DELFT UNIVERSITY OF TECHNOLOGY

MASTERS' THESIS

Reservoir Characterization using a Geometric Approach

Author:
Anshul Khandelwal

Supervisors/Committee:
Prof. Dr. Elmar Eisemann
Dr. Klaus Hildebrandt
Prof. Dr. Nick van de Giesen
Dr. Cynthia C. S. Liem

This thesis was submitted to obtain the degree
Master of Science in Computer Science
Data Science & Technology Track

in the

Computer Graphics & Visualization Group
Faculty of Electrical Engineering, Mathematics and Computer Science

September 12, 2018

DELFT UNIVERSITY OF TECHNOLOGY

Abstract

Computer Graphics & Visualization Group
Faculty of Electrical Engineering, Mathematics and Computer Science

Master of Science

Reservoir Characterization using a Geometric Approach

by Anshul Khandelwal

Quantifying the anthropogenic impacts such as reservoir characterization is a big challenge in the field of water management. In this work, a computer graphics based geometric approach is presented which can predict the underlying topology of large-scale reservoirs. The proposed algorithm uses freely available, satellite based landscape data of the surrounding regions to predict reservoir characteristics. The premise of the presented approach is that the slope of the surrounding landscape is an important determinant to understand the underlying landscape of the reservoirs. This method outperforms the existing state-of-the-art techniques used to estimate the storage capacities drastically, both in terms of estimated maximum volume stored and estimated volume area curves. Evaluation of the geometric model presented is done on 28 reservoirs using the HydroSHEDs data which was developed using the Shuttle RADAR Topography Mission conducted by NASA. This HydroSHEDs data was obtained in 2000 which acts as ground truth data for the reservoirs built after 2000. Further, model parameters are introduced to improve the modeling capabilities of the reconstructed reservoirs. This approach further intensifies the case of using computer graphics techniques for raster based analysis and provides a platform for further research in the field of water management.

Acknowledgements

This research was conducted in collaboration with the Computer Graphics & Visualization Group at the Faculty of Electrical Engineering, Mathematics and Computer Science and the Water Management Group at the Faculty Of Civil Engineering and Applied Earth Sciences of the Delft University of Technology.

The successful completion of this thesis project was primarily attributed to my thesis supervisors - Dr. Klaus Hildebrandt, Prof. Elmar Eisemann and Prof. Nick van de Giesen. My heartiest gratitude to Dr. Klaus Hildebrandt, whose door was always open for me whenever I ran into trouble or wanted guidance. He always answered my questions with all his heart and with a smile on his face. All the meetings and discussions with him shaped the structure of my thesis project. Next, I would like to thank Prof. Elmar Eisemann for accepting my request to supervise my thesis project. I would like to thank him for his proposal to start the thesis and valuable inputs and ideas during our discussions. His constant support and feedback were extremely valuable throughout the course of the project. I would also like to thank Prof. Nick van de Giesen for entrusting me with this project and helping me with the Water Management side of the project. His feedback and weekly discussions were invaluable to me. Further, this project was cross-disciplinary, however, constant guidance and motivation from all three supervisors contributed significantly to the successful completion of this project.

The biggest gratitude goes to my parents for their endless support and giving me the opportunity to pursue this Masters' degree. Finally, I would like to thank my friends for constantly supporting and motivating me throughout the ups and downs of this journey.

Anshul Khandelwal,
September 2018

Contents

Abstract iii

Acknowledgements v

1 Introduction 1
- 1.1 Cape Town - Day Zero Case Study . . . 1
- 1.2 Motivation . . . 4
- 1.3 Research Goal . . . 6
- 1.4 Challenges . . . 7
- 1.5 Our Contribution - The Geometric Approach . . . 7
- 1.6 Thesis Outline . . . 8

2 Literature Review and Background 9
- 2.1 Related Work . . . 9
  - 2.1.1 Storage Capacities of Reservoirs . . . 9
  - 2.1.2 Computer Graphics techniques . . . 11
- 2.2 Water Management - Definitions and Related Work . . . 12
  - 2.2.1 Reservoirs and its Impacts . . . 12
  - 2.2.2 Global Hydrological Models . . . 13
  - 2.2.3 Digital Elevation Models . . . 14
  - 2.2.4 Volume-Area Curves . . . 15

3 Mathematical Background 17
- 3.1 Laplacian Matrix . . . 17
- 3.2 Least Squares Setting for Quadratic Polynomials . . . 17
  - 3.2.1 Univariate Quadratic Polynomial . . . 18
  - 3.2.2 Bivariate Quadratic Polynomial . . . 18
  - 3.2.3 Multivariate Quadratic Polynomial . . . 18
- 3.3 LDL Solver - MA57 . . . 19
- 3.4 Haversine Formula . . . 20

4 Methodology 21
- 4.1 Data Description . . . 21
- 4.2 Data Preprocessing and Pipeline . . . 23
- 4.3 Algorithmic Setup . . . 26

5 Implementation 29
- 5.1 Laplacian Matrix . . . 29
- 5.2 Quadratic Optimization . . . 31
- 5.3 Maximum Storage Capacities . . . 31
- 5.4 Volume Area Curves . . . 33
- 5.5 Introduction of Control Parameters . . . 34

6 Results and Analysis **35**
6.1 Basic Geometric Model . . . 37
6.1.1 RMSE Analysis . . . 37
6.1.2 Maximum Storage Estimates . . . 37
6.1.3 VA Curves Analysis . . . 41
6.2 Global Kappa parameter . . . 43
6.2.1 RMSE Analysis . . . 43
6.2.2 Relationship between Kappa and size of the reservoirs . . . 46
6.2.3 Analysis of Overestimating Reservoirs . . . 46
6.3 Local Kappa Parameter . . . 48
6.4 Discussion of different reservoir cases . . . 49
6.4.1 Reservoirs with islands . . . 49
6.4.2 Dendritic shaped reservoirs . . . 51
6.4.3 Large Reservoirs . . . 52
6.4.4 Narrow Valley Reservoirs . . . 53

7 Conclusions and Future Work **55**
7.1 Conclusion and Insights . . . 55
7.2 Recommendations . . . 56

A Appendix A: Reservoir Characteristics **57**
A.1 Shape of Reservoirs and JRC Water Mask . . . 57
A.2 RMSE . . . 65
A.2.1 Basic Geometric Model . . . 65
A.2.2 Geometric Model - Optimal Kappa . . . 66
A.3 Maximum Storage Capacities . . . 67
A.4 Volume Area Curves . . . 68

Bibliography **83**

List of Figures

1.1 Percentage of annual normal rainfall (2012-2017) in South Africa 1
1.2 Rainfall in (mm/year) for the region of Cape Town 2
1.3 Volume of water stored in Big Six dams of WCWSS 3
1.4 Gross storage of major dams (in million m^3) in Cape Town 2017-18 . . 4
1.5 List of major cities that may face day zero in the near future 5

2.1 Virtual Dam Approach for Foss Reservoir, United States 10
2.2 Volume-Area Curve for Rules Reservoir, Spain 14
2.3 Illustration of VA curves for Rules Reservoir, Spain 15

4.1 Overview of AQUASTAT database 22
4.2 Illustration of cliiping procedure for HydroSHEDS data 22
4.3 Illustration of Clipping procedure of the JRC Water Occurrence Data . 24
4.4 Data pipeline 26

