.M., Battista, T. A. and Pittman, S. J. (2009) Comparative evaluation of airborne lidar and ship-based multibeam sonar bathymetry and intensity for mapping coral reef ecosystems. *Remote Sens. Environ* 113, pp.1082-1100.

Fonstad, M. A., and Marcus, W. A. (2005) Remote sensing of stream depths with hydraulically assisted bathymetry (HAB) models. *Geomorphology* 72, pp.320-339.

Fotheringham, A.S., Charlton, M. E. and Brunsdon, C. (1998) Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis, *Environ. Plann. A* 30 (11), pp.1905-1927.

Gao, J. (2009) Bathymetric mapping by means of remote sensing methods, accuracy and limitations. *Progress in Physical Geography* 33(1), pp.103-116.

George, D.G. (1997) Bathymetric mapping using a compact airborne spectrographic imager (CASI). *International Journal of Remote Sensing* 18, pp.2067-2071.

Gholamalifard, M., A. Esmaili S., Abkar, A. and Naimi, B. (2013) Bathymetric modeling from satellite imagery via single band algorithm (SBA) and principal components analysis (PCA) in southern caspian sea. *International Journal of Environmental Research* 7(4), pp.877-886.

Gollini, I., Lu, B., Charlton, M., Brunsdon, C. and Harris, P. (2013) GWmodel: an R Package for exploring spatial heterogeneity using geographically weighted models. *Journal of Statistical Software* 63(17), pp.1548-7660.

Goto, C., Ogawa, Y., Shuto, N. and Imamura, F. (1997). IUGG/IOC Time Project: Numerical Method of Tsunami Simulation with the Leap-Frog Scheme. Paris: IOC Manuals and Guides No. 35, UNESCO.

Griffin J. D., Latief, H., Kongko, W., Harig, S., Horspool, N., Hanung, R., and Rojali, A ., (2015) An evaluation of onshore digital elevation models for modeling tsunami inundation zones, *Frontiers in Earth Science*, 3(32), pp.1-16.

Griffin, J. D., Pranantyo, I. R., Kongko, W., Haunan, A., Robiana, R., Miller, V. (2016) Assessing Tsunami Hazard Using Heterogeneous Slip Models in the Mentawai Islands, Indonesia. *Geological Society of London Special Publication*.

Grilli, S.T., Harris,J.C., Tajalli Bakhsh, T.S., and Masterlark, T.L (2013) Numerical simulation of the 2011 Tohoku tsunami based on a new transient FEM co-seismic source: Comparison to farand near-field observations, *Appl. Geophys*., 170, pp.1333–1359

Guo, L., Ma, Z. and Zhang., L. (2008) Comparison of bandwidth selection in application of geographically weighted regression: a case study. *Canadian Journal of Forest Research* 38, pp.2526-2534.

Hamilton, M.K., and Davis., C.O. (1993) Estimating chlorophyll content and bathymetry of lake Tahoe using AVIRIS data. *Remote Sensing of Environment* 44(2), pp.217-230.

Harris, P., Fotheringham, A.S., Crespo, R., and Charlton. M. (2010) The use of geographically weighted regression for spatial prediction: an evaluation of models using simulated data Sets. *Math. Geosci* 42, pp.657-680.

Harris, P., C. Brunsdon, and A.S. Fotheringham. 2011. Links, comparisons and extensions of the geographically weighted regression model when used as a spatial predictor. *Stochast. Environ. Res. Risk Assess* 25, pp.123-138.

Hernandez, W.J., and Armstrong, R.A. (2016) Deriving bathymetry from multispectral remote sensing data. *Journal of Marine Science and Engineering* 4(8), pp.1-16.

Ibrahim, M., and Cracknell, A.P. (1990) Bathymetry using Landsat MSS data of penang island in Malayasia. *International Journal of Remote Sensing* 11, pp.557-559.

Jakeman, J., Nielsen, O., Van Putten, K., Mleczko, R., Burbidge, D. and Horspool, N. (2010) Towards spatially distributed quantitative assessment of tsunami inundation models. *Ocean Dyn*. 60, pp.1115–1138.

Jupp, D.L.B. (1988) Background and extensions to depth of penetration (DOP) mapping in shallow coastal Waters. *Symposium on Remote Sensing of the Coastal Zone*, Gold Coast, Queensland.

Kanno, A., Koibuchi, Y.T. and Isobe, M. (2011) Statistical combination of spatial interpolation and multispectral remote sensing for shallow water bathymetry. *IEEE Geoscience and Remote Sensing Letters* 8(1), pp.64-68.

Kanno, A., and Tanaka, Y. (2012). Modified lyzenga's method for estimating generalized coefficients of satellite-based predictor of shallow water depth. *IEEE Geoscience and Remote Sensing Letters* 9, pp.715-719.

Lee, Z., Carder, K.L., Mobley, C.D., Steward, R.G. and Patch, J.S. (1998) Hyperspectral remote sensing for shallow waters. 2. Deriving bottom depths and water properties by optimization. *Appl. Opt.* 38(18), pp.3831-3843.

Lu, B., Harris, P., Charlton, M. and Brunsdon, C. (2014) The GWmodel R package: further topics for exploring spatial heterogeneity using geographically weighted models. *Geo-spatial Information Science* 17(2), pp.85-101.

Lyzenga, D.R., (1978) Passive remote sensing techniques for mapping water Depth and

bottom features. *Applied Optics* 17 (3), pp.379-383.

Lyzenga, D.R., (1981) Remote sensing of bottom reflectance and water attenuation parameters in shallow water using aircraft and Landsat data. *International Journal of Remote Sensing* 2, pp.72–82.

Lyzenga, D.R., Malinas, N.R. and Tanis, F.J. (2006) Multispectral bathymetry using a simple physically based algorithm. *IEEE Trans. Geosci. Remote Sensing* 44(8), pp.2251-2259.

Manessa, M.D.M., Kanno, A., Sekine, M., Ampou, E.E., Widagti, N. and As-syakur, A.R. (2014) Shallow-water benthic identification using multispectral satellite imagery: Investigation on the effects of improving noise correction method and spectral cover. *Remote Sensing* 6, pp.4454-4472.

Mobley, C. D. (2001) Radiative transfer in the ocean. *Scientific, Inc*., WA, USA, pp.2321-2330

Monteys, X., Harris, P., Caloca, S. and Cahalane, C. (2015) Spatial prediction of coastal bathymetry based on multispectral satellite imagery and multibeam data. *Remote Sens*. 7, pp.13782-13806.

Mori, N. and Takahashi, T. (2012) Nationwide post event survey and analysis of the 2011 Tohoku earthquake tsunami. *Coastal Engineering Journal*, 54, pp.1250001-27

Muslim, A.M., and Foody, G.M. (2008) DEM and bathymetry estimation for mapping a tide-coordinated shoreline from fine spatial resolution satellite sensor imagery. *International Journal of Remote Sensing* 29, pp.4515-4536.

Nielsen, O., Roberts, S., Gray, D., McPherson, A. and Hitchman, A. (2005) Hydrodynamic modelling of coastal inundation in MODSIM (Melbourne, VIC).

Pacheco, A., Horta, J., Loureiro, C. and Ferreira, O. (2015) Retrieval of nearshore bathymetry from Landsat-8 Images: a tool for coastal monitoring in shallow waters. *Remote Sensing of Environment* 159, pp.102-116.

Philpot, W.D. (1989) Bathymetric mapping with passive multispectral imagery. *Applied Optics* 28(8), pp.1569-1579.

Rakowsky, N., Androsov, A., Fuchs, A., Harig, S., Immerz, A., Danilov, S. (2013)

Operational tsunami modelling with TsunAWI—recent developments and applications. *Nat. Hazards Earth Sys. Sci.*, 13, pp.1629-1642.

Ribeiro, S. R. A., Centeno, J. A. S. and Krueger, C. P. (2008) An estimate of depth from a bathymetric survey and IKONOS II data by means of artificial neural network. *Bol. Ciênc. Geod*., 14, pp.171–185

Solomon C. Y., Kwok, F.C., Michael J.O. and Yoshiki Y. (2012) Tohoku tsunami survey, modeling and probabilistic load estimation applications proceedings of the International Symposium on Engineering Lessons Learned from the 2011 Great East Japan Earthquake, Tokyo, Japan, pp.430-443

Stoffle, R.W., and Halmo, D.B. (1991) Satellite monitoring of coastal marine ecosystems: a case from the dominican republic. *consortium for integrated earth science information network (CIESIN)*. Saginaw, Michigan.

Stumpf, R.P., Holderied, K. and Sinclair, M. (2003) Determination of water depth with high-resolution satellite imagery over variable bottom types. *Limnol. Oceanogr* 48(2), pp.547-556.

Su, H., Liu, H., Lei, W., Philipi, M., Heyman, W. and Beck. A. (2013) Geographically adaptive inversion model for improving bathymetric retrieval from multispectral satellite imagery. *IEEE Transaction on Geosciences and Remote Sensing* 52(1), pp.465-476.

Tripathi, N.K. and Rao, A.M. (2002) Bathymetric mapping in kakinada bay, India, using IRS- 1D LISS-III Data. *International Journal of Remote Sensing* 23(6), pp.1013-1025.

Van Hengel, W., and Spitzer. D. (1991) Multi-temporal water depth mapping by means of Landsat TM. *International Journal of Remote Sensing* 12(4), pp.703-712.

Vinayaraj, P., Raghavan, V., Masumoto, S. and Glejin. J. (2015) Comparative evaluation and refinement of algorithm for water depth estimation using medium resolution remote sensing data. *International Journal of Geoinformatic* 11(3), pp.17-29.

Vinayaraj, P., Raghavan, V. and Masumoto, S. (2016) Satellite derived bathymetry using adaptive-geographically weighted regression model, *Marine Geodesy*, 39(6), pp.458-478.

Warne, D.K., 1972. Landsat as an aid in the preparation of hydrographic charts. *Photogrammetric Engineering and Remote Sensing* 44, pp.1011-1016.

Yi, G., and Li, T. (1988) The ocean information contents of remotely sensed image and the acquisition of water depth message. *Proceedings of the Ninth Asian Conference on Remote Sensing, Bangkok, Thailand* (Asian Association of Remote Sensing, University of Tokyo, Japan) F-3-1–F-3-8.

Yrigoyen, C.C., Rodrigouz, I.G. and Otera. J.V. (2008) Modeling spatial variations in household disposable income with geographically weighted regression. *Estadística española* 50(168), pp.321-360.

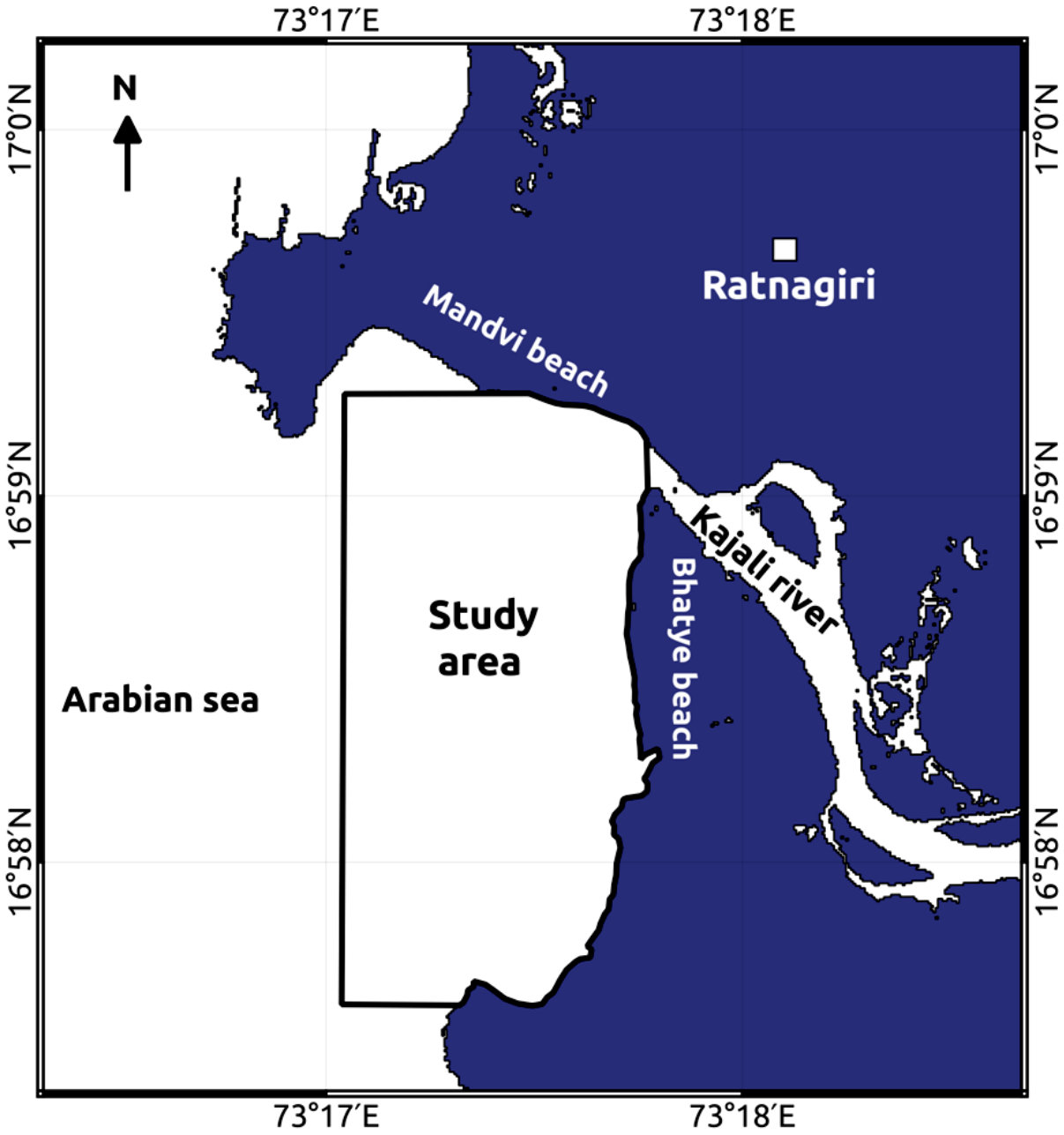


Figure 2.1: Study area, Ratnagiri, India (after Vinayaraj *et al.*, 2015)