5.1 Optimization Flowchart 30
5.2 Effect of Kappa on the geometric model 32

6.1 Evaluation phase 35
6.2 Block Diagram of Data flow - Comparison phase 36
6.3 Block Diagram of Data flow - Evaluation phase 36
6.4 Ground Truth vs Geometric Model - Maximum Storage estimates . . . 40
6.5 Ground Truth vs GRanD Model - Maximum Storage estimates 40
6.6 VA Curves for Karkheh reservoir, Iran 41
6.7 VA Curves for Maguga reservoir, Swaziland 42
6.8 Log-Log plot between Kappa and Surface Area 45
6.9 Log-Log plot between Kappa and Surface Area excluding exceptional cases when Kappa = 0 45
6.10 Karkheh Reservoir, Iran 47
6.11 Thaphanseik Reservoir, Myanmar 47
6.12 Illustration of Global Kappa vs Local Kappa 48
6.13 Hum Thuan 1 - Reservoir with island 50
6.14 Hum Thuan 1, Vietnam 50
6.15 Comparison Curves for Dendritic shaped reservoirs. 51
6.16 3D view of Peribonka Reservoir, Canada. Arrow pointing out the area where the model overestimates. 54
6.17 Peribonka - Narrow valley reservoir 54

A.1 Porce 2, Columbia 57
A.2 Peribonka, Canada 57
A.3 Caruachi, Venezuela 58
A.4 Cana Brava, Brazil 58
A.5 Manso, Brazil 58
A.6 Quiemado, Brazil 58

A.7 Paraitinga, Brazil . . . 59
A.8 Quebra Qeixo, Brazil . . . 59
A.9 Itapebi, Brazil . . . 59
A.10 Itioz, Spain . . . 59
A.11 Alqueva, Portugal . . . 59
A.12 Rules, Spain . . . 60
A.13 Kozjak, Macedonia . . . 60
A.14 Capanda, Angola . . . 60
A.15 Mohale, Lesotho . . . 60
A.16 Maguga, Swaziland . . . 61
A.17 Karkheh, Iran . . . 61
A.18 Masjed Solayman, Iran . . . 61
A.19 Karoon 3, Iran . . . 61
A.20 Sardar Sarovar, India . . . 62
A.21 Bansagar Dam, India . . . 62
A.22 Bennithora, India . . . 62
A.23 Nina, China . . . 62
A.24 Khuga, India . . . 63
A.25 Thaphanseik, Myanmar . . . 63
A.26 Jinpen, China . . . 63
A.27 Ham Thuan 1, Vietnam . . . 63
A.28 Da Mi 1, Vietnam . . . 64
A.29 VA Curves 3 . . . 68
A.30 VA Curves 4 . . . 68
A.31 VA Curves 6 . . . 69
A.32 VA Curves 7 . . . 69
A.33 VA Curves 8 . . . 70
A.34 VA Curves 9 . . . 70
A.35 VA Curves 11 . . . 71
A.36 VA Curves 14 . . . 71
A.37 VA Curves 18 . . . 72
A.38 VA Curves 20 . . . 72
A.39 VA Curves 21 . . . 72
A.40 VA Curves 22 . . . 73
A.41 VA Curves 24 . . . 73
A.42 VA Curves 25 . . . 74
A.43 VA Curves 26 . . . 74
A.44 VA Curves 28 . . . 75
A.45 VA Curves 29 . . . 75
A.46 VA Curves 30 . . . 76
A.47 VA Curves 31 . . . 76
A.48 VA Curves 32 . . . 77
A.49 VA Curves 34 . . . 77
A.50 VA Curves 35 . . . 78
A.51 VA Curves 36 . . . 78
A.52 VA Curves 37 . . . 79
A.53 VA Curves 38 . . . 79
A.54 VA Curves 40 . . . 80
A.55 VA Curves 45 . . . 80
A.56 VA Curves 46 . . . 81

List of Tables

4.1 The characteristics of reservoirs considered in this analysis are detailed. 25
4.2 Summary of comparison between Colorization Approach and the Geometric Model 27

6.1 Root Mean Squared Error and Mean Absolute Error between the Computed elevation values and Ground truth elevation values is listed. Surface area and depth for the reservoirs are also given. 38
6.2 Optimal Kappa values computed for the 28 reservoirs at different storage capacities. 44

Chapter 1

Introduction

Water is one of the most essential resource for life on earth. About 79% of the earth's surface is covered with water, out of which only 2.5% constitutes freshwater resources. With climate change gradually becoming a reality and the global population growing exponentially, a large portion of this population is facing water shortage and water management is under stress. About 10,800 cubic kilometers of these freshwater sources, have been impounded in human-made reservoirs since 1900, making the study of reservoirs and their characteristics crucial in terms of water management and hydrology (Chao, Wu, and Li, 2008). These artificial water impoundments have a considerable impact on global processes and hence should be included in their analyses (Dean and Gorham, 1998, St. Louis et al., 2000, Downing et al., 2006).

1.1 Cape Town - Day Zero Case Study

In this section, the case study of the city of Cape Town is presented which analyzes the drought scenario in 2017-18 and elaborates as to how efficient water management practices are essential to avert such issues which threatens to be a major concern in metropolitan cities in the years to come.

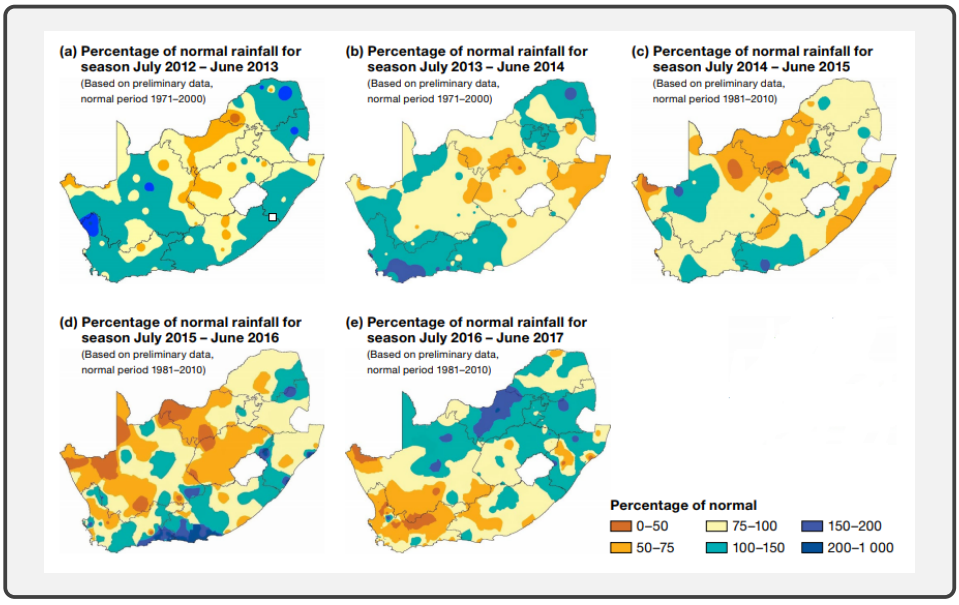


FIGURE 1.1: Percentage of annual normal rainfall from July to June (2012-2017) in South Africa.

Traditionally, South Africa is classified as an arid country with a mean annual precipitation of 497 mm which is considerably less than the world average of 860 mm. The rainfall is unevenly distributed both seasonally and spatially as can be seen in Figure. 1.1[1], which results in sporadic availability of freshwater resources (Smakhtin et al., 2001). In Figure 1.1, a consistent decrease in rainfall can be observed in South Africa from 2012 to 2017, especially in the south western region. Due to this erratic nature of rainfall, in 2018 "Day Zero" was announced for the city of Cape Town - a cosmopolitan city of 3.7 million people - located in the Western Cape province of South Africa which was a global first for a major metropolitan city. Day Zero is the dystopian term for the date by which the region is expected to run out of freshwater resources and hence all the water in taps would run dry.

From 2015 to 2017, rainfall around the region of Cape Town was the least in an 84-year period, which is illustrated in Figure 1.2[2]. The figure shows rainfall from 1933 to 2017 in the Western Cape Water region, from which a considerable decrease below the trend line can be observed for 2015-2017. This reduced rainfall was an important factor that led to the water crisis in the city of Cape Town.

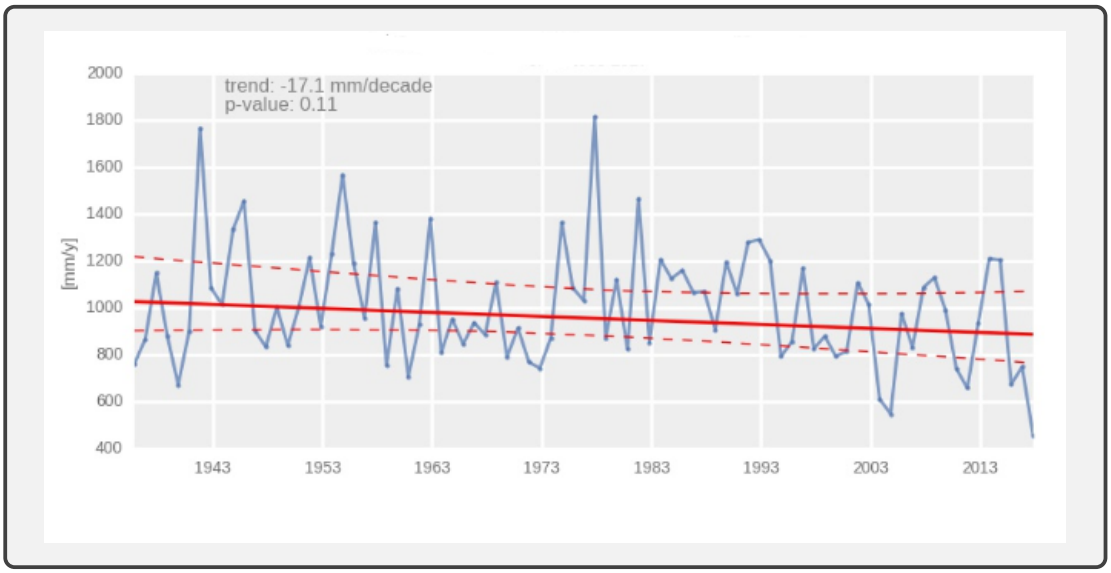


FIGURE 1.2: Rainfall in (mm/year) for the region of Cape Town from 1933 to 2017

However, this situation could have been averted if efficient water management practices were incorporated by the municipality of Cape Town. The first warnings of water crisis and potential distress on water resources in the region of Cape Town was published by the South African Water Research Commission (WRC) in 1990[3]. They predicted that the fresh water resources would dry up in the upcoming decades due to factors such as population explosion and industrialization. Scientists, meteorologists and engineers have raised multiple warnings ever since (Smakhtin et al., 2001), (Basson, Van Niekerk, and Van Rooyen, 1997). In 2009, the Western Cape Water Reconciliation Strategy (WCRS) was put into action in the form of Berg River

[1] South African Weather Service
[2] It's True: Cape Town's Water Supply Is Three Months Away from a Shutdown
[3] WaterCrisis: 1990 article shows Day Zero plans should've began years ago - IOL

dam and various transfer schemes[4], through which water would be transferred from other catchment areas into the Western Cape Water Supply System (WCWSS). City of Cape Town's water conservation and water demand management (WC/WDM) strategy further aimed to significantly reduce its growth in water demand and was projected to solve the crisis till 2019. However, these practices proved to be insufficient, because in 2017, drought hit the city of Cape Town.

This drought scenario was hence attributed to two factors, namely: decreased rainfall and inefficient water management practices. Due to the consistent decrease in rainfall from 2015 to 2018, dam water levels were affected. The decrease in dam water levels between 2015 and 2018 is imminent in Figure. 1.3[5]. Figure. 1.3 shows the estimated volume of water stored in the large six dams of the Western Cape Water Supply System where a gradual and consistent decline can be observed in the water levels of dams. For the City of Cape Town, it was estimated that day zero would arrive when the storage levels of water in the reservoirs of the Western Cape Water Supply System (WCWSS) would reach a level of 13.5% of maximum storage levels. The dam levels reached 19% of maximum capacity at the end of May 2018, hence entering the **Danger Zone**. This water crisis could have been avoided if serious and efficient water management practices would be put into action from 2015.

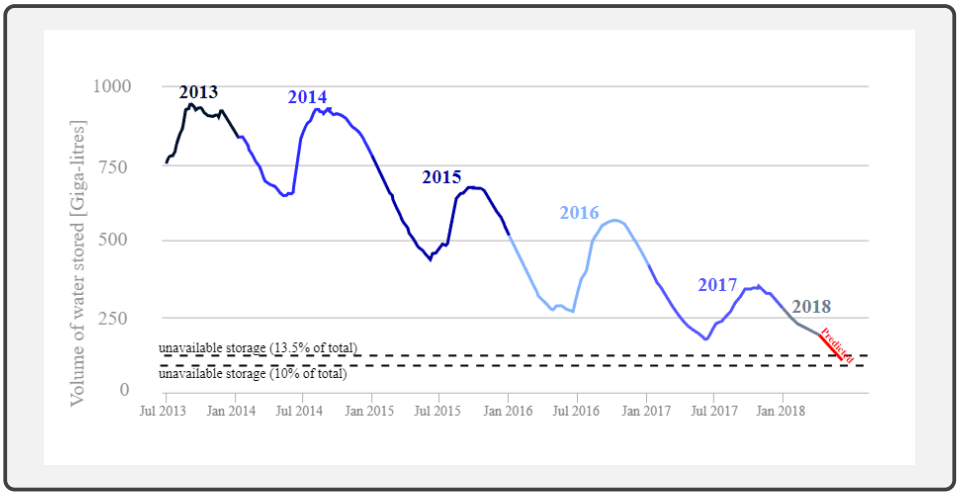


FIGURE 1.3: Volume of water stored in Big Six dams of WCWSS (Source: CSAG)

However, the seriousness of the issue was realized in later stages of 2017 by the City of Cape Town. This resulted in controlled water usage by imposing level 6B restrictions and a new water tariff structure, along with reduced water usage in the agricultural sector. Further, the monsoon season of 2018 saw heavy rainfall in Cape Town, resulting in the dam levels reaching as high as 65 percent in the month of August shown in Figure 1.5[6]. This controlled water usage and heavy rainfall resulted in the day zero being pushed away from Cape Town into 2019.

From this case, two important conclusions can be drawn with respect to water management. Firstly, we need efficient planetary scale models to monitor and assess fresh water resources to avert such water-related crisis. A family of such models

[4] *Western Cape Water Supply System Reconciliation Strategy Study, Overview of Water Conservation and Demand Management in the City of Cape Town. Prepared for the DWAF by UWP Consulting (Pty) Ltd. DWAF Report No. P WMA 19.*