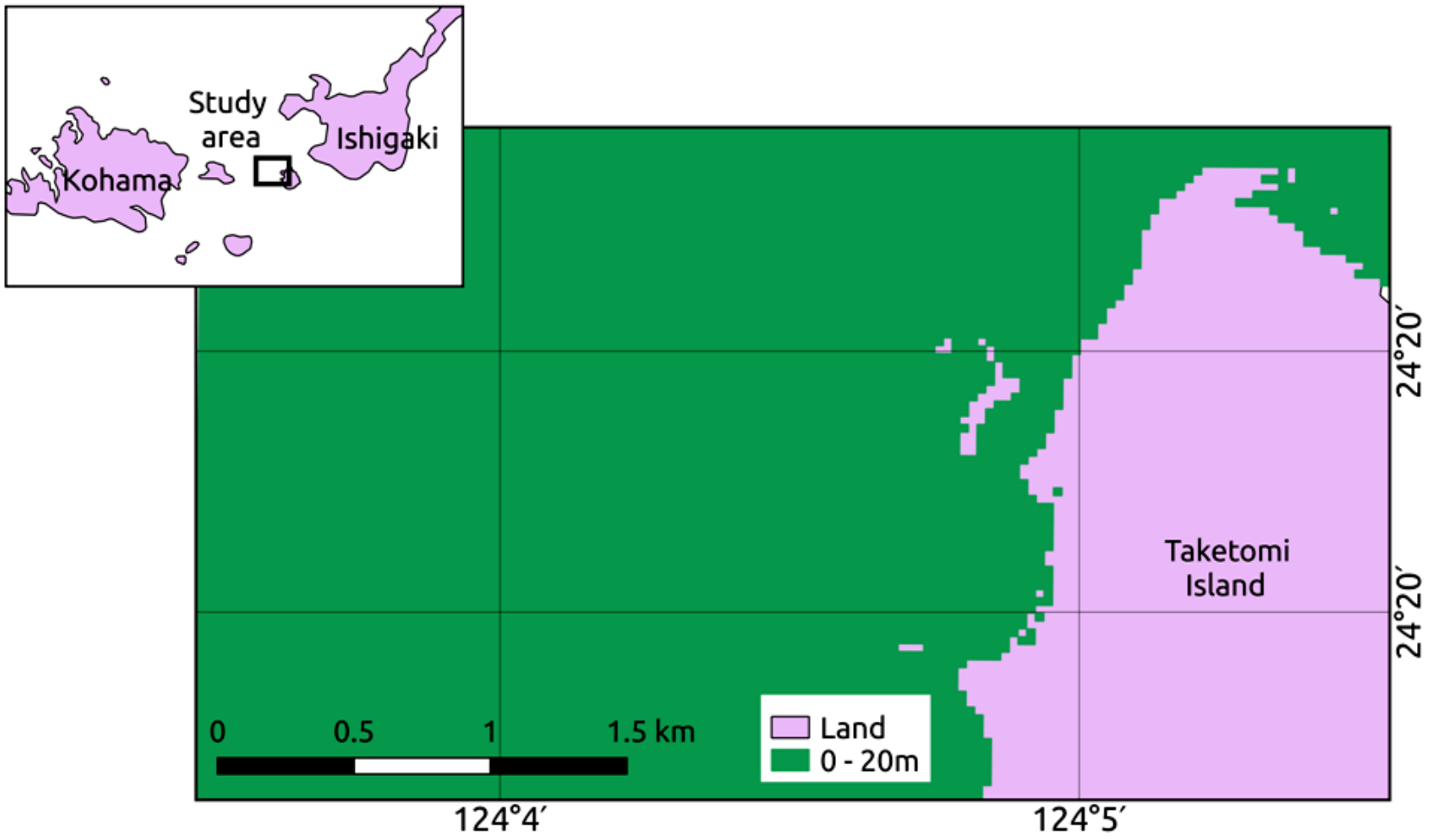


Figure 2.2: Study area, Taketomi Island, Japan

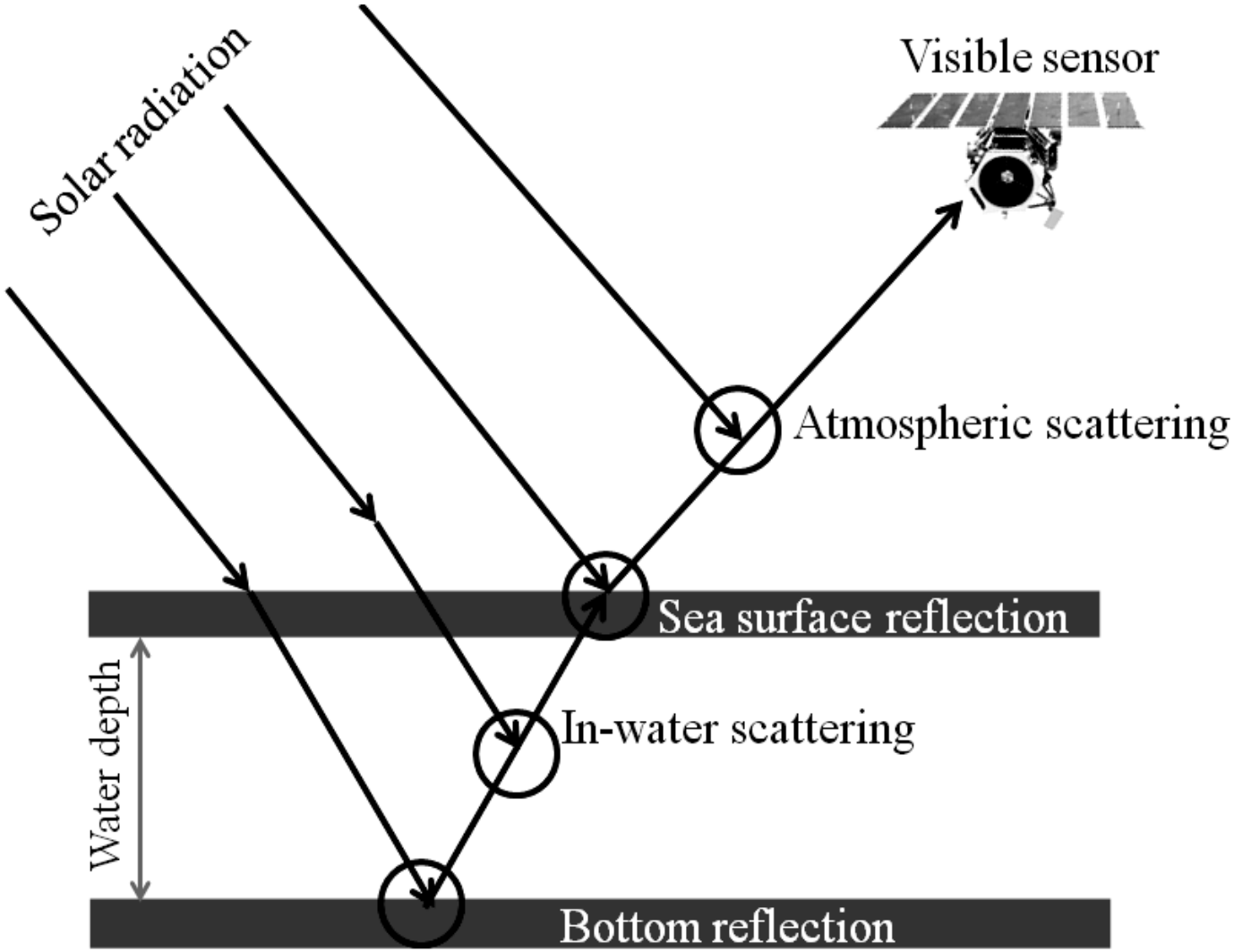


Figure 2.3: Schematic view of spectral radiance components observed using the visible sensor over optically shallow water (after Vinayaraj *et al*., 2016)

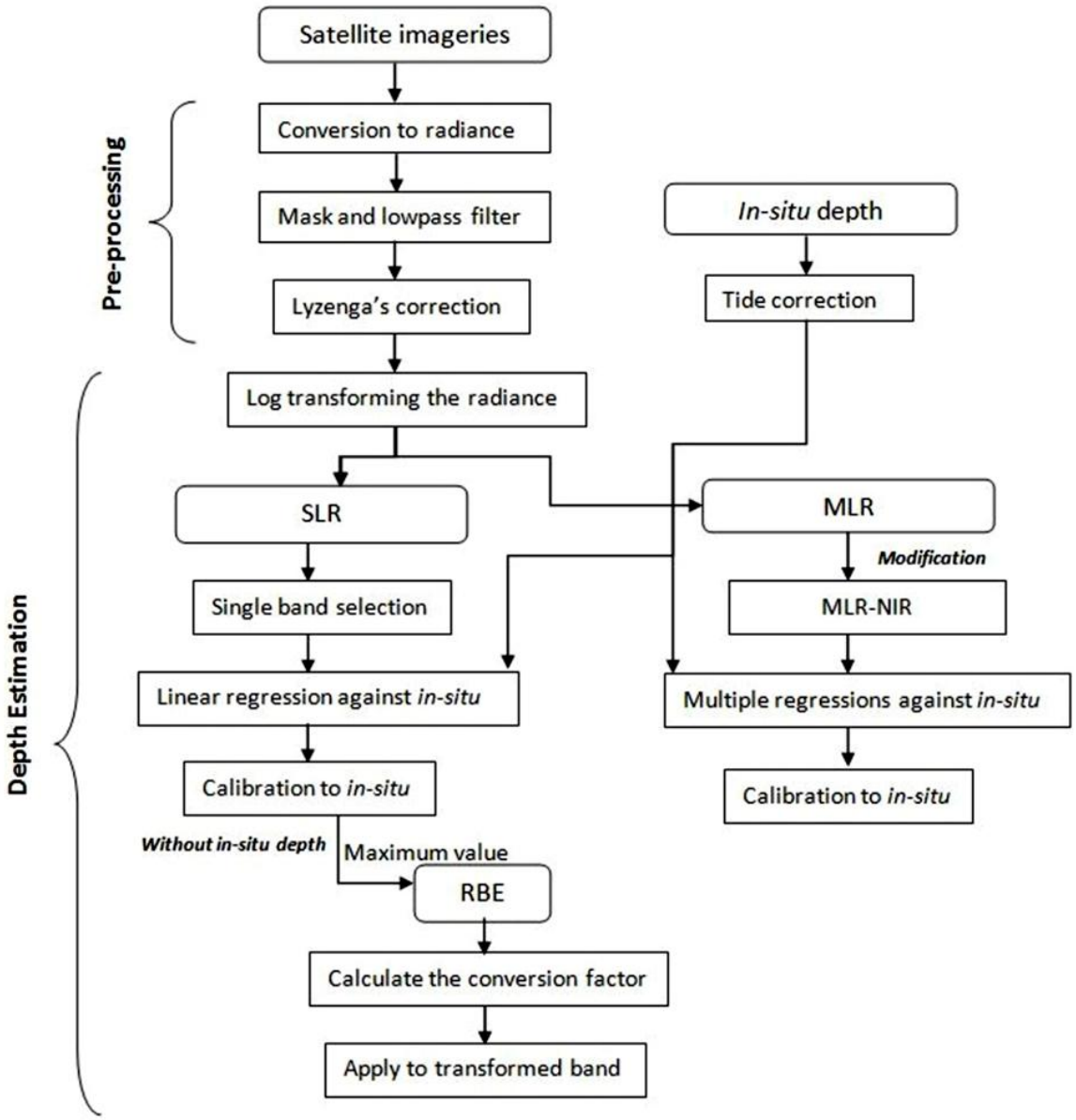


Figure 2.4: Flowchart of proposed models (after Vinayaraj *et al*., 2015)

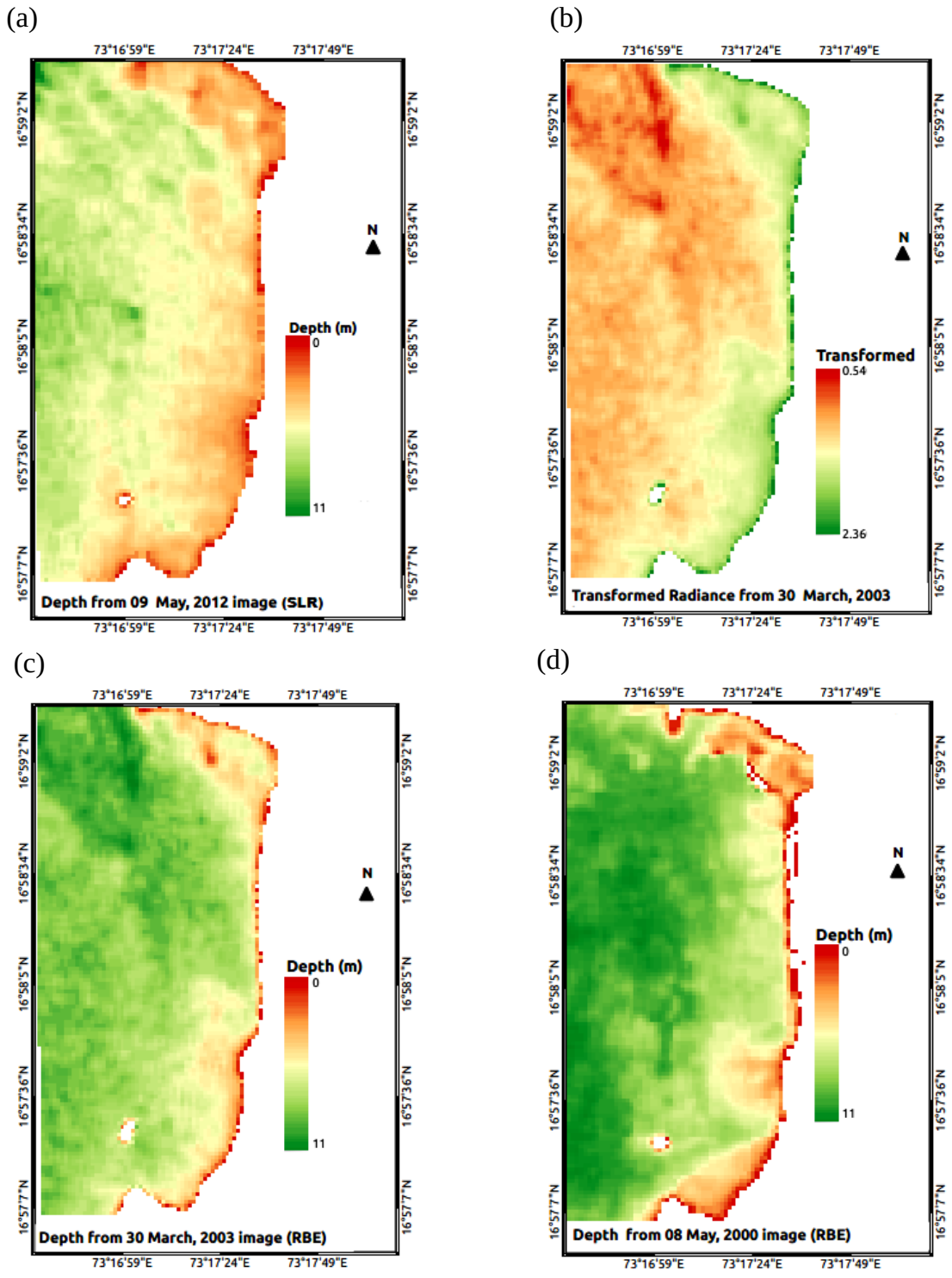


Figure 2.5: SDB by SLR method from Landsat-7, (a) depth for 09 May, 2012, (b) Transformed radiance of NIR band of 30 March, 2003, (c) SDB for 30 March, 2003 and (d) SDB for 08 May, 2000 (after Vinayaraj *et al.*, 2015)

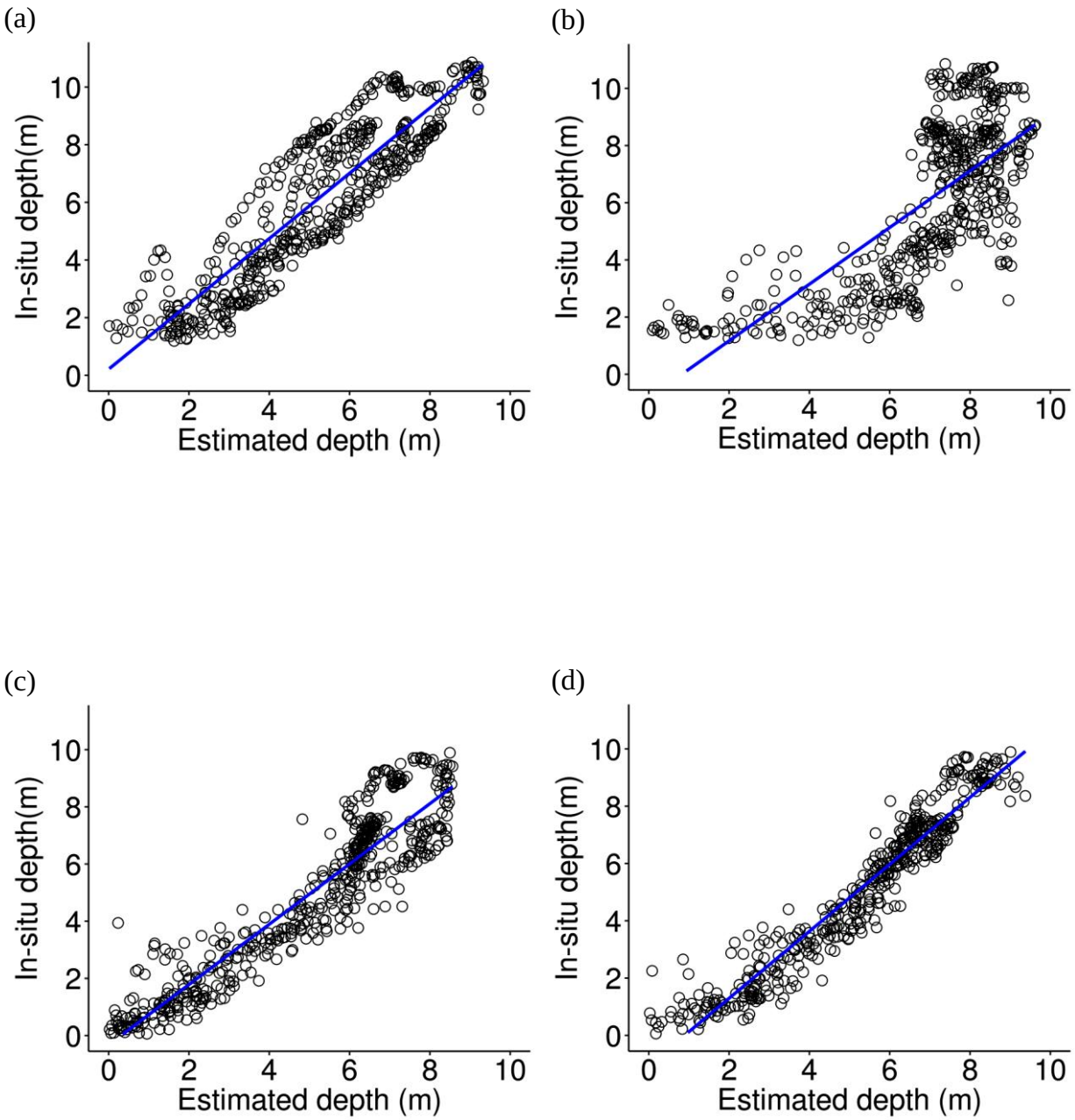


Figure 2.6: Bivariate Scatter plot between SDB from Landsat-8 and reference depth
RBE, (a) 31 Jan, 2014; (b) 04 Mar, 2014
MLR, (c) 12 Nov, 2013; (e) 31 Jan, 2014; (g) 04 Mar, 2014
MLR-NIR, (d) 12 Nov, 2013; (f) 31 Jan, 2014; (h) 04 Mar, 2014 (continued).

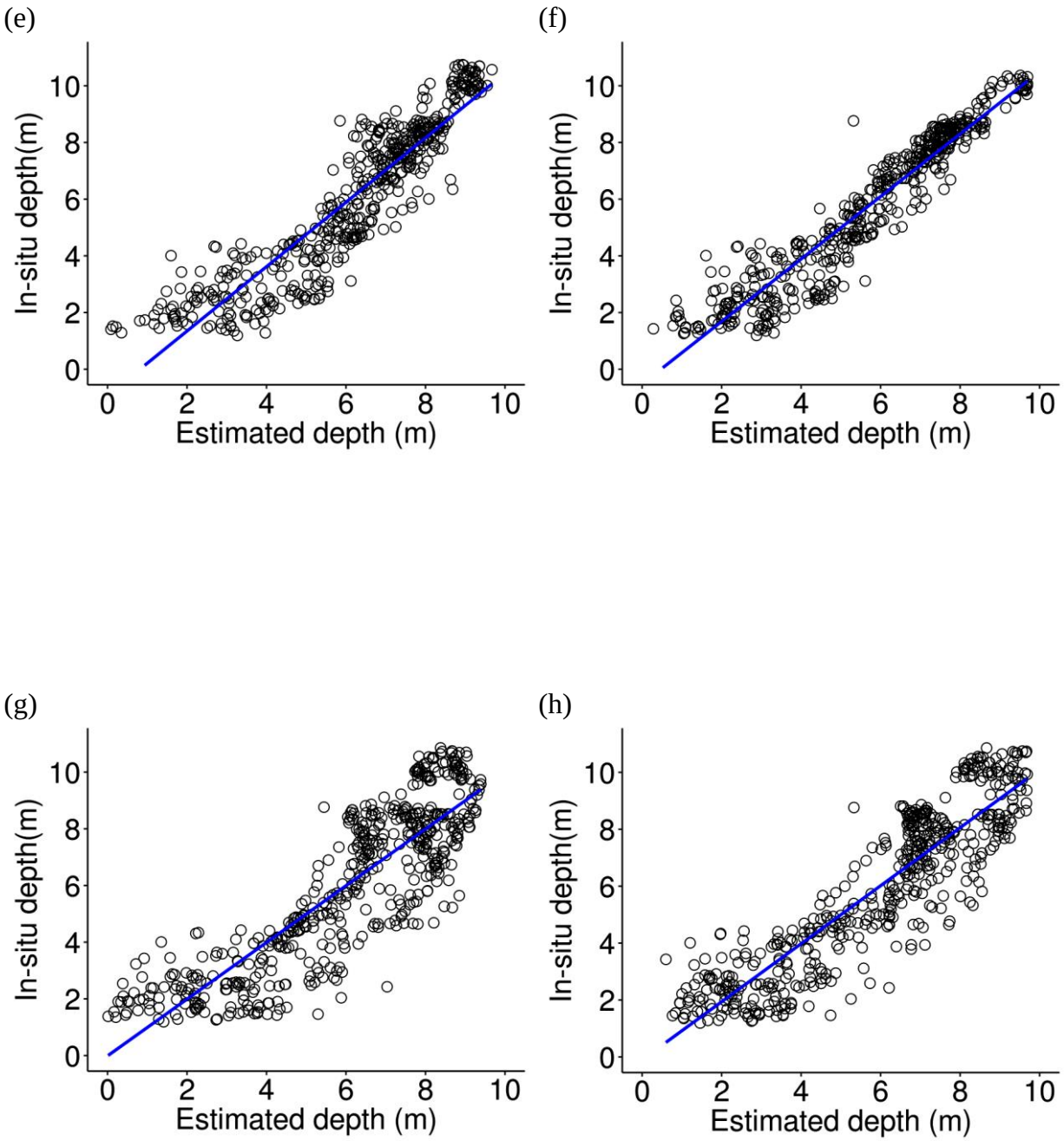


Figure 2.6: Bivariate scatter plot between SDB from Landsat-8 and reference depth.
RBE, (a) 31 Jan, 2014; (b) 04 Mar, 2014
MLR, (c) 12 Nov, 2013; (e) 31 Jan, 2014; (g) 04 Mar, 2014
MLR-NIR, (d) 12 Nov, 2013; (f) 31 Jan, 2014; (h) 04 Mar, 2014

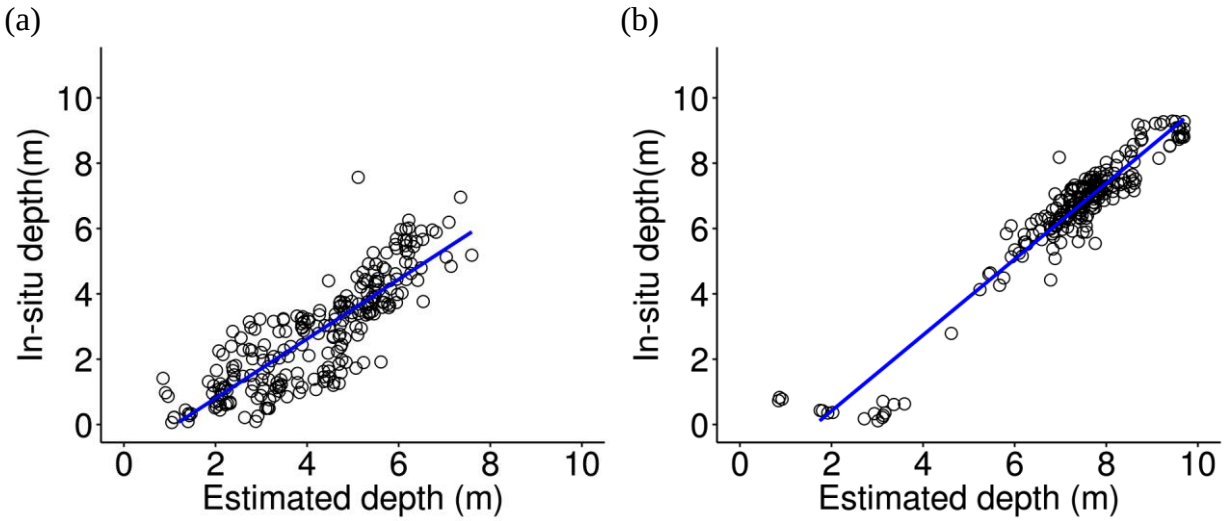


Figure 2.7: Bivariate scatter plot of SDB from Landsat-8 (31 Jan, 2014) in terms turbidity (a) at high turbid (b) and low turbid region (after Vinayaraj et al., 2015)

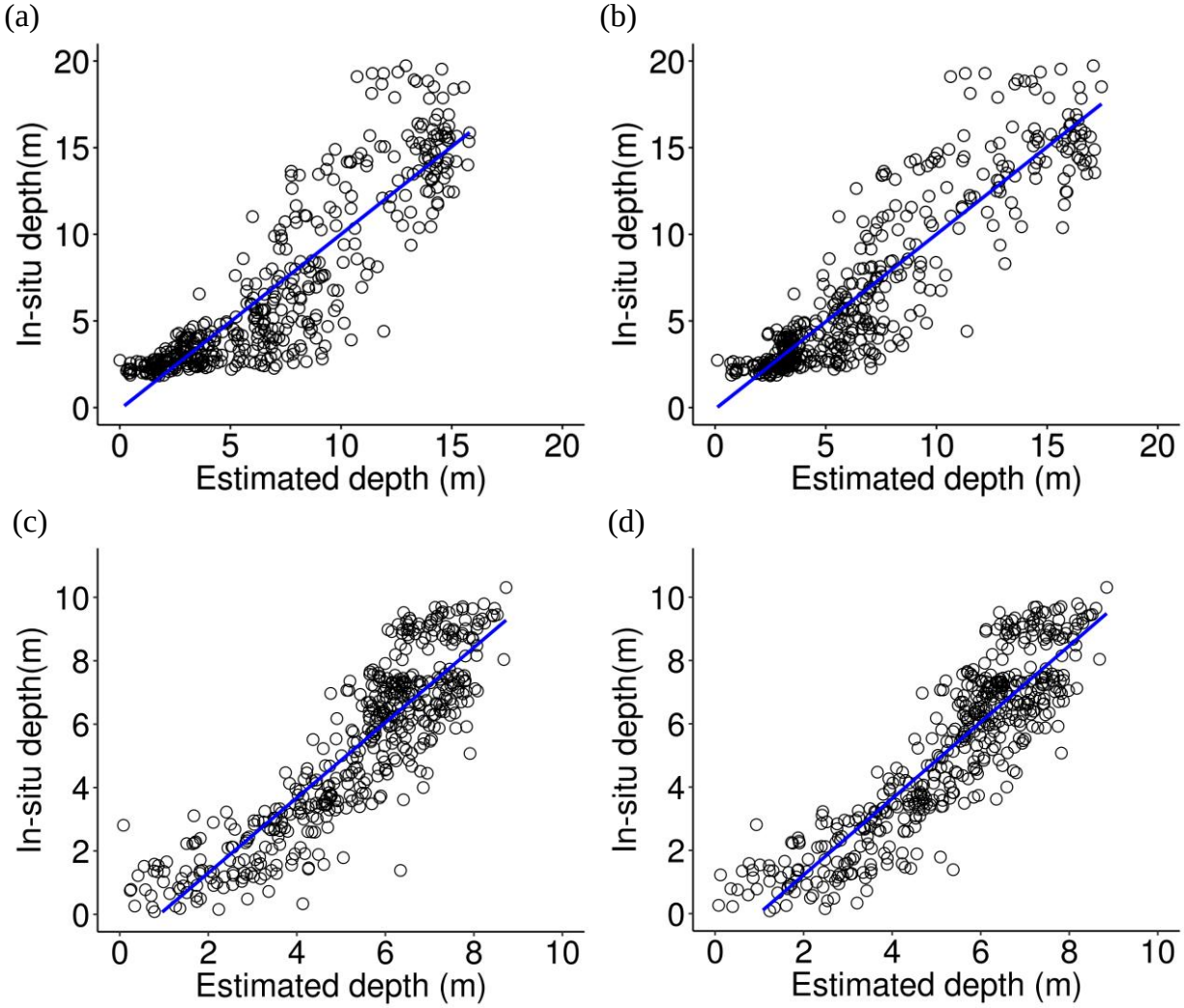


Figure 2.8: Bivariate scatter plots between reference depth and SDB estimated with NIR and SWIR corrections at (a), (b): Taketomi and (c), (d): Ratnagiri, respectively.

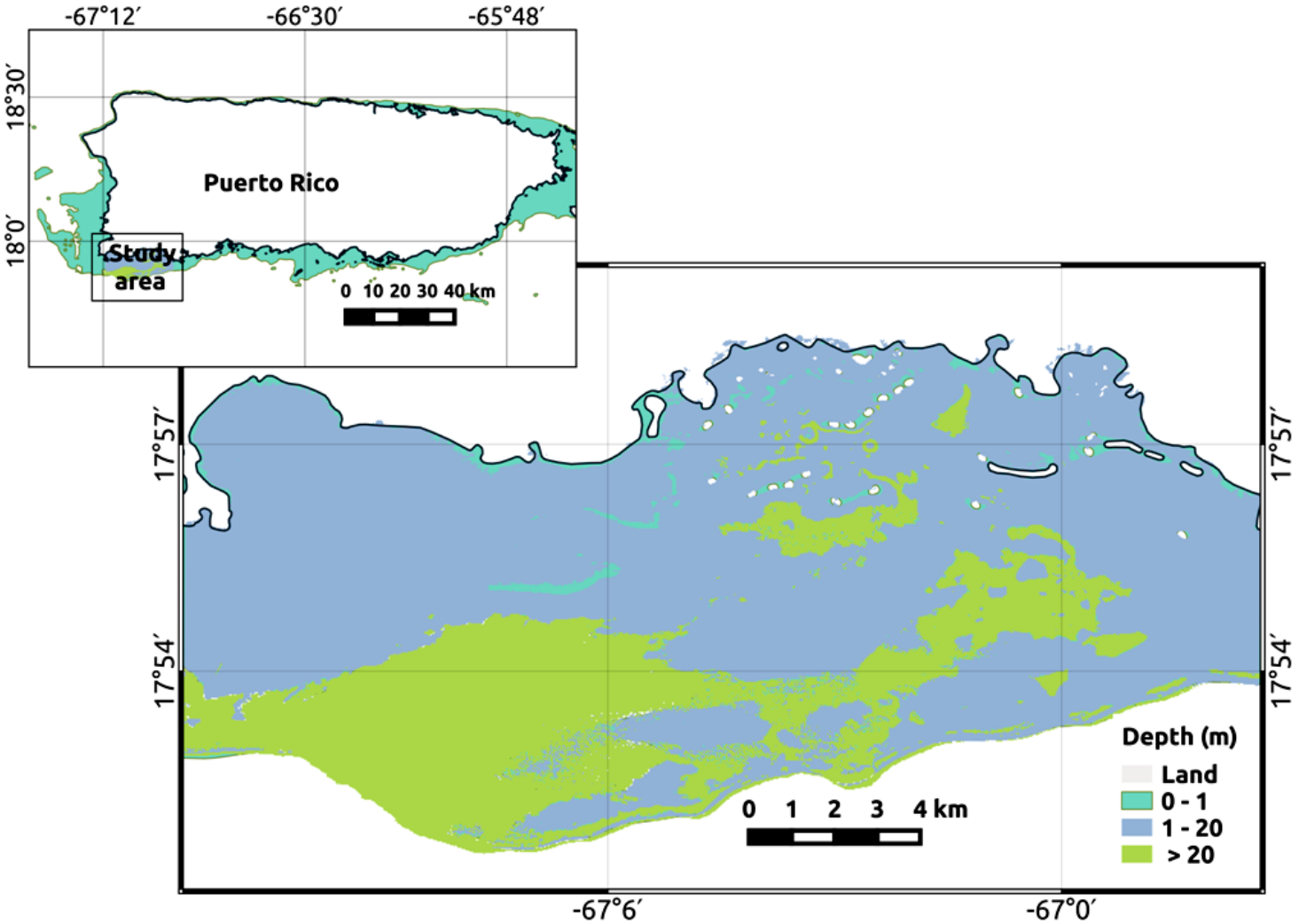


Figure 3.1: Study area used in South-western Puerto Rico (after Vinayaraj *et al.*, 2016)

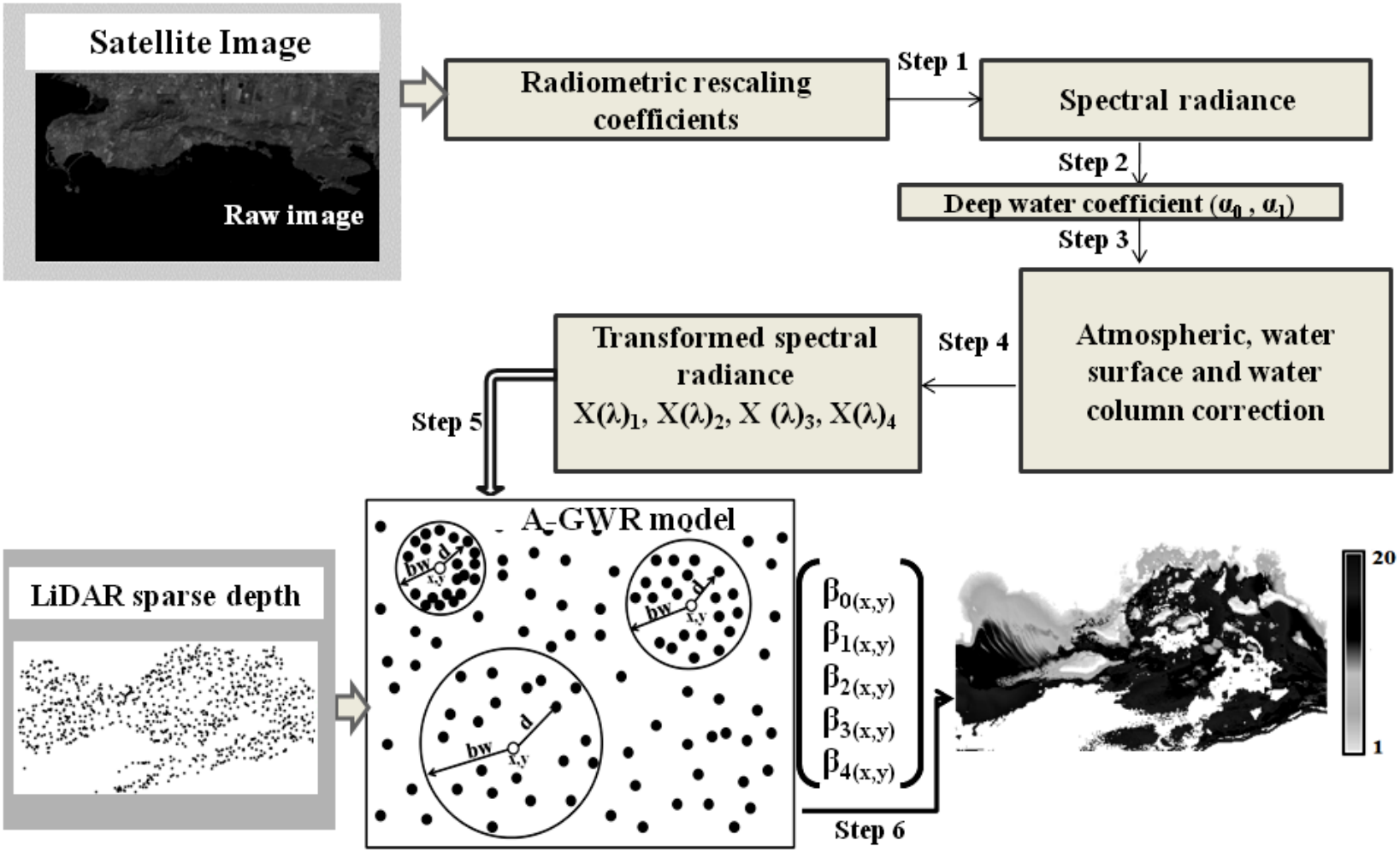


Figure 3.2: Workflow of the A-GWR model (after Vinayaraj *et al*., 2015)