[5] Big Six WCWSS Dams - Climate System Analysis Group

[6] City of Cape Town - Dam Levels Report

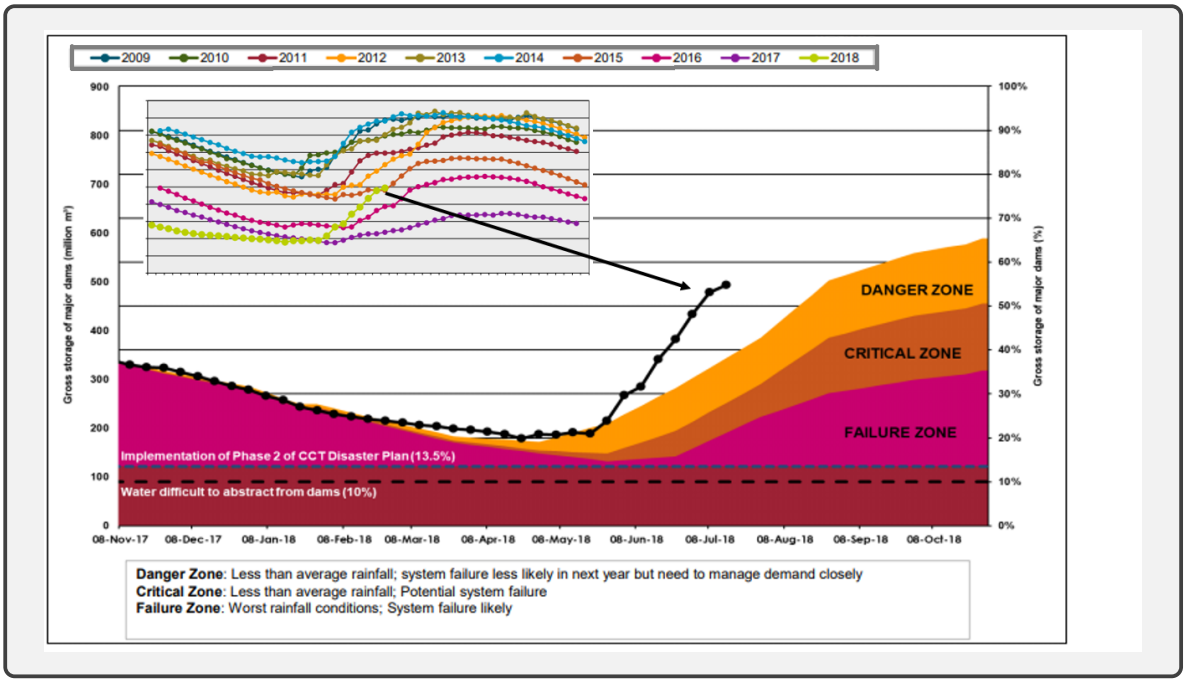


FIGURE 1.4: Gross storage of major dams (in million m^3) in Cape Town 2017-18

known as Global Hydrological Models (GHMs) are being developed to tackle this issue and is hence an active research topic in the field of water management. Further, due to the widespread impact of water stored in reservoirs, reservoir characteristics must be included in these global scale models. Secondly, the models used to predict the storage capacities use a novel volume area relationship that gives a rough estimate of the volume of water stored in the reservoirs[7]. Inaccurate estimations of storage capacities could lead to serious social and financial implications, hence accurate models are required to estimate water available in reservoirs to mitigate such water related crisis in the future.

1.2 Motivation

Due to heavy rainfall in monsoon season of 2018 and control over water usage, temporarily the water crisis scenario has been turned away from Cape Town; however, possible water crisis scenario remains an active problem in other metropolitan cities across the world (Figure 1.5). Satisfying water demands of the ever growing global population whilst simultaneously protecting the hydrological ecosystem will remain a serious challenge for the future generations. This calls for development and research of hydrological models to accurately monitor and assess global water cycles. Further, quantification of human impact is posed as a major challenge in the development of these global hydrological models (Wood et al., 2011). The characteristics of reservoirs on a global scale and its impact have been included recently in the prediction of current and future water availability (Coerver, Rutten, and Giesen, 2018). This has increased the accuracy of prediction of water availability, droughts, and floods; however, better models need to be developed to characterize these reservoirs on a global scale.

[7] Big Six WCWSS Dams - Climate System Analysis Group

FIGURE 1.5: List of major cities that may face day zero in the near future

Lack of information about reservoirs and poor record keeping affects the quality of data making it insufficient for global analysis. The correct representation of the internal states and process dynamics of these reservoirs can often not be verified due to missing in situ measurements. These inaccuracies when combined to model the reservoirs globally can lead to erroneous measurements.

The Global Water System Project (GWSP), a joint project of the Earth System Science Partnership (ESSP), initiated an international effort to gather the existing dam and reservoir data sets with the aim of providing a single, geographically explicit and reliable database for the scientific community: the Global Reservoir and Dam (GRanD) database (Lehner et al., 2011a). The model derived to estimate the storage capacities for large scale reservoirs was linear, novel and a mere linear regression of area and volume values available across the globe. However, this model when used to plot volume-area curves of reservoirs does not give good representations for reservoirs with large storage capacities. The characterization of reservoirs to estimate volume-area curves is computed using the virtual dam placement approach (Bemmelen et al., 2016). However, the model involves subjective human judgement to decide the location of virtual dams and number of virtual dams required to accurately characterize a reservoir. The model is based on the concept of geomorphological homogeneity but geomorphology is a subjective quantity affected by various factors like time, tectonics, geology, and climate (Schumm and Lichty, 1965).

Further, reservoirs also have extensive economic, political, social, and ecological impacts, making their research a platform for further work and applications in the field of water management and hydrology.

1.3 Research Goal

Using freely available satellite based digital elevation models (DEMs) of the surrounding landscape, how can reservoir characteristics such as maximum storage capacities and volume-area curves be estimated for large-scale reservoirs?

This research question leads to the following sub-tasks:

- Design an optimization problem that generates the topology of a large-scale reservoir without user interaction and exploiting available topographic data such as shape of the reservoir and slope of the surrounding landscape?
- Evaluate the geometric model developed, by comparing it to the ground truth data and the results produced by state-of-the-art techniques, both in terms of maximum volume stored and volume-area relationships?
- Based on the performance of the basic geometric model, define control parameters to improve the efficiency of the model and find predictive power of these parameters with objective measures like size of the reservoir and distance to the boundary for a multitude of reservoirs?

1.4 Challenges

This section elaborates the challenges to be addressed throughout the scope of this work.

Ideally, to compute the bathymetry of a landscape under a water body, SONAR techniques are used. However, SONAR techniques require either manpower or instrumentation or a combination of both for carrying out global scale analyses. So conducting global-scale analyses for large-scale reservoirs using SONAR is not feasible with current state of the art satellites. Hence, a model must be developed which can be implemented for reservoirs across the world.

Reservoirs are static in terms of geographic location, however, storage capacities vary drastically depending on the use of dam on which the reservoir is constructed. Hence, the model developed should account for seasonal variations in reservoir capacities. The model built should be able to compute the storage capacities or volume area curves irrespective of the season when the analysis is conducted.

The shape of the valley in which reservoirs are built vary globally. For example, hydroelectric dams are built in narrrow valleys to generate high potential energy and simultaneously the cost of construction of the dam is low. On the other hand, reservoirs built for irrigational purposes are generally spread out and hence dendritic in shape to reach a large number of settlements. Hence, the model developed should account for different reservoir cases and should be independent of the shape or size of the reservoir. For instance, prior work does not account for reservoirs which have islands in the catchment or reservoirs with dendritic shaped catchments.

The most challenging aspect of modeling large-scale reservoirs is the evaluation strategy to assess the accuracy of the modeled reservoirs. This is because of lack of bathymetric data and inefficient record keeping of reservoirs on a global level.

To sum it up, research challenges include global-scale reservoir analysis, incorporation of seasonal changes, implementation on reservoirs of varying shapes and sizes and the evaluation strategy to assess the accuracy of the approach.

1.5 Our Contribution - The Geometric Approach

In this work, a geometric optimization based approach is presented which uses the slopes of the surrounding landscape to extrapolate the underlying topology of the reservoir. This approach depends on the accuracy of elevation values of the surrounding landscape and the mask separating land from water.

In addition, this work uses an elaborate evaluation scheme to test the performance of the approach for reservoirs of different shapes and sizes. The SRTM data available from 2000 is used, which acts as ground truth data for the reservoirs that are built after 2000. Further, the evaluation is performed in terms of RMSE, Maximum storage estimates and Volume-Area curves.