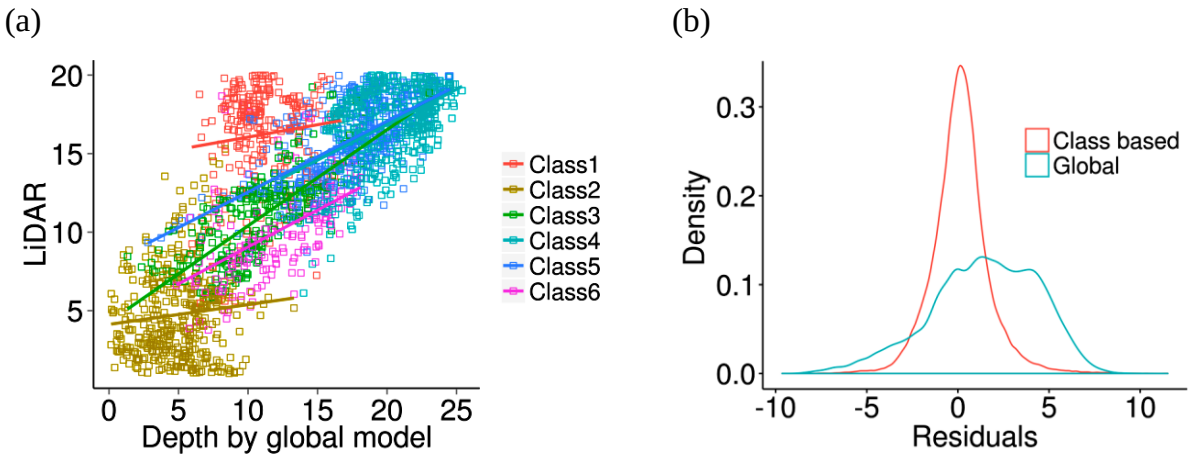


Figure 3.3: (a)Trend lines of scatter plots of SDB (Landsat-8) from global model and LiDAR depth for different classes; (b) and density plot of differences between SDB and LiDAR depth (after Vinayaraj *et al.*, 2016)

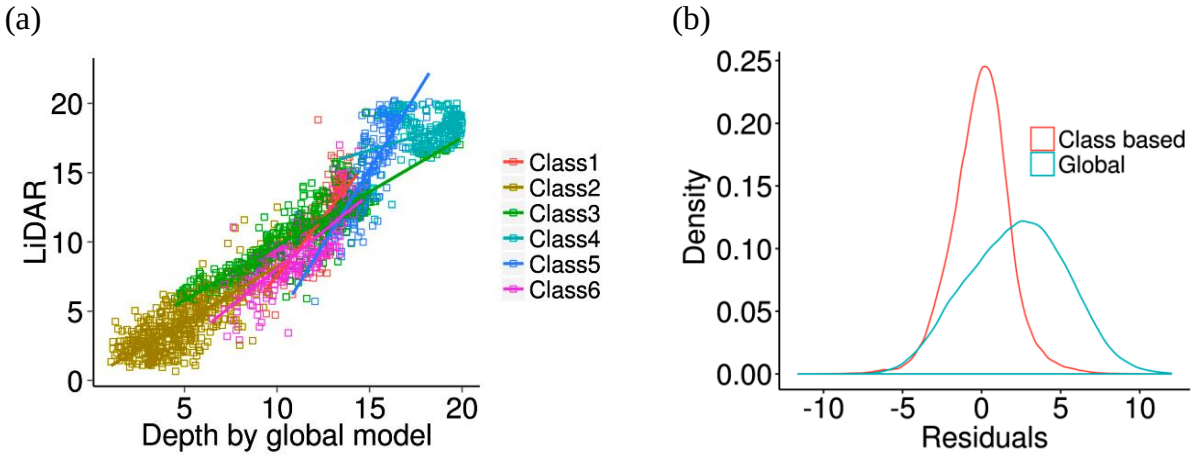


Figure 3.4: (a)Trend lines of scatter plots of SDB (RapidEye) from global model and LiDAR depth for different classes; (b) and density plot of differences between SDB and LiDAR depth (after Vinayaraj *et al.*, 2016)

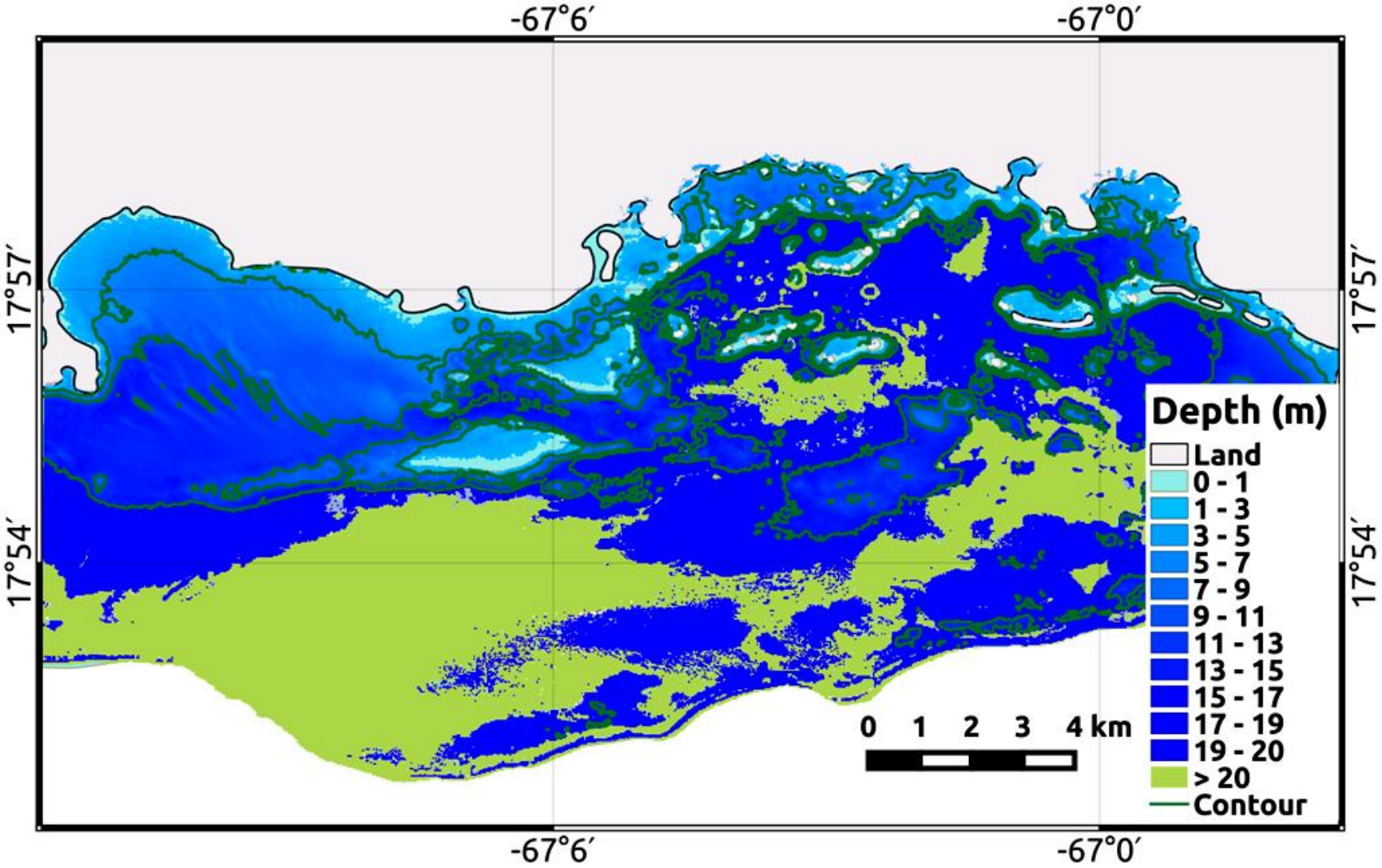


Figure 3.5: SDB from Landsat-8 using 10,000 LiDAR depth points as calibration depth (after Vinayaraj *et al*., 2016)

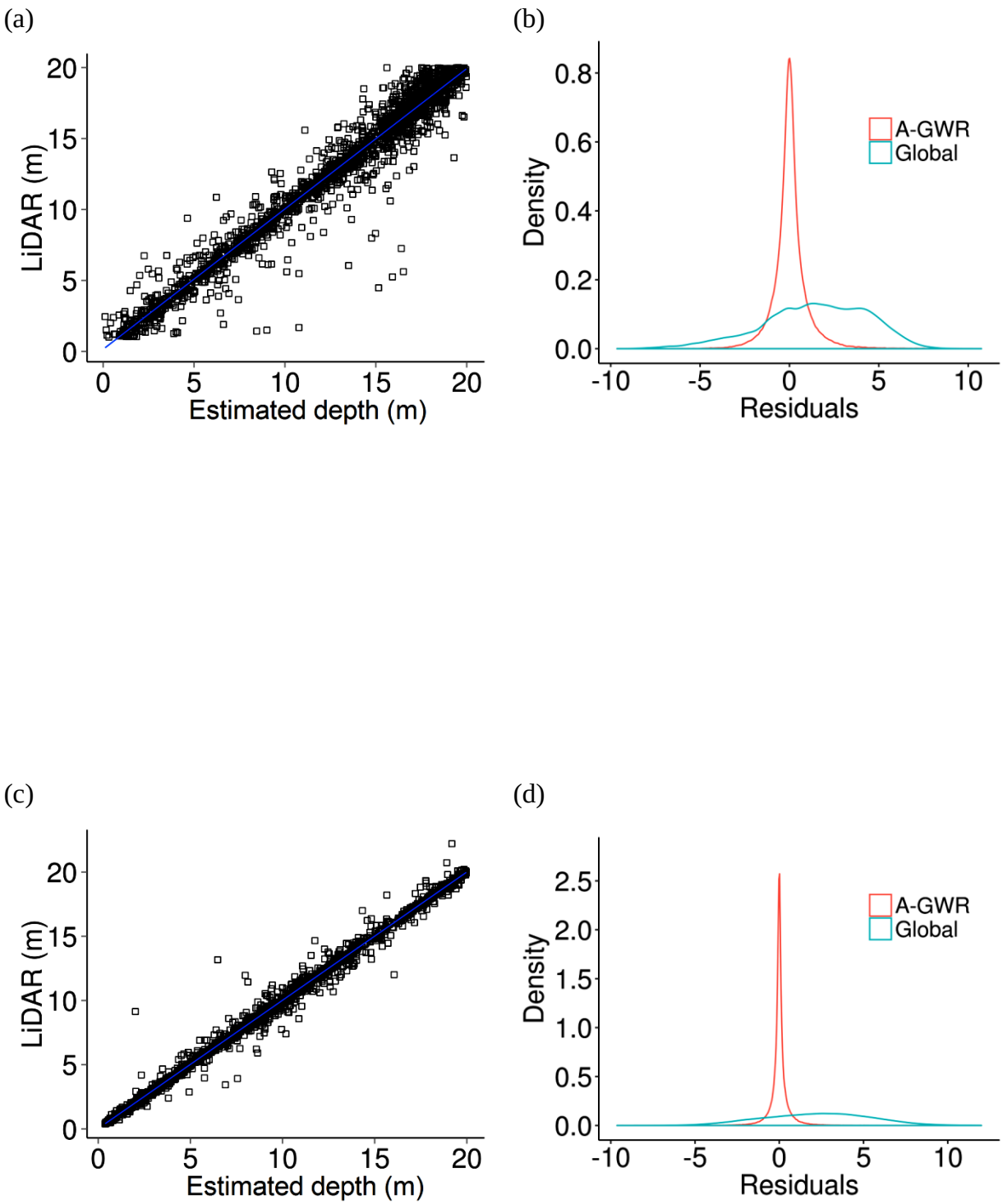


Figure 3.6: Evaluation and comparison of SDB from Landsat-8 (a), (b) and RapidEye (c), (d) images by A-GWR model (after Vinayaraj *et al*., 2016)

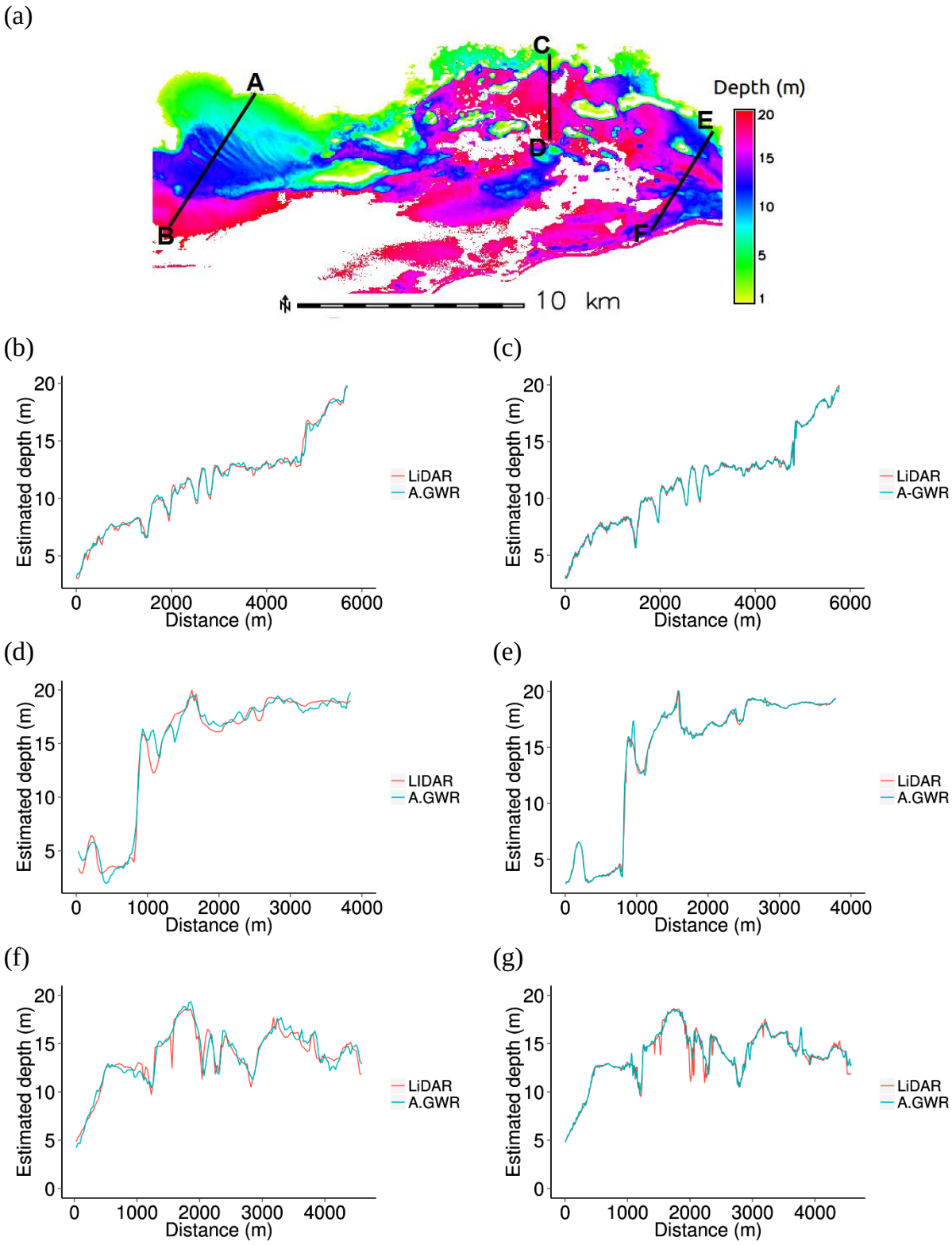


Figure 3.7: Evaluation of SDB from Landsat-8 (b, d and f) and RapidEye (c, e and g) images by A-GWR model using random cross profiles (after Vinayaraj *et al.*, 2016)