1.6 Thesis Outline

The structure of the thesis is as follows: Chapter 2 provides a concise literature review describing the state-of-the-art techniques used in the field of water management to predict storage capacities. Further, various techniques have been reviewed from a computer graphics perspective that can be used to tackle the research challenges. In addition, water management concepts used throughout the course of this work have been detailed. Chapter 3 lists the mathematical background required to for development of the geometric approach. Chapter 4 describes the datasets that are used for this work and gives a detailed explanation of the methodology, which leads to the formulation of the algorithmic setup. Chapter 5 elaborates on the implementation details of the methodology. Chapter 6 evaluates the model built for the dams built after SRTM was conducted in 2000 in two phases - Comparison and Evaluation. Further, reservoir cases are presented to better analyze the performance of the geometric model as compared to prior work in the field. Finally, Chapter 7 provides a summary of results obtained, final discussions about the approach and discusses as to how this work can be a platform for research in the field of water management and hydrology.

Chapter 2

Literature Review and Background

This chapter details the prior work in the field of water management such as the models used currently to estimate storage capacities of reservoirs and discusses the research gaps associated with these models. Further, relevant literature is discussed in the form of potential computer graphics techniques that can be used for raster based analysis of large-scale reservoirs. In addition, water management concepts relevant to the study of reservoirs such as Global Hydrological Models (GHMs), Digital Elevation Models (DEMs) and Volume-Area curves have been defined and discussed in detail.

2.1 Related Work

In this section, prior work to estimate storage capacities of reservoirs is discussed leading to a comprehensive survey of possible computer graphics techniques that can be used to assess the challenges discussed in the research problem.

2.1.1 Storage Capacities of Reservoirs

Global Lakes and Wetlands Database (Lehner and Döll, 2004) used the model developed by Takeuchi, 1997 to compute missing reservoir storage capacities. The model derived a linear relationship between the gross capacity and the inundated area of reservoirs, under the assumption that the topography of global land surface is covered by various scales of valleys that have a common shape. However, this assumption is not valid for large scale reservoirs as the shape of the valley is affected by subjective factors like time, tectonics, geology and climate (Schumm and Lichty, 1965). Also, the storage estimates by both GLWD 1 and GLWD 2 show high degree of uncertainty due to scaling issues and model inaccuracies. These inaccuracies may arise because a linear relationship is assumed between the gross capacity and the inundated areas of the reservoirs. GLWD 2 comprises of both natural lakes and man-made reservoirs, and because of the relatively high quantity of natural lakes, this dataset is skewed in nature. These characteristics of GLWD calls for a model to analyze artificial reservoirs, separately from natural lakes.

Global Reservoirs and Dams Database (Lehner et al., 2011a) proposes a globally valid average area-volume relationship based on regression analysis of bathymetric data of 5824 reservoirs to estimate missing reservoir storage capacities. This proposed model is linear and skewed in nature as the dataset contains significantly greater number of small dams as compared to large dams. Further, Lehner et al., 2011 states that previous approximations does not separate the lakes from reservoirs in their analysis generating inaccurate estimates in terms of storage capacities. The model considered polygons to separate land from water which were below full

capacity in many cases. Hence the model can be deemed inappropriate to make estimates of reservoirs at varying storage levels across different seasons.

Another approach to predict reservoir capacities and to make accurate estimations of volume-area curves is based on the method of virtual dam placement (Bemmelen et al., 2016). The method uses the aggregate measures of virtual reservoirs built upstream and downstream to derive area-volume relationships of the existing reservoir. An illustration is shown in Figure. 2.1, in which virtual reservoirs are shown in light blue and the reservoir for which volume-area relationship is to be generated is shown in purple. However, this method involves subjective human judgement to decide the location of the virtual reservoirs and the number of virtual dams required to accurately characterize a reservoir. Another drawback of the method is that the model is based on the concept of homogeneity of the landscape, because of which extending this study for global-scale analyses would not be feasible.

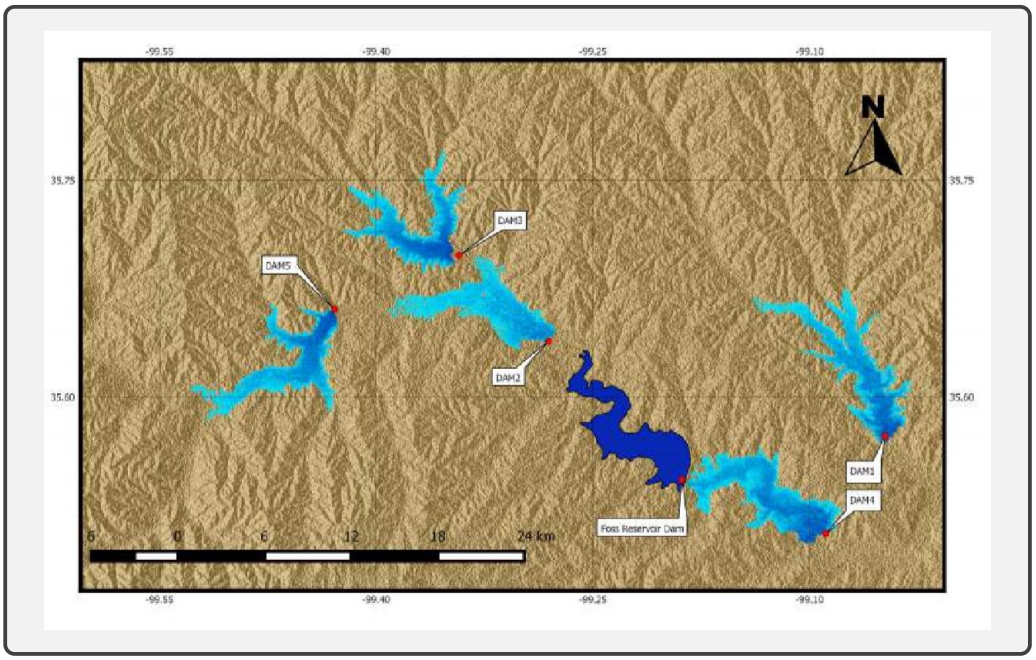


FIGURE 2.1: Virtual Dam Approach for Foss Reservoir, United States (Bemmelen et al., 2016)

HydroLAKES[1] is a comprehensive database that develops a novel geo-statistical model to produce consistent volume estimates of lakes and reservoirs on a global scale using the assumption that the land surface topography surrounding a given water body can be used as a predictor for bathymetry (Messager et al., 2016). Messager et al., 2016 states that the role of lakes and reservoirs in the global hydrological and biochemical cycles is intimately tied to the geometric characteristics of surface area, depth, stored water volume and shoreline length. It is used in combination with the elevation data provided by EarthEnv-DEM90 (Robinson, Regetz, and Guralnick, 2014) at 90m resolution to calculate estimates of average depths for every polygon. The selected prediction model applies size-specific multiple regression equations using surface area and the average terrain slope within a 100m buffer surrounding the water body. The equations developed were based on bathymetric data records for more than 7000 lakes globally, and the results were tested against

[1] HydroLAKES - Technical Documentation

independent validation data of more than 5000 lakes. The validation confirmed satisfying regional results, but interpreting the volume of singular water bodies was not satisfactory as individual errors and uncertainties were projected to be large. Further, this model was developed for lakes and reservoirs combined; however, the two ecosystems lentic and lotic are different aquatic systems in terms of physical, biological and chemical characteristics[2], which calls for a separate model for these ecosystems.

2.1.2 Computer Graphics techniques

Following the research gaps discussed to characterize reservoirs globally, we hypothesize that the surrounding landscape properties such as slopes of the neighboring raster cells are important determinants to understand the underlying landscape of a reservoir. Hence, potential computer graphics techniques that can incorporate these neighborhood landscape properties to overcome the research challenges have been detailed.

In order to incorporate slopes of surrounding landscape for modeling, gradient-domain filters can be used that manipulate pixel differences such as first order image-gradients in addition to pixel values of an image. This challenge is formulated as gradient-domain filtering in the computer graphics community and uses graph Laplacian matrices which are solved using discrete Poisson equations. Examples in two-dimensional space include gradient-domain tone mapping (Fattal, Lischinski, and Werman, 2002), Poisson blending (Pérez, Gangnet, and Blake, 2003), alpha matting (Sun et al., 2004), image colorization (Levin, Lischinski, and Weiss, 2004), tonal adjustment (Lischinski et al., 2006), edge-preserving smoothing (Farbman et al., 2008), and image relighting and non photo realistic rendering (Bhat et al., 2010). Three-dimensional geometric processing applications include mesh segmentation (Liu and Zhang, 2007) and geodesic distance computation (Crane, Weischedel, and Wardetzky, 2013).

Colorization using optimization (Levin, Lischinski, and Weiss, 2004) is the computer assisted process of adding color to a monochrome image or movie. This method was developed using a quadratic cost function to obtain an optimization problem based on the premise that neighboring pixels in space and time should have similar intensity values. This cost function is represented using the Laplacian matrix which is optimized using relevant quadratic optimization techniques. This algorithm represents a wide spectrum of tasks in the field of image processing.

In image segmentation algorithms based on normalized cuts (Shi and Malik, 2000), second smallest eigenvector of the graph Laplacian matrix is used. The graph Lapalcian matrix is defined as $A = D - W$ where W is the *Adjacency Matrix* and D is the *Degree Matrix* . The second smallest eigenvector of any symmetric matrix A is a unit norm vector x that minimizes $x^T Ax$ and is orthogonal to the first eigenvector. Thus, this algorithm optimizes the same cost function as (Levin, Lischinski, and Weiss, 2004) but under slightly different constraints. In image denoising algorithms based on anistropic diffusion a similar cost function is minimized (Perona and Malik, 1990, Tang, Sapiro, and Caselles, 2001).

The optimization formulation developed in this work is similar to the method adopted by Levin, Lischinski, and Weiss, 2004, in which the intensity pixel values can be inferred as raster elevation values, the colored annotations can be inferred as the surrounding landscape to a reservoir and the resulting colored image can be inferred as the underlying landscape of the reservoir.

[2] Physical, Chemical and Biological Characteristics of Lentic and Lotic Ecosystems

2.2 Water Management - Definitions and Related Work

In this section, reservoirs are introduced and their widespread economic, political, social and environmental impacts are discussed. Further, this section presents the case of importance of global assessment and monitoring of reservoirs, following which, Global Hydrological Models (GHMs) are introduced. The research gaps related to incorporation of anthropogenic impacts of reservoirs on these GHMs have been detailed. Finally, Digital elevation Models (DEMs) and Volume-Area curves are defined which are the input features and output curves to the geometric model introduced in this work.

2.2.1 Reservoirs and its Impacts

Reservoirs are artificial impoundments, enlarged lakes or storage ponds built to store water for various purposes like water supply - domestic or industrial use, irrigation for agriculture, hydroelectricity and flood control. As of 2011, it was estimated that 16.7 million reservoirs larger that 0.01 ha, with a combined gross capacity of approximately 8070 km^3, existed worldwide (Lehner et al., 2011b). Because of the large volume of freshwater sources in these reservoirs, they have widespread economic, societal, political and environmental impacts making their global-scale analyses crucial and challenging.

Large-scale reservoirs have widespread economic impacts such as impacts on employment, trade, agriculture, fishery, real estate prices (Bhatia et al., 2008). For example, the High Dam at Aswan, Egypt was built in late 1950s, which led to an employment surge from 2.9$thousand$ to 5.7$thousand$ and the value of production also quadrupled from 3.5$million$ to 14.0$million$ in a period from 1957 to 1961. Further, in 1997 the value of the high Aswan Dam to the Egyptian economy was estimated to account for a total gain of EGP 7.1 billion to 10.3 EGP billion which equals to 2.7% to 4.0% of the annual GDP of Egypt (Strzepek et al., 2008).

Social impacts of large scale reservoirs include the migration and resettlement of people near the dam sites, changes in the rural economy and employment structure, effects on infrastructure and housing, and community heath. Large-scale reservoirs also have political implications. For example, the construction of The Three Gorges Dam in China provoked a political unrest due to relocation of over 1.3 million people, even though the project was expected to boost the economy (Jackson and Sleigh, 2000). In Sweden, Norway and Switzerland, large public opposition has stopped the construction of numerous projects (Dorcey et al., 1997). A brief description about the history and politics of dam building worldwide is discussed by McCully, 1996 and gives an account as to why dam building is one of the most controversial technologies.

Finally, ecological impacts of reservoirs include changing sea level, erratic seasonal patterns, changes in delivery of sediments and changes in bio-aquatic connectivity of river streams. Further, reservoirs are directly responsible for generation of hydroelectricity, which accounted for 71% of the total electricity generated by renewable sources in 2016, according to the World Energy Council [3]. However, global warming emissions produced during the installation and dismantling of hydroelectric power plants when combined with the facility's operation can be significant. Further, dammed reservoirs affect the land use pattern and have a major impact on the aquatic ecosystems[4].

[3] World Energy Resources Hydropower, 2016
[4] Environmental Impacts of Hydroelectric Power

Characterization of reservoirs is of further interest to hydrologists, aquatic biologists, river and coastal morphologists and biochemists (Bemmelen et al., 2016). Despite the widespread importance of reservoirs, global datasets and hydrological models describing their characteristics and geographical distribution are incomplete. The lack of knowledge of the number of existing reservoirs, their distribution, and their storage volumes hinders efficient water management and reservoir planning. These large-scale impacts of reservoirs make their study and analysis crucial, both from an engineering and research perspective.

2.2.2 Global Hydrological Models

Global Hydrological Models (GHMs) are hydrological models which are developed with the aim of modeling real-world hydrological processes. GHMs can also be defined as hydrological models that incorporate spatial and temporal data at coarse resolutions that help us better understand, predict and manage water resources on a global scale. The advent of GHMs over the course of the last two decades, is primarily because of the explosion of global scale data availability such as low-resolution hydrological states and high-resolution parameter fields from remote sensing satellites (Tang et al., 2009). This when combined with improved computational capabilities and data storage have provided a platform for the development of high-resolution GHMs (Coerver, Rutten, and Giesen, 2018). GHMs are prevalently used to forecast drought or flood related scenarios to mitigate disasters. However, widespread uses of GHMs include describing global economy, ecology, trade, biodiversity, energy balance, land-use change detection, climate change and crop growth (Sood and Smakhtin, 2015) (Bierkens et al., 2015) (Wood et al., 2011).

The input features into GHMs can be broadly classified into:

- Natural features such as climate, temperature data which is spatio-temporal in nature.
- Anthropogenic features such as reservoir operational characteristics, reservoir storage.

Formerly, anthropogenic impacts were not incorporated in GHMs due to data unavailability and the fact that modeling reservoir characteristics in terms of inflow-outflow data and storage capacities was complicated. However, reservoirs have significantly altered the hydrological cycle such as reduced global sea level (Fiedler and Conrad, 2010), changing seasonal patterns (Haddeland, Skaugen, and Lettenmaier, 2006), change in delivery of sediments (Syvitski et al., 2005) and change in bioacquatic connectivity of river systems (Pringle, 2001). Because of these widespread impacts, a new generation of GHMs aim to take impacts of reservoirs into account, in fact, incorporation of reservoirs has been posed as a major challenge in the development of GHMs (Wood et al., 2011).