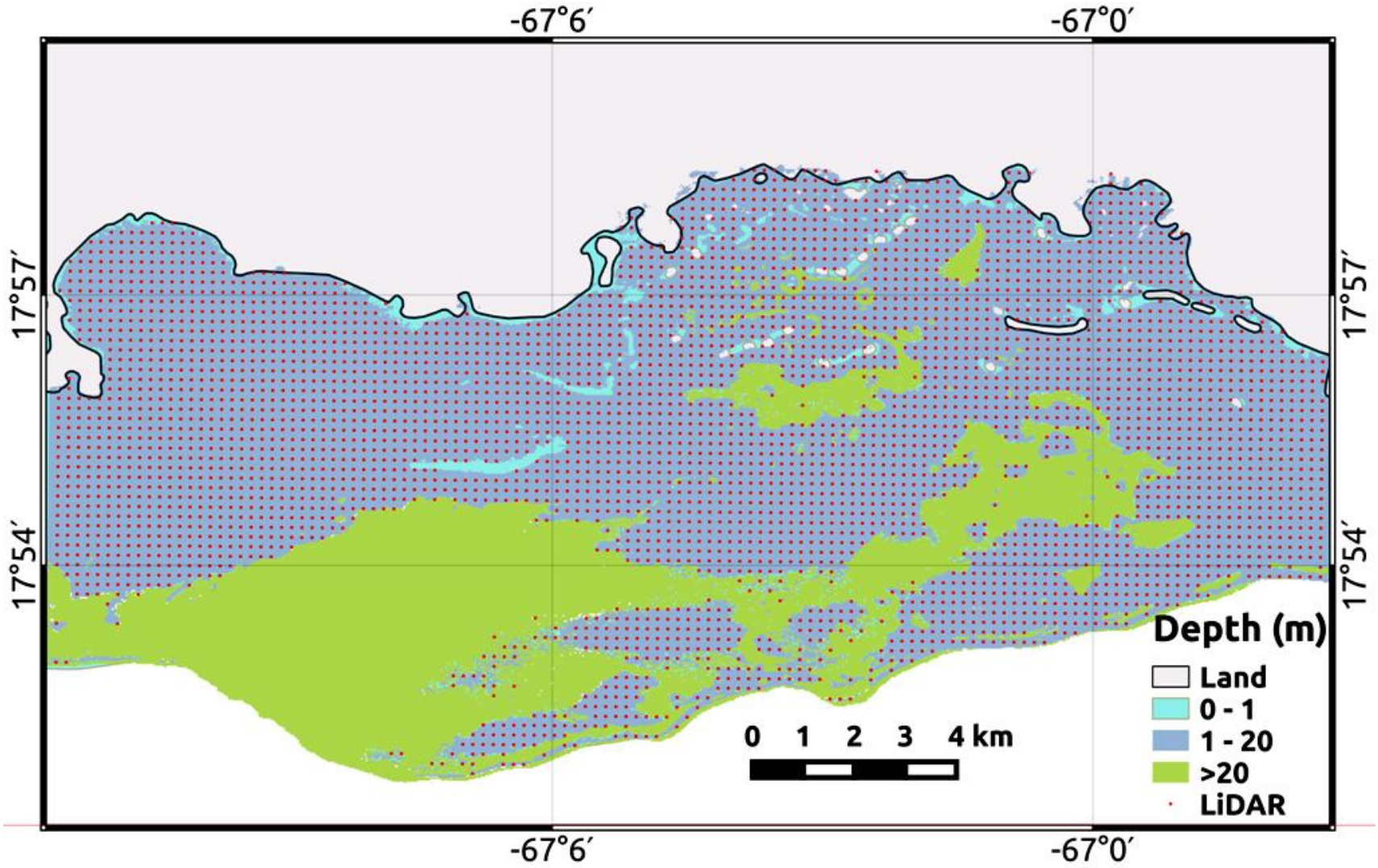


Figure 3.8: Demonstration of LiDAR depth points distribution as in scenario-1

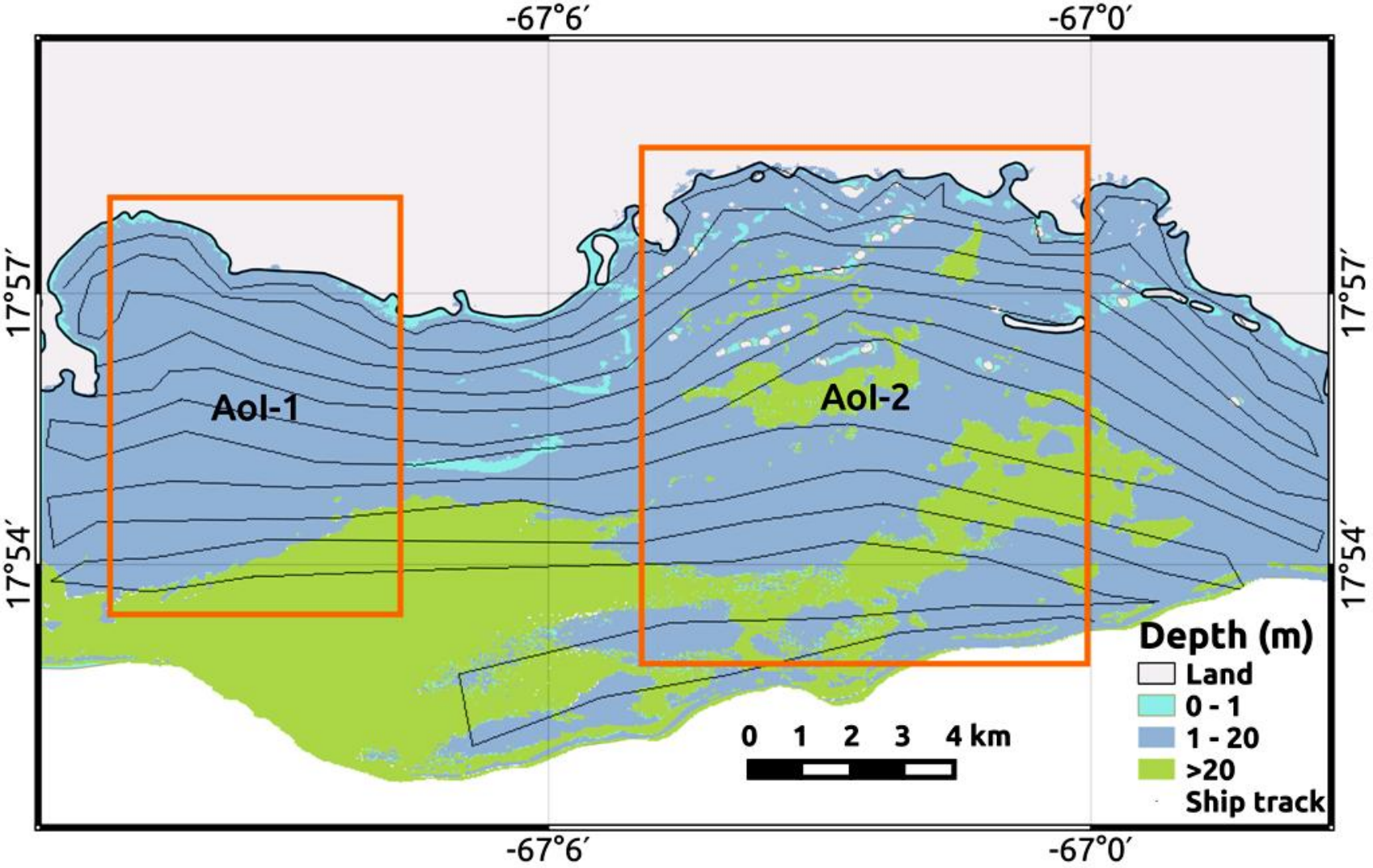


Figure 3.9: Demonstrate the arbitrary ship tracks for SDB in scenario-2 (after Vinayaraj *et al*., 2016)

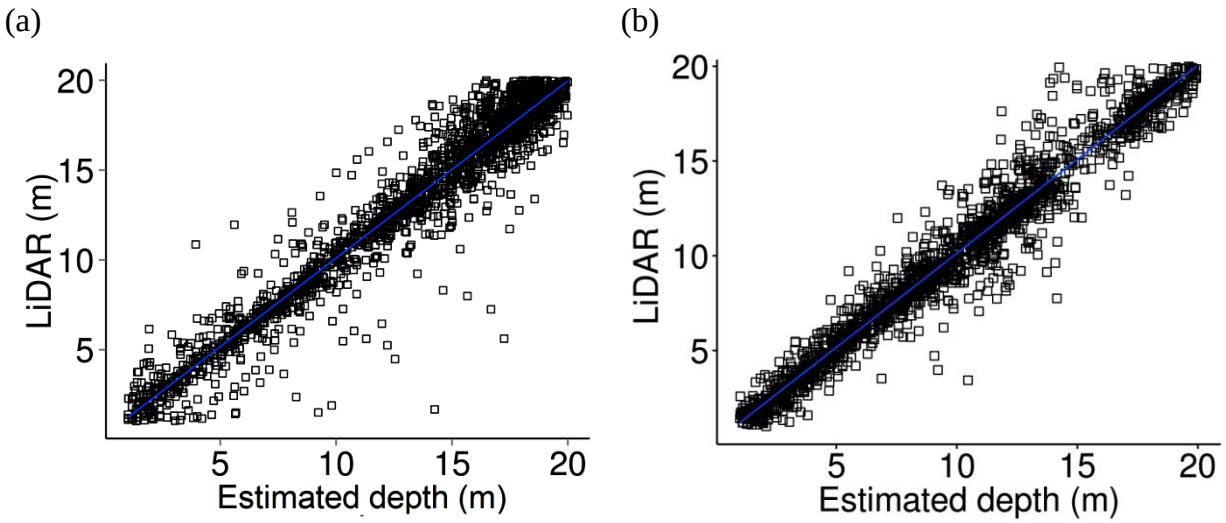


Figure 3.10: Evaluation of SDB from Landsat-8 (a) and RapidEye (b) images with LiDAR depth for Scenario-1 using A-GWR model (after Vinayaraj *et al*., 2016)

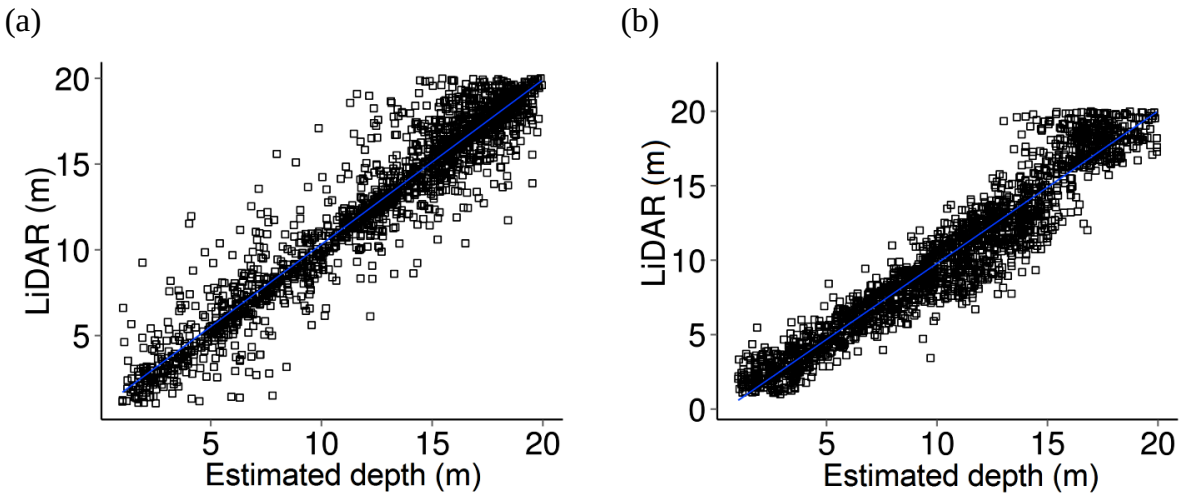


Figure 3.11: Evaluation of SDB from Landsat-8 (a) and RapidEye (b) images with LiDAR depth for Scenario-2 using A-GWR model (after Vinayaraj *et al*., 2016)

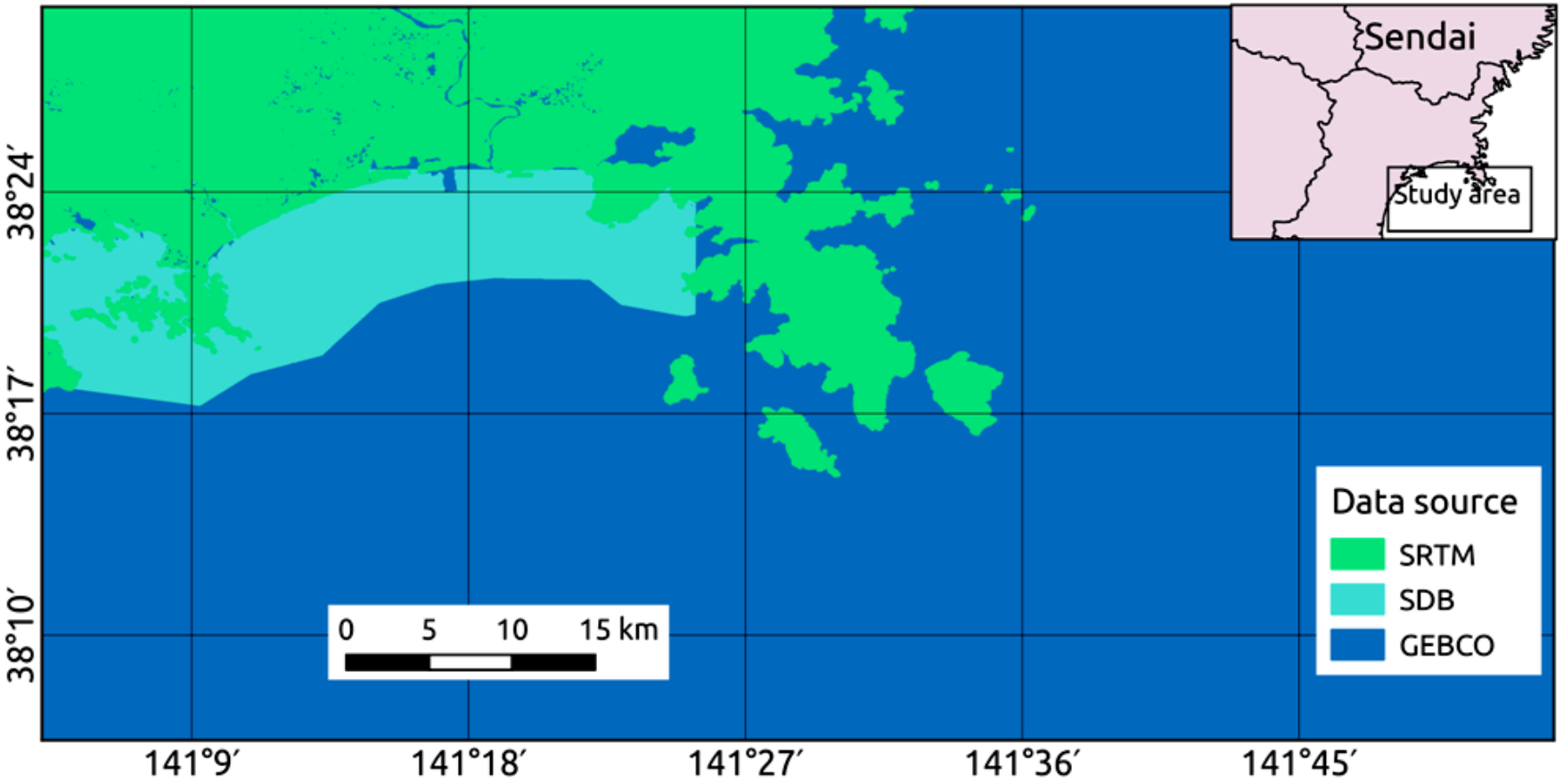
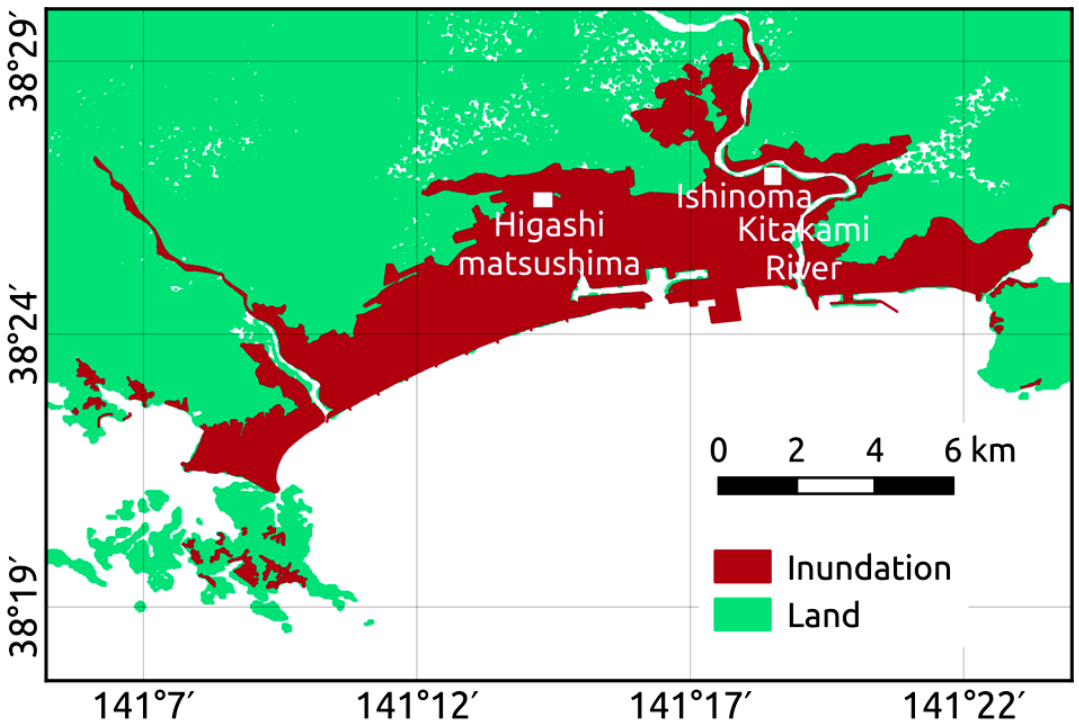