Examples of prevalently used GHMs include PCRGLOB-WB [5], waterGAP 3 (Döll, Fiedler, and Zhang, 2009), WBMplus (Wisser et al., 2010), SWBM (Orth and Seneviratne, 2013), WR3A (Dijk et al., 2014), HBV-SIMREG (Beck et al., 2016) and eWaterCycle[6]. A review of algorithms currently used in GHMs calls for inclusion of water resource management in these earth system models (Nazemi and Wheater, 2015). Lack of representation of open water bodies such as lakes and reservoirs on a grid

[5] PCRGLOB-WB
[6] E-Water Cycle

and sub-grid scale has also been addressed. Macro-scale algorithms, generally cannot rely on detailed information on reservoir operation policies used in small-scale models; hence reservoir release and storage have been proposed as viable alternatives. A simulation based scheme is proposed that uses storage capacity, purpose, simulated inflow and downstream demand of a reservoir for individual reservoirs in global river routing models (Hanasaki, Kanae, and Oki, 2006).

In conclusion, global modeling of various ecosystems, especially hydrology is an active research topic in the field of water management and inclusion of anthropogenic impacts is posed as one of the biggest challenges. For incorporating these anthropogenic impacts, reservoir characteristics such as reservoir operating characteristics and reservoir storage capacities are essential. Increasing the accuracy of these models on a global scale can drastically improve the efficiency of Global Hydrological Models.

2.2.3 Digital Elevation Models

Digital Elevation Models (DEMs) are a representation of the terrain's surface obtained in the form of elevation values via remote sensing techniques. They are available as spatially geo-referenced datasets that encode topographical features for environmental modeling purposes. DEMs provide an accurate representation of the landscape features depending on the spatial resolution of the data. DEMs can be obtained in two formats: raster based DEMs and vector based DEMs; however for this work raster based DEMs are used.

Some examples of widely used DEMs include Shuttle RADAR Topography Mission (SRTM) (Farr et al., 2007), Advanced Spaceborne Thermal Emission and Reflectance Radiometer Global DEM (ASTER GDEM) (Tachikawa et al., 2011), ALOS PRISM (Tadono et al., 2014), Global multi-resolution terrain elevation data 2010 (GMTED2010) (Danielson and Gesch, 2011), to name a few. Further, detailed analysis of global and regional DEMs can be found at Gamache, 2004.

Digital Elevation Models can be used for a wide range of geological applications including surface analysis, satellite navigation, terrain analysis in geomorphology, among others. For this work, we shall use Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales (HydroSHEDs) (Lehner, Verdin, and Jarvis, 2006) because of global scale coverage at high resolution (3 arc seconds).

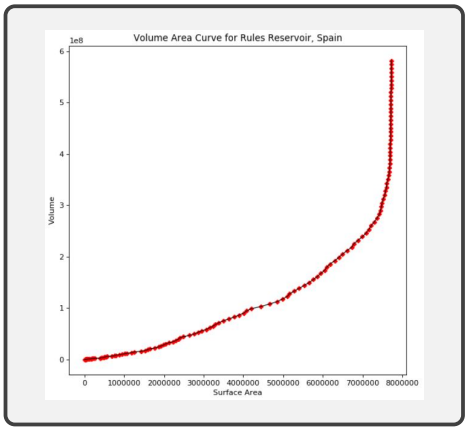
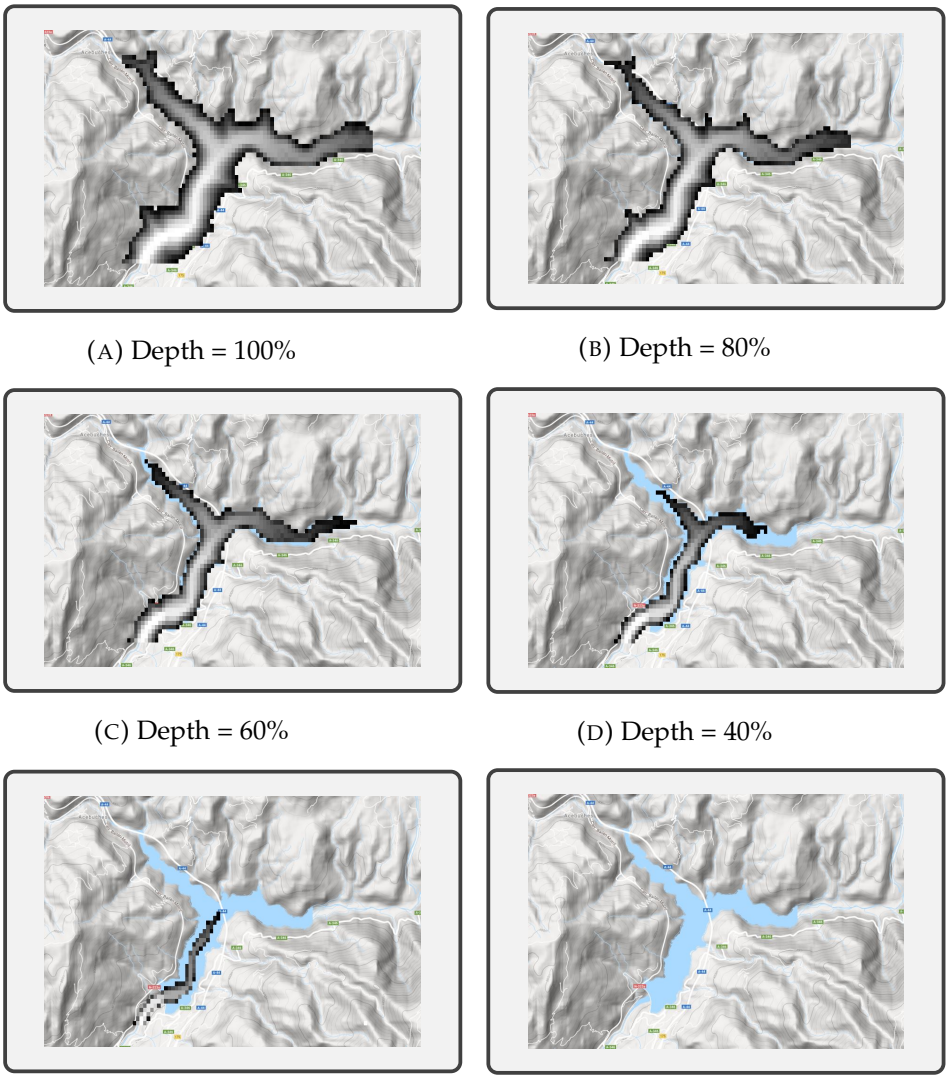


FIGURE 2.2: Volume-Area Curve for Rules Reservoir, Spain



(A) Depth = 100% (B) Depth = 80%

(C) Depth = 60% (D) Depth = 40%

(E) Depth = 20% (F) Depth = 0%

FIGURE 2.3: Illustration of VA curves for Rules Reservoir, Spain

2.2.4 Volume-Area Curves

To survey a cavity or a water body bathymetrically, it is important to understand the shape of the valley. Volume Area curves are representations used to give an idea about the shape of the cavity or valley. These curves are obtained by virtually filling the water body to the maximum storage capacity and releasing the water stored in terms of depth 1m at a time. Capacity of the reservoir and surface area are plotted from maximum to minimum depth and the curves obtained are called VA curves. As, VA curves help in better understanding of the shape of a water body they have great significance practically. An illustration of the process of development of VA curves for Rules reservoir in Spain is shown in Figure 2.3. The figure illustrates the emptying process of the reservoirs starting from 100% capacity (Figure 2.3a) to 0% capacity (Figure 2.3f). The resulting VA curve for the Rules reservoirs is shown in Figure 2.2.