Figure 4.1: Study area parts of Miyagi prefecture, Japan

(a)



(b)

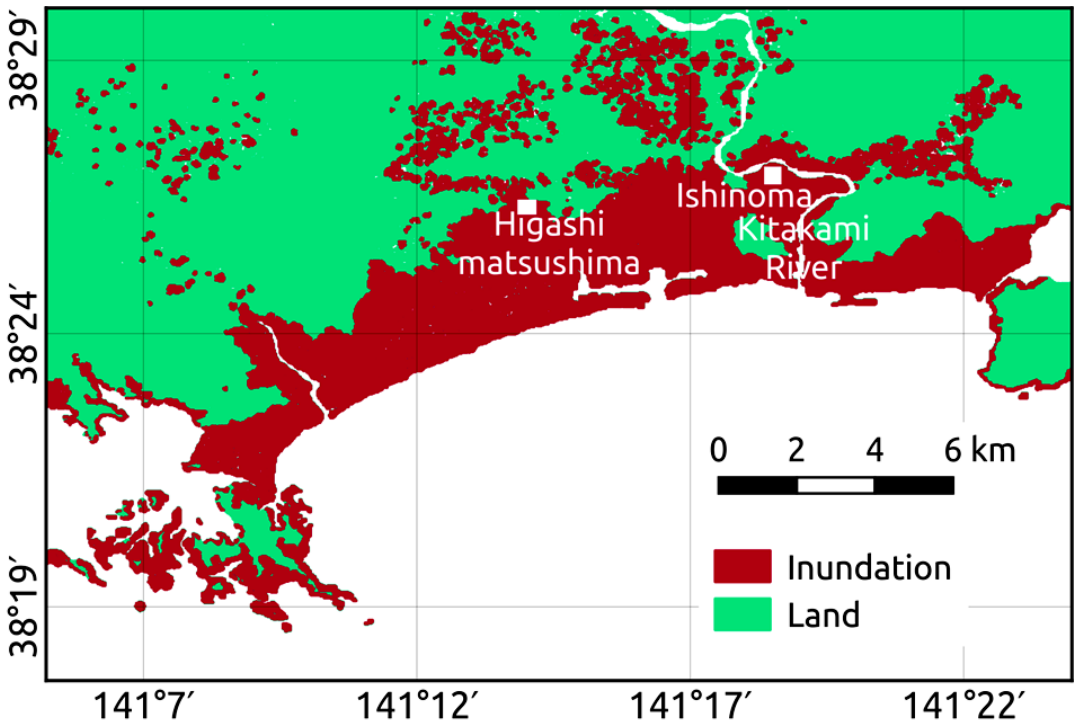
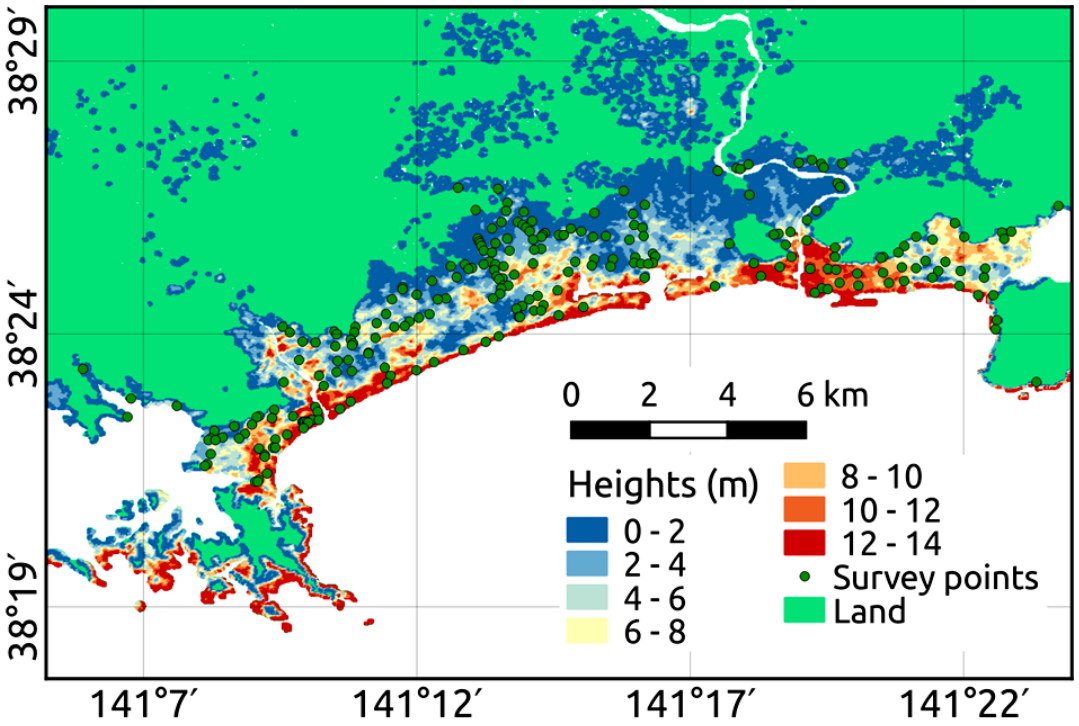


Figure 4.2: Evaluation of Tohoku tsunami simulation results. (a) Actual tsunami inundation extent, (b) simulation inundation extent (continued).

(a)



(b)

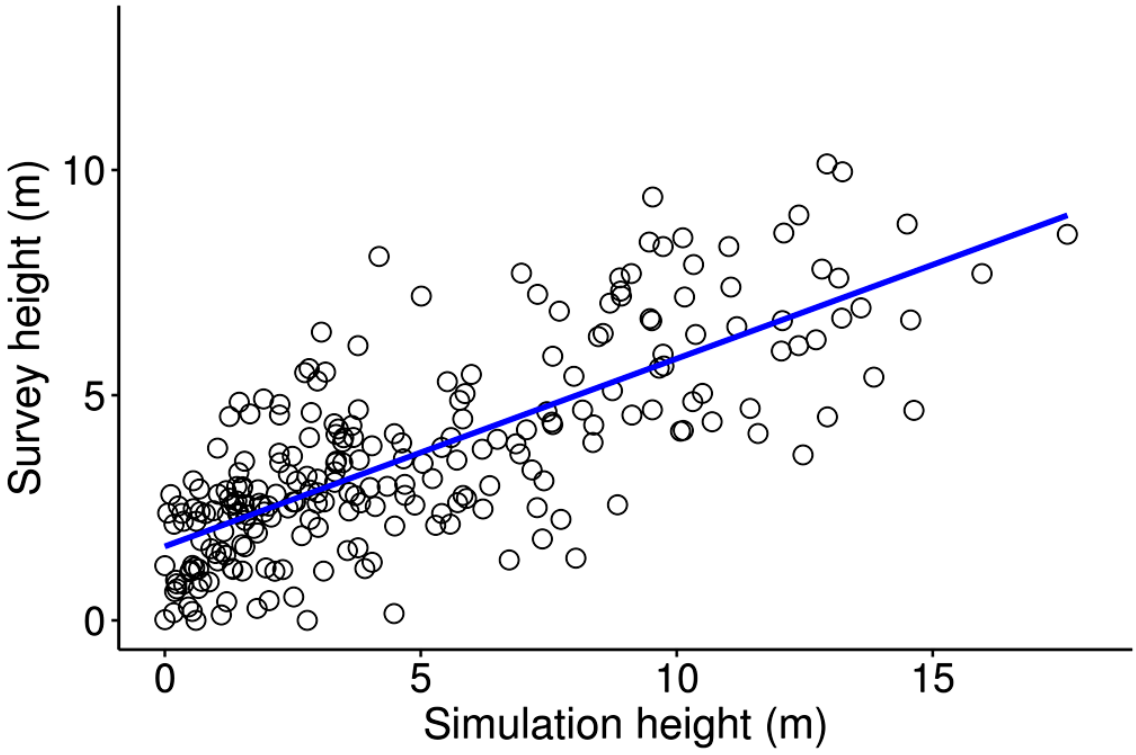


Figure 4.3: Evaluation of Tohoku tsunami simulation results. (a) survey tsunami inundation height points overlaid on simulated tsunami inundation height map (b) scatter plot between surveyed and simulated tsunami inundation height

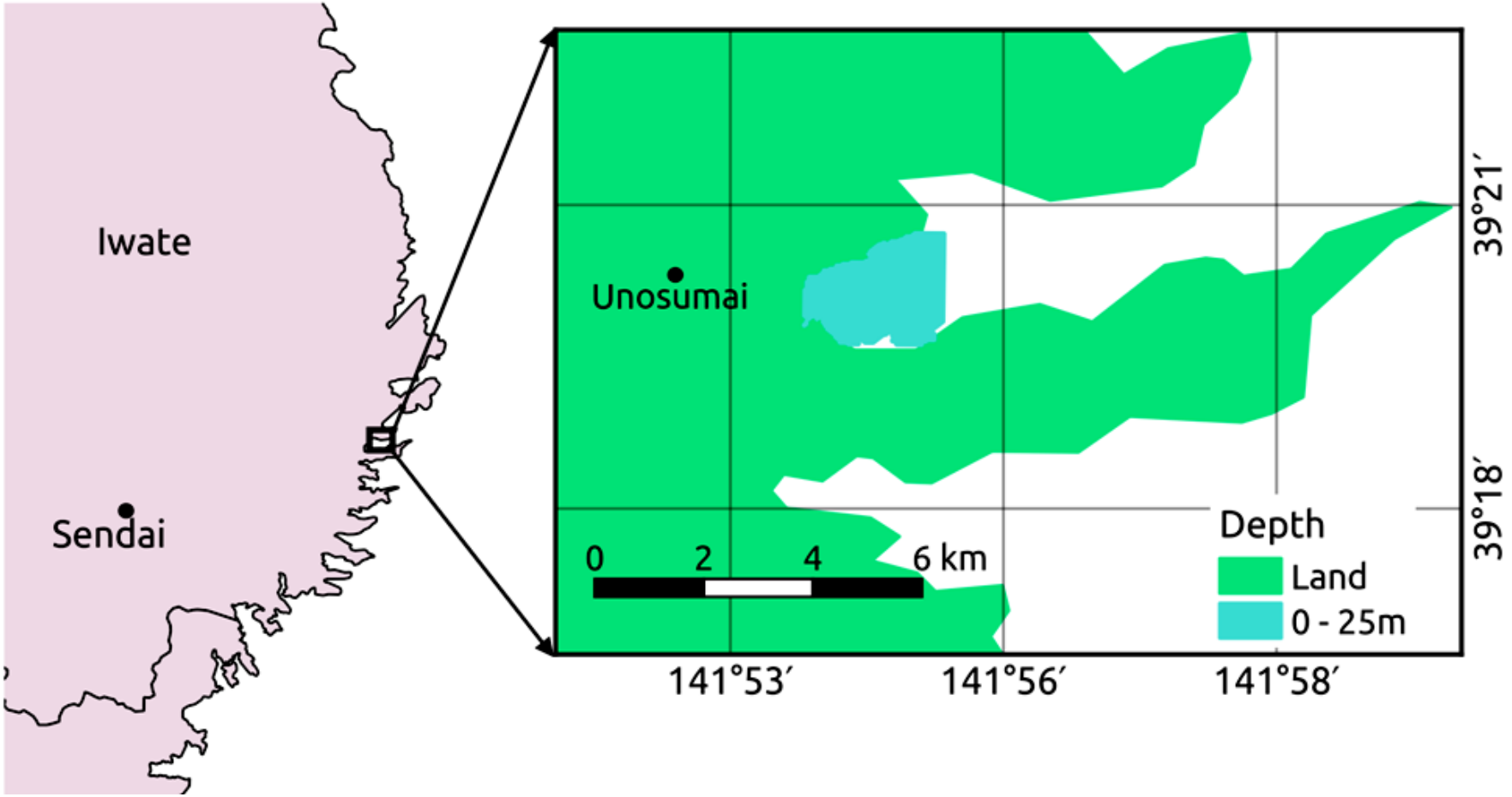


Figure 5.1: Study area, Parts of Iwate prefecture, Japan.

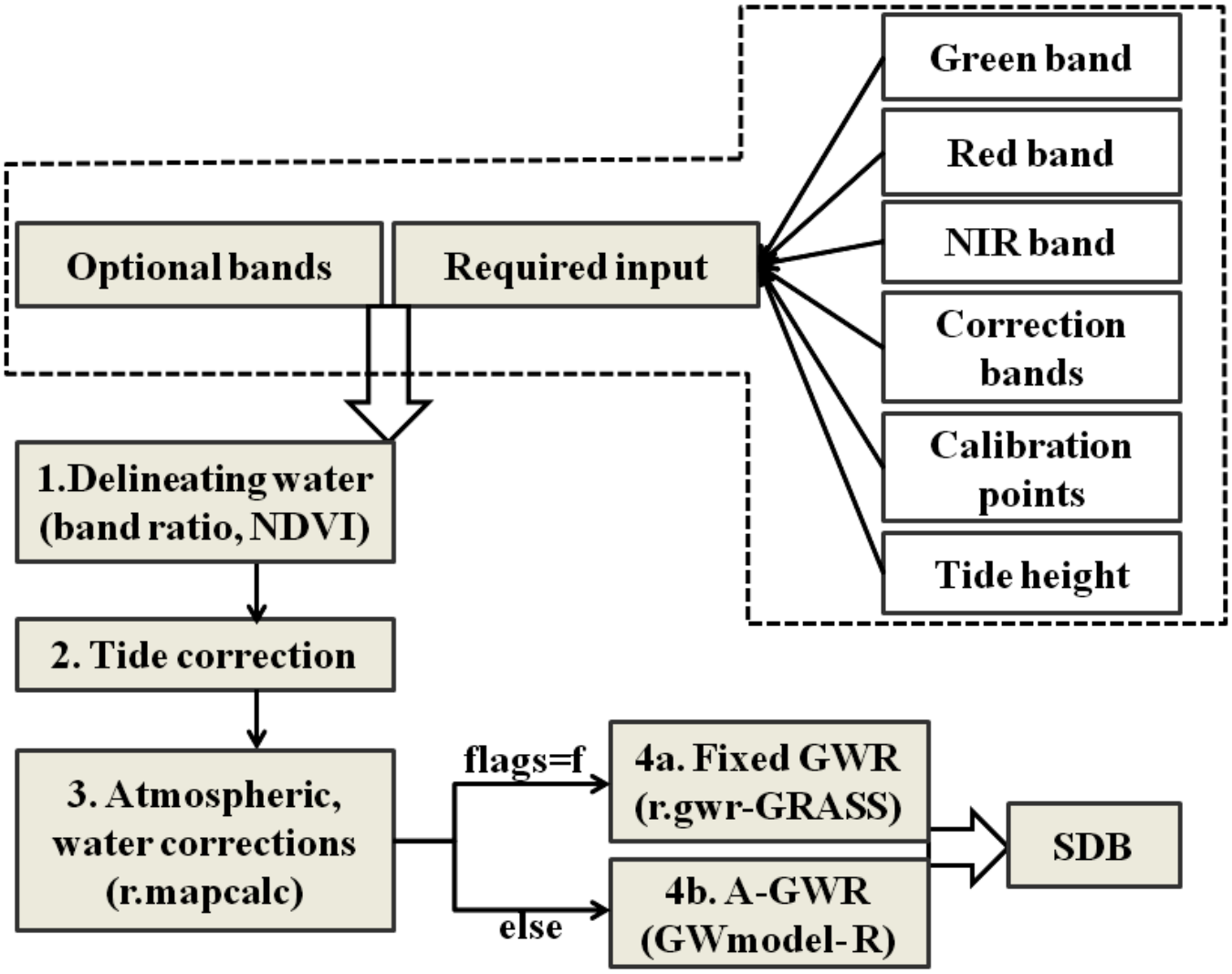


Figure 5.2: Flowchart of workflow of *i.image.bathymetry*

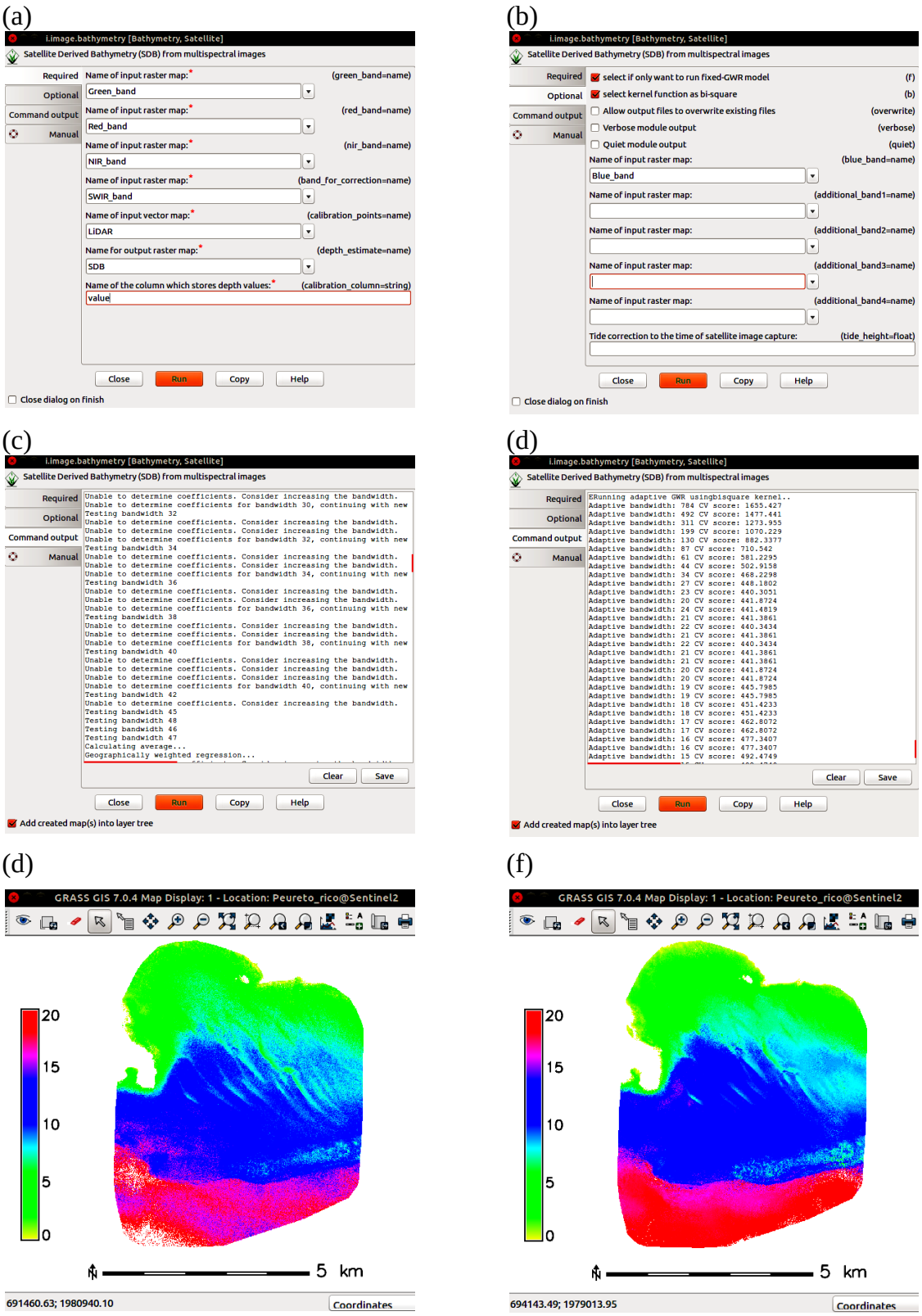


Figure 5.3: Screenshots demonstrate the processing of *i.image.bathymetry*, (a) required and, (b) optional band and flags selected for Fixed-GWR, (c) optimal bandwidth for Fixed-GWR, (d) optimal bandwidth for A-GWR (e) SDB from Fixed-GWR and (f) SDB from A-GWR

Table 2.1: Characteristics of data used

Date	Source	Tide(m)	Res.(m)
09-05-2012	Landsat-7	1.15	30
27 May, 2012	*In-situ* depth	1.53-2.3	10
30 Mar, 2003	Landsat-7	2	30
2003	NHO chart (1:60,000)	2.3	30
08 May, 2000	Landsat-7	1.36	30
12 Nov, 2013	Landsat-8	1.25	30
31 Jan, 2014	Landsat-8	2.40	30
04 Mar, 2014	Landsat 8	1.66	30
17 Nov, 2013	*In-situ* depth	2.24-2.16	7
10 Dec, 2012	ASTER	1.03	15
16 Jan, 2003	ASTER	1.8	15
5 Apr, 2014	Landsat-8	1.37	30
30 Nov, 2014	*In-situ* depth	0	2

Table 2.2: The estimated multi-temporal SDB from various data sets using single spectral band model

Data source	Date	Method	R	R^2	RMSE
Landsat-7	09 May, 2012	SLR	0.67	0.46	2.14
	30 Mar, 2003	RBE	0.63	0.41	2.83
	8 May, 2000	RBE	0.74	0.56	2.88
Landsat-8	12 Nov, 2013	SLR	0.85	0.74	1.46
	31 Jan, 2014	RBE	0.89	0.81	1.39
	04 Mar, 2014	RBE	0.80	0.66	2.09
ASTER	10 Dec, 2012	SLR	0.70	0.50	2.10
	16 Jan, 2003	RBE	0.57	0.34	2.38

Table 2.3: Comparison of accuracy of SDB from MLR and MLR-NIR

Data source	Date	MLR			MLR-NIR		
		R	R^2	RMSE	R	R^2	RMSE
Landsat-7	8 May, 2000	0.80	0.66	1.64	0.82	0.69	1.57
	30 Mar, 2003	0.65	0.44	2.44	0.77	0.61	2.07
	09 May, 2012	0.53	0.30	1.77	0.68	0.48	1.53
Landsat-8	12 Nov, 2013	0.91	0.84	1.77	0.95	0.91	1.53
	31 Jan, 2014	0.91	0.84	1.12	0.95	0.91	0.83
	04 Mar, 2014	0.85	0.74	1.43	0.90	0.83	1.23
ASTER	16 Jan, 2003	0.76	0.59	1.90	0.78	0.62	1.86
	10 Dec, 2012	0.69	0.50	1.75	0.76	0.60	1.56

Table 2.4: Comparison of accuracy of SDB estimated from Landsat-8 data in terms of turbidity

Date	Methods	NDWI	R	R^2	RMSE
12 Nov, 2013	RBE	Low turbid	0.88	0.78	1.21
		High turbid	0.82	0.67	1.50
	MLR-NIR	Low turbid	0.97	0.94	0.85
		High turbid	0.90	0.81	0.90
31 Jan, 2014	RBE	Low turbid	0.82	0.68	1.27
		High turbid	0.77	0.60	1.36
	MLR-NIR	Low turbid	0.95	0.91	0.97
		High turbid	0.84	0.71	1.65
04 Mar, 2014	RBE	Low turbid	0.77	0.59	2.70
		High turbid	0.65	0.43	2.82
	MLR-NIR	Low turbid	0.93	0.86	1.48
		High turbid	0.85	0.72	1.85

Table 2.5: Comparison of SDB obtained from NIR and SWIR band correction methods

Study area	Correction band	R	R^2	RMSE
Taketomi	SWIR	0.90	0.81	2.19
	NIR	0.87	0.77	2.42
Ratnagiri	SWIR	0.90	0.81	1.24
	NIR	0.85	0.74	1.38

Table 3.1: Characteristics of the data used

Data	Date	Res.(m)	Estimation bands	Correction band
Landsat-8	25 Nov, 2013	30	0.43-0.45µm (coastal) 0.45-0.51µm (blue) 0.53-0.59µm (green) 0.64-0.67µm (red) 0.85 - 0.88µm (NIR)	1.57-1.65µm (SWIR)
RapidEye	01 May, 2010	5	0.44-0.51µm (blue) 0.52-0.59µm(green) 0.63-0.69µm (red) 0.69-0.73µm (Red-edge)	0.76-0.85µm (NIR)
LiDAR	07 Apr, 2006 – 15 May, 2006	4	0.523 µm	1.064µm

Table 3.2: Comparison of accuracy of SDB from global model and class based model

Satellite data	Global model			Class based model		
	R	R^2	RMSE	R	R^2	RMSE
Landsat-8	0.88	0.78	2.63	0.94	0.90	1.70
RapidEye	0.88	0.78	2.48	0.94	0.88	1.82

Table 3.3: Comparison of accuracy of SDB from global, Fixed-GWR and A-GWR model

Data	Global model			Fixed-GWR model			A-GWR model		
	R	R^2	RMSE	R	R^2	RMSE	R	R^2	RMSE
Landsat-8	0.88	0.78	2.63	0.96	0.93	1.41	0.98	0.95	1.13
RapidEye	0.88	0.78	2.48	0.98	0.96	1.06	0.99	0.99	0.41

Table 3.4: Statistics of differences for each cross-shore profile (A-B, C-D and E-F)

Data	Profile	LiDAR					A-GWR model					Comparison		
		N	Min	Max	Mean	STD	N	Min	Max	Mean	STD	R	R^2	RMSE
Landsat-8	A-B	240	2.71	19.65	11.68	4.08	240	2.97	19.57	11.68	4.02	0.99	0.99	0.29
	C-D	127	2.85	20.00	15.23	5.47	127	1.94	19.58	15.08	5.43	0.99	0.98	0.85
	E-F	181	4.56	18.61	13.53	2.85	181	3.94	19.32	13.63	2.99	0.92	0.95	0.84
RapidEye	A-B	502	2.65	19.79	11.77	4.11	502	2.70	19.79	11.78	4.11	0.99	0.99	0.21
	C-D	480	2.92	20.00	14.95	5.65	480	2.96	19.97	14.77	5.68	0.99	0.99	0.29
	E-F	502	5.21	18.70	13.72	2.61	502	5.36	18.56	13.77	2.56	0.97	0.95	0.59

Table 3.5: Comparison of accuracy in each regular depth categories from A-GWR model.

Data		1-5m	5-10m	10-15m	15-20m	20-25m
Landsat-8	N	21919	29126	45064	77829	35359
	R	0.73	0.77	0.73	0.74	0.63
	R^2	0.53	0.59	0.54	0.54	0.39
	RMSE	0.92	1.18	1.26	1.09	1.50
RapidEye	N	1005762	1013220	1269665	2118627	1313220
	R	0.97	0.93	0.89	0.90	0.78
	R^2	0.95	0.88	0.80	0.82	0.62
	RMSE	0.33	0.53	0.69	0.62	2.01

Table 3.6: Comparison of SDB was from A-GWR based model from Scenario-1 and Scenario-2

Scenarios	Landsat-8			RapidEye		
	R	R^2	RMSE	R	R^2	RMSE
Scenario-1	0.96	0.94	1.38	0.97	0.95	1.55
Scenario-2	0.96	0.93	1.40	0.96	0.92	1.44

Table 4.1: Characteristics of data used

Data	Date	Res.(m)	Estimation bands	Correction band
ASTER	10 Sep, 2010	15	0.52 - 0.60µm (green) 0.63 - 0.69µm (red)	0.76 - 0.86 µm (NIR)
JDOC	Before 1998	500		

Table 5.1: Characteristics of data used

Data	Date	Res.(m)	Estimation bands	Correction band
Landsat-8	31 Oct, 2014	30	0.43-0.45µm (coastal) 0.45-0.51µm (blue) 0.53-0.59µm (green) 0.64-0.67µm (red) 0.85 - 0.88µm (NIR)	1.57-1.65µm (SWIR)
Sentinel-2	25 Dec, 2015	10 10 10 20 20 10	0.49µm (blue) 0.56µm(green) 0.66µm (red) 0.70µm (Red-edge) 0.74µm (Red-edge) 0.84µm (NIR)	1.61µm (SWIR)
LiDAR	07 Apr, 2006 - 15 May, 2006	4	0.52µm	1.06µm
SoNAR	10 Jun, 2012	2		

Table 5.2: Performance of *i.image.bathymetry* in case studies 1, 2 and 3

Area	Data	Kernel	GWR model	Time(minute)	SDB results		
					R	R^2	RMSE
Puerto Rico	Sentinel-2	Gaussian	Fixed-GWR	2.22	0.99	0.98	0.67
			A-GWR	180	0.99	0.98	0.62
		Bi-square	Fixed-GWR	2.24	0.99	0.98	0.64
			A-GWR	184	0.99	0.99	0.61
Miyagi	ASTER	Gaussian	Fixed-GWR	3.02	0.91	0.83	1.93
			A-GWR	265	0.89	0.80	2.20
		Bi-square	Fixed-GWR	3.08	0.93	0.87	1.65
			A-GWR	255	0.91	0.84	1.95
Iwate	Landsat-8	Gaussian	Fixed-GWR	2.50	0.86	0.74	2.91
			A-GWR	6.00	0.96	0.93	1.54
		Bi-square	Fixed-GWR	2.50	0.88	0.77	2.77
			A-GWR	5.55	0.97	0.94	1.50

Appendix

As mentioned in the Chapter Five, *i.image.bathymetry* the module developed using python GRASS scripting as shown below. The script is also available in GRASS GIS SVN repository. *i.image.bathymetry* can be considered as the first Open Source SDB module. Like all the GRASS GIS module *i.image.bathymetry* uses the GRASS parser mechanism. This very advanced parser helps to check the user input format and optionally create a graphical user interface for the module. In Python, parser() function used to call the parser from grass.script package. These comments start with #% and can be referred to as interface definition comments.

```
#!/usr/bin/env python

############################################################################
###
#
# MODULE: i.image.bathymetry
# AUTHOR(S): Vinayaraj Poliyapram <vinay223333@gmail.com> and Luca Delucchi
# PURPOSE:   Script for estimating bathymetry from optical satellite images
# COPYRIGHT: (C) Vinayaraj Poliyapram and by the GRASS Development Team
#            This program is free software under the GNU General
#            Public License (>=v2). Read the file COPYING that
#            comes with GRASS for details.
#
############################################################################
###
```

 # Input parameters and flags definition

```
#%module
#% description: Estimates Satellite Derived Bathymetry (SDB) from multispectral images.
#% keyword: imagery
#% keyword: bathymetry
#% keyword: satellite
#%end
#%option G_OPT_R_INPUT
#% key: blue_band
#%required: no
#%end
#%option G_OPT_R_INPUT
#% key: green_band
#%required: yes
```

```
#%end
#%option G_OPT_R_INPUT
#% key: red_band
#%required: yes
#%end
#%option G_OPT_R_INPUT
#% key: nir_band
#%required: yes
#%end
#%option G_OPT_R_INPUT
#% key: band_for_correction
#%required: yes
#%end
#%option G_OPT_V_INPUT
#% key: calibration_points
#%required: yes
#%end
#%option G_OPT_V_INPUT
#% key: area_of_interest
#%required: no
#%end
#%option G_OPT_R_INPUT
#% key: additional_band1
#%required: no
#%end
#%option G_OPT_R_INPUT
#% key: additional_band2
#%required: no
#%end
#%option G_OPT_R_INPUT
#% key: additional_band3
#%required: no
#%end
#%option G_OPT_R_INPUT
#% key: additional_band4
#%required: no
#%end
#%option G_OPT_R_OUTPUT
#% key: depth_estimate
#%required: yes
#%end
#%option
#% key: tide_height
#%type: double
#%multiple: no
#%required: no
#%description: Tide correction to the time of satellite image capture
#%end
```

```
#%option G_OPT_DB_COLUMN
#% key: calibration_column
#%required: yes
#%description: Name of the column which stores depth values
#%end
#%flag
#%key: f
#%description: select if only want to run Fixed-GWR model
#%end
#%flag
#%key: b
#%description: select kernel function as bi-square
#%end

# importing required python libraries
import atexit
import grass.script as g
from grass.pygrass.raster import RasterRow
import subprocess
import os
crctd_lst = []

# Define the input option names
def main():
    options, flags = g.parser()

    Blue = options['blue_band']
    Green = options['green_band']
    Red = options['red_band']
    NIR = options['nir_band']
    SWIR = options['band_for_correction']
    Calibration_points = options['calibration_points']
    Area_of_interest = options['area_of_interest']
    Additional_band1 = options['additional_band1']
    Additional_band2 = options['additional_band2']
    Additional_band3 = options['additional_band3']
    Additional_band4 = options['additional_band4']
    bathymetry = options['depth_estimate']
    tide_height = options['tide_height']
    calibration_column = options['calibration_column']
    bisquare = flags['b']
    fixed_GWR = flags['f']

    # setting the grass region same as input spectral band
    res = g.parse_command('g.region', raster=Green, flags='g')
    g.run_command('v.to.rast', input=Calibration_points, type='point',
                  use='attr', attribute_column=calibration_column,
                  output='tmp_Calibration_points')
```

```
# hull generation from calibration depth points
g.run_command('v.hull', input=Calibration_points, output='tmp_hull',
          overwrite=True)
# buffer the hull to ceate a region for including all calibration points
g.run_command('v.buffer', input='tmp_hull', output='tmp_buffer',
          distance=float(res['nsres']), overwrite=True)
# applying tide correction
if tide_height:
   cal = g.parse_command('r.univar', map='tmp_Calibration_points',
                 flags='g')
   if float(cal['min']) >= 0:
      t = float(tide_height)
      g.mapcalc(exp="{d}=({d}+float({t}))".format(d='tmp_Calibration_points',
                                    t=t),
            overwrite=True)
   if float(cal['min']) < 0:
      t = float(tide_height) * -1
      g.mapcalc(exp="{d}=({d}+float({t}))".format(d='tmp_Calibration_points',
                                    t=t),
            overwrite=True)
# estimating band ratio
g.mapcalc(exp="{tmp_ratio}=({Green}/{SWIR})".format(tmp_ratio='tmp_ratio',
       Green=Green, SWIR=SWIR))
# estimating NDVI
g.mapcalc(exp="{tmp_NDVI}=float({NIR}-{Red})/float({NIR}+{Red})"
       .format(tmp_NDVI='tmp_NDVI', NIR=NIR, Red=Red))
g.mapcalc(exp="{tmp_water}=if({tmp_ratio} < 1, null(), if({tmp_NDVI} <"
          "0, {tmp_ratio}, null()))".format(tmp_NDVI='tmp_NDVI',
                                tmp_water='tmp_water',
                                tmp_ratio='tmp_ratio'))
g.run_command('r.mask', raster='tmp_water', overwrite=True)
li = [Green, Additional_band1, Additional_band2, Additional_band3,
    Additional_band4, Blue, Red]
for i in li:
   j, sep, tail = i.partition('@')
   tmp_ = RasterRow(str(i))
   if tmp_.exist() is False:
      continue
   g.message("Ditermining minimum value for %s" % i)
   g.run_command('g.region', vector=Calibration_points)
   # ignore zero values may be available in the DN value
   g.mapcalc(exp="{tmp_b}=if({x}>1, {x},null())".format(tmp_b='tmp_b',
          x=str(i)), overwrite=True)
   tmp_AOI = g.parse_command('r.univar', map='tmp_b', flags='g')
   tmp_AOI_min = float(tmp_AOI['min'])
   g.run_command('g.region', raster=Green)
   try:
      # estimate the deep water pixels
```

```
        g.mapcalc(exp="{tmp_deep}=if({tmp_band}<{band_min}, {tmp_band},"
                  "null())".format(tmp_deep='tmp_deep',
                                   band_min=tmp_AOI_min,
                                   tmp_band=str(i)),
                  overwrite=True)
        g.run_command('r.mask', raster='tmp_deep', overwrite=True)
        tmp_coe = g.parse_command('r.regression.line', mapx=SWIR,
                                  mapy=str(i), flags='g')
        g.message("Deep water ditermination for %s" % i)

        # bathymetry to be estimated in the area which user provided as polygon file
        if Area_of_interest:
           g.run_command('r.mask', vector=Area_of_interest, overwrite=True)
           g.run_command('g.region', vector=Area_of_interest)
        else:
           # bathymetry to be estimated in the area according to the calibration points
           g.run_command('r.mask', vector='tmp_buffer', overwrite=True)
           g.run_command('g.region', vector=Calibration_points)
        # atmospheric and water corrections using infrared band
        g.mapcalc(exp="{tmp_crct}=log({tmp_band}-{a}-{b}*{SWIR})"
                  .format(tmp_crct='tmp_crct' + str(j),
                          tmp_band=str(i), a=float(tmp_coe['a']),
                          b=float(tmp_coe['b']), SWIR=SWIR),
                   overwrite=True)
        g.run_command('r.mask', raster='tmp_water', overwrite=True)
        g.mapcalc("{tmp_crctd} = ({tmp_crct} * 1)"
                  .format(tmp_crct='tmp_crct'+str(j), tmp_crctd='tmp_crctd' + str(j)))
    except:
      # if the deep water pixels are not included in the satellite image, correction will
        be carried out with  using deep water coefficients
        g.message("Cannot find deep water pixels")
        if Area_of_interest:
           g.run_command('r.mask', vector=Area_of_interest, overwrite=True)
           g.run_command('g.region', vector=Area_of_interest)
        else:
           g.run_command('r.mask', vector='tmp_buffer', overwrite=True)
           g.run_command('g.region', vector=Calibration_points)
        g.mapcalc(exp="{tmp_crct}=log({tmp_band}-{a}-{b}*{SWIR})"
                  .format(tmp_crct='tmp_crct' + str(j),
                          tmp_band=str(i), a=float(tmp_coe['a']),
                          b=float(tmp_coe['b']), SWIR=SWIR),
                   overwrite=True)
        g.run_command('r.mask', raster='tmp_water', overwrite=True)
        g.mapcalc("{tmp_crctd} = ({tmp_crct} * 1)"
                  .format(tmp_crct='tmp_crct'+str(j), tmp_crctd='tmp_crctd' + str(j)))
    crctd_lst.append('tmp_crctd' + str(j))
# if user selected –f flag for fixed-GWR based SDB model
if fixed_GWR:
```

```
    if not g.find_program('r.gwr', '--help'):
        g.run_command('g.extension', extension='r.gwr')
    if bisquare:
        g.message("Calculating optimal bandwidth using bisqare kernel...")
        bw = g.parse_command('r.gwr', mapx=crctd_lst,
                             mapy='tmp_Calibration_points',
                             kernel='bisquare', flags='ge')
        g.message("Running Fixed-GWR using bisqare kernel...")
        g.run_command('r.gwr', mapx=crctd_lst,
                      mapy='tmp_Calibration_points',
                      estimates='tmp_bathymetry', kernel='bisquare',
                      bandwidth=int(bw['estimate']))
    else:
        g.message("Calculating optimal bandwidth using gaussian kernel...")
        bw = g.parse_command('r.gwr', mapx=crctd_lst,
                             mapy='tmp_Calibration_points', flags='ge')
        g.message("Running Fixed-GWR using gaussian kernel...")
        g.run_command('r.gwr', mapx=crctd_lst,
                      mapy='tmp_Calibration_points',
                      estimates='tmp_bathymetry',
                      bandwidth=int(bw['estimate']))
else:
    global r
    global predict
    try:
        # in case A-GWR model use GWmodel in R
        r = g.tempfile()
        r_file = open(r, 'w')
        # check if R libraries installed in the if not install
        libs = ['GWmodel', 'data.table', 'rgrass7', 'rgdal', 'raster']
        for i in libs:
            install = 'if(!is.element("%s", installed.packages()[,1])){\n' % i
            install += "cat('\\n\\nInstalling %s package from CRAN\n')\n" % i
            install += "if(!file.exists(Sys.getenv('R_LIBS_USER'))){\n"
            install += "dir.create(Sys.getenv('R_LIBS_USER'), recursive=TRUE)\n"
            install += ".libPaths(Sys.getenv('R_LIBS_USER'))}\n"
            install += 'install.packages("%s", repos="http://cran.us.r-' \
              'project.org")}\n' % i
            r_file.write(install)
            libraries = 'library(%s)\n' % i
            r_file.write(libraries)
        Green_new, sep, tail = Green.partition('@')
        r_file.write('grass_file = readRAST("tmp_crctd%s")\n' % Green_new)
        r_file.write('raster_file = raster(grass_file)\n')
        frame_file = 'pred = as.data.frame(raster_file,na.rm = TRUE,xy = TRUE)\n'
        r_file.write(frame_file)
        for i in li:
            j, sep, tail = i.partition('@')
```

```
    Green_new, sep, tail = Green.partition('@')
    tmp_ = RasterRow(str(i))
    if tmp_.exist() is False:
        continue
   # exports files from GRASS to R
    r_file.write('grass_file = readRAST("tmp_crctd%s")\n' % j)
    r_file.write('raster_file = raster(grass_file)\n')
    r_file.write('frame_pred%s = as.data.frame(raster_file, na.rm = TRUE,'
                 'xy = TRUE)\n' % j)
    pred_file = 'frame_pred_green=data.frame(frame_pred%s)\n' % Green_new
    pred_file += 'pred=merge(pred, frame_pred%s)\n' % j
    r_file.write(pred_file)
    g.run_command('r.mask', raster='tmp_Calibration_points',
                          overwrite=True)
    r_file.write('grass_file=readRAST("%s")\n' % 'tmp_Calibration_points')
    r_file.write('raster_file = raster(grass_file)\n')
    frame_file = 'calib = as.data.frame(raster_file,na.rm = TRUE ,' \
         'xy = TRUE)\n'
    r_file.write(frame_file)
# creating independent variables for GWmodel processing
for i in li:
    j, sep, tail = i.partition('@')
    tmp_ = RasterRow(str(i))
    if tmp_.exist() is False:
        continue
    r_file.write('grass_file = readRAST("tmp_crctd%s")\n' % j)
    r_file.write('raster_file = raster(grass_file)\n')
    r_file.write('frame_ref%s = as.data.frame(raster_file,na.rm = TRUE,' \
         'xy = TRUE)\n' % j)
    ref_file = 'calib = merge(calib, frame_ref%s)\n' % j
    r_file.write(ref_file)
g.run_command('g.remove', type='raster', pattern='MASK', flags='f')
ref_file = 'Rapid_ref.sdf=SpatialPointsDataFrame(calib[,1:2],calib)\n'
ref_file += 'Rapid_pred.sdf=SpatialPointsDataFrame(pred[,1:2],' \
        'pred)\n'
ref_file += 'DM_Rapid_ref.sdf=gw.dist(dp.locat=coordinates' \
        '(Rapid_ref.sdf))\n'
r_file.write(ref_file)
l = []
predict = g.read_command("g.tempfile", pid=os.getpid()).strip() + '.txt'
# join the corrected bands in to a string
le = len(crctd_lst)
for i in crctd_lst:
    l.append(i)
    k = '+'.join(l)
if bisquare:
    # bi-square Adaptive kernel processing
    ref_flag = "cat('\nCalculating optimal bandwidth using " \
```

```
                "bisquare kernel..\n')\n"
            ref_flag += 'BW_Rapid_ref.sdf=bw.gwr(tmp_Calibration_points~%s,' \
                  'data=Rapid_ref.sdf, kernel="bisquare",' \
                  'adaptive=TRUE, dMat=DM_Rapid_ref.sdf)\n' % k
            ref_flag += "cat('\nCalculating euclidean distance\n')\n"
            ref_flag += 'DM_Rapid_pred.sdf=gw.dist(dp.locat=coordinates' \
                  '(Rapid_ref.sdf), rp.locat=coordinates' \
                  '(Rapid_pred.sdf))\n'
            ref_flag += "cat('\nRunning A-GWR using bisquare kernel\n')\n"
            ref_flag += 'GWR_Rapid_pred.sdf=gwr.predict(tmp_Calibration_poi' \
                  'nts~%s,data=Rapid_ref.sdf, bw = BW_Rapid_ref.sdf,' \
                  'predictdata = Rapid_pred.sdf, kernel = "bisquare",' \
                  'adaptive = TRUE, dMat1 = DM_Rapid_pred.sdf,' \
                  'dMat2 = DM_Rapid_ref.sdf)\n' % k
            r_file.write(ref_flag)
        if not bisquare:
            # gaussian Adaptive kernel processing
            ref_fla = "cat('\nCalculating optimal bandwidth using " \
                   "gaussian kernel..\n')\n"
            ref_fla += 'BW_Rapid_ref.sdf=bw.gwr(tmp_Calibration_points~%s,' \
                 'data=Rapid_ref.sdf, kernel="gaussian",' \
                 'adaptive=TRUE, dMat= DM_Rapid_ref.sdf)\n' % k
            ref_fla += "cat('\nCalculating euclidean distance\n')\n"
            ref_fla += 'DM_Rapid_pred.sdf=gw.dist(dp.locat=coordinates' \
                 '(Rapid_ref.sdf), rp.locat=coordinates' \
                 '(Rapid_pred.sdf))\n'
            ref_fla += "cat('\nRunning A-GWR using gaussian kernel\n')\n"
            ref_fla += 'GWR_Rapid_pred.sdf = gwr.predict(tmp_Calibration_poi' \
                 'nts~%s,data=Rapid_ref.sdf, bw=BW_Rapid_ref.sdf,' \
                 'predictdata = Rapid_pred.sdf, kernel = "gaussian",' \
                 'adaptive = TRUE, dMat1 = DM_Rapid_pred.sdf,' \
                 'dMat2 = DM_Rapid_ref.sdf)\n' % k
            r_file.write(ref_fla)
        ref_fil = 'Sp_frame = as.data.frame(GWR_Rapid_pred.sdf$SDF)\n'
        r_file.write(ref_fil)
        # export estimated depth to a grass temporary file
        export = 'write.table(Sp_frame, quote=FALSE, sep=",",' \
                 '"%s")\n' % predict
        r_file.write(export)
        r_file.close()
        subprocess.check_call(['Rscript', r], shell=False)
        g.run_command('r.in.xyz', input=predict,
                output='tmp_bathymetry', skip=1, separator=",",
                x=(int(le) + 5), y=(int(le) + 6), z=(int(le) + 3),
                overwrite=True)
    except subprocess.CalledProcessError:
        g.message("Integer outflow... ")
      # check if r.gwr addon installed and if not install
```

```
        if not g.find_program('r.gwr', '--help'):
            g.run_command('g.extension', extension='r.gwr')
        if bisquare:
            # bi-square Fixed kernel processing
            g.message("Running Fixed-GWR using bisqare kernel...")
            bw = g.parse_command('r.gwr', mapx=crctd_lst,
                        mapy='tmp_Calibration_points',
                        kernel='bisquare', flags='ge')
            g.run_command('r.gwr', mapx=crctd_lst,
                    mapy='tmp_Calibration_points',
                    estimates='tmp_bathymetry', kernel='bisquare',
                    bandwidth=int(bw['estimate']))
        else:
            # gaussian Fixed kernel processing
            g.message("Running Fixed-GWR using gaussian kernel...")
            bw = g.parse_command('r.gwr', mapx=crctd_lst,
                        mapy='tmp_Calibration_points', flags='ge')
            g.run_command('r.gwr', mapx=crctd_lst,
                    mapy='tmp_Calibration_points',
                    estimates='tmp_bathymetry',
                    bandwidth=int(bw['estimate']))
    tmp_rslt_ext = g.parse_command('r.univar', map='tmp_Calibration_points',
                           flags='g')
    g.mapcalc(exp="{bathymetry}=if({tmp_SDB}>{max_}, null(),"
                "if({tmp_SDB}<{min_}, null(), {tmp_SDB}))".format
                (tmp_SDB='tmp_bathymetry', bathymetry=bathymetry,
                 max_=float(tmp_rslt_ext['max']),
                 min_=float(tmp_rslt_ext['min'])))


def cleanup():
    # clean the temp files
    g.run_command('g.remove', type='raster', name='MASK', flags='f')
    g.run_command('g.remove', type='all', pattern='*tmp*', flags='f')
    try:
        g.try_remove(predict)
        g.try_remove(r)
    except:
        pass

if __name__ == '__main__':
    atexit.register(cleanup)
    main()
```