Volume area curves are specifically useful to understand reservoir processes such as reservoir flood routing, dam operation, reservoir classification, evaporation from reservoirs and prediction of sediment distribution in reservoirs. In addition, to simulate the effects of reservoirs, depth-storage curves and area-volume relationships are required but challenging to obtain (Wood et al., 2011). Area-volume relationships are essential as they allow operational observation of stored volumes by observing areas through satellite remote sensing (Liebe et al., 2009, Eilander et al., 2014).

For this work, VA curves of individual reservoirs are computed and used to evaluate the performance of the geometric model. VA curves are obtained both from the ground truth elevation values obtained from DEMs and the geometrically computed elevation values and then compared to understand if the method was successfully implemented to predict the shape of the underlying topology.

Chapter 3

Mathematical Background

In this chapter, the mathematical background is discussed which introduces the laplacian matrix which is an integral part of the algorithm. Following, the least squares setting based on univariate, bivariate leading to n-dimensional variables is detailed. In addition, LDL Solver and Haversine formula are discussed.

3.1 Laplacian Matrix

The Laplacian matrix (Cvetković, Doob, and Sachs, 1998) (Babić et al., 2002) is used in a multitude of applications ranging from spectral graph theory to image processing. In this work, Laplacian matrix $L \in \mathbb{R}^{nxn}$ of a 2D GeoTIFF image (I) with n raster cells is defined as:

$$L_{ij}(I) = \begin{cases} \sum_{j \in N(i)} w_{ij} & i \leq j \\ -w_{ij} & j \in N(i) \\ 0 & elsewhere \end{cases}$$

where, i is the raster taken in consideration, $N(i)$ are the neighboring raster cells to i which are denoted by j and w_{ij} are the weights between i and its corresponding neighbor. We take the weights w_{ij} to be unity. The properties of Laplacian matrix required for the scope of this work are as follows:

- Laplacian matrix is symmetric.
- Laplacian matrix is positive-semidefinite.

An example of Laplacian matrix for a 2-D 3x3 image is:

$$\begin{bmatrix} 2 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 3 & -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 3 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 2 \end{bmatrix}$$

3.2 Least Squares Setting for Quadratic Polynomials

In this section we define a least squares problem for a quadratic polynomial which is used as an integral part of the algorithm defined in the next chapter. First, a quadratic polynomial is defined for a univariate case which is then extended to the

case of two variables, following an n-dimensional case is developed using the Laplacian matrix discussed in the previous section.

3.2.1 Univariate Quadratic Polynomial

The general form of a univariate quadratic equation is as follows:

$$f(x) = \frac{1}{2}ax^2 + bx + c \tag{3.1}$$

$$\text{where } a, b, c \in \mathbb{R}$$

The critical points of this equation satisfies the linear equation obtained by setting the derivative to zero.

$$\frac{d}{dx}f(x) = ax + b = 0 \tag{3.2}$$

If a > 0, the solution thus obtained is the minimum value of $f(x)$ and if a < 0, the solution obtained is maximum value of $f(x)$.

3.2.2 Bivariate Quadratic Polynomial

The general form of the bivariate quadratic polynomial is as follows:

$$f(x_1, x_2) = a_1{x_1}^2 + a_2x_1x_2 + a_3{x_2}^2 + b_1x_1 + b_2x_2 + c \tag{3.3}$$

This equation can be written in the matrix form as:

$$f(x_1, x_2) = \frac{1}{2}\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 2a_1 & a_2 \\ a_2 & 2a_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 & b_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + c \tag{3.4}$$

or,

$$f(x) = \frac{1}{2}x^TAx + B^Tx + C \tag{3.5}$$

$$\text{where } x \in \mathbb{R}^2$$

The critical points of this equation satisfies the linear equation obtained by setting the derivative to zero.

$$\frac{d}{dx}f(x) = Ax + B = 0 \tag{3.6}$$

And, the solution to the equation is:

$$x = \begin{cases} minimum & ;if\ A\ is\ positive\ definite \\ maximum & ;if\ A\ is\ negative\ definite \end{cases}$$

3.2.3 Multivariate Quadratic Polynomial

The bivariate equation stated in 3.5 can be written in multivariate case (n-dimensional) as:

$$f(x) = \frac{1}{2}x^TAx + B^Tx + C \tag{3.7}$$

where $x \in \mathbb{R}^n$

Similar to the bivariate case, the critical points of this equation satisfies the linear equation obtained by setting the derivative to zero.

$$\frac{d}{dx}f(x) = Ax + B = 0 \tag{3.8}$$

And, the solution to the equation is:

$$x = \begin{cases} minimum & ;if\ A\ is\ positive\ definite \\ maximum & ;if\ A\ is\ negative\ definite \end{cases}$$

A particular example of the multivariate quadratic polynomial, that we will use in the following sections, can be constructed using the Laplacian matrix ($L \in \mathbb{R}^{nxn}$) and constant matrix ($k \in \mathbb{R}^m$):

$$\begin{aligned} f(x) &= \frac{1}{2}||Lx - k||^2 \\ &= \frac{1}{2}(Lx - k)^T(Lx - k) \\ &= \frac{1}{2}x^T L^T Lx - k^T Lx + \frac{1}{2}k^T k \end{aligned} \tag{3.9}$$

Interpreting this example in the general form Eqn. 3.7, we see that:

$$A = L^T L, B = -L^T k, C = k^T k$$

Hence, the minimizer of the quadratic function in Eqn. 3.9 satisfies the equation:

$$L^T Lx = L^T k \tag{3.10}$$

3.3 LDL Solver - MA57

LDL solver solves a linear equation using the LDL factorization of matrix A from Eqn. 5.2. The factorization is as follows:

$$A = LDL^T \tag{3.11}$$

where L is a lower triangular square matrix with unity diagonal elements, D is a diagonal matrix and

Hence, combining Eqn. 5.2 and Eqn. 3.11, we get:

$$LDL^T * X = B \tag{3.12}$$

This equation is solved using the following steps:

- Substitute $Y = DL^T X$
- Substitute $Z = L^T X$
- Hence, we get $LY = R$, $DZ = Y$ and $L^T X = Z$.

These are one diagonal and two triangular systems which are solved independently. Similar to this factorization, the MA57 code is implemented using a multifrontal approach (Duff, 2004) to obtain a direct solution of sparse symmetric linear equations. This MA57 algorithm is used with the backslash operator in MATLAB as LDL solver.

3.4 Haversine Formula

Haversine formula is used to determine the great-circle distance between two points in a sphere given their geo coordinates. This formula is based on the law of haversines that relates the sides and angles of spherical trangles in spherical trignometry. The haversine function is defined as:

$$haversine(\theta) = \sin^2(\theta/2) \tag{3.13}$$

The haversine of the central angle between two vectors is defined as:

$$haversine(d/r) = hav(\psi_2 - \psi_1) + cos(\psi_1)cos(\psi_2)hav(\lambda_2 - \lambda_1) \tag{3.14}$$

where, d is the required distance between the points on the sphere, r is the radius of the sphere, ψ_1 and $\psi 2$ are the latitudes of the two points (in radians), λ_1 and λ_2 are the longitudes of the two points (in radians).

Rearranging the Eqn 3.14, we get:

$$d = r \times c \tag{3.15}$$

$$c = 2 \arctan 2(\sqrt{a}, \sqrt{1-a}) \tag{3.16}$$

$$a = \sin^2((\psi_2 - \psi_1)/2) + \cos(\psi_1) \times (\cos \psi_2) \times \sin^2((\lambda_2 - \lambda_1)/2) \tag{3.17}$$

Overall, this formula is used to find the distance between two points and hence the resolution of Digital Elevation Models which varies significantly when we go from the equator to the poles.

Chapter 4

Methodology

This chapter introduces the various sources of data such as AQUASTAT, HydroSHEDs and JRC used in this work, following which preprocessing and flow of data is discussed. In addition, the geometric model is elaborated starting from the basic premise of the approach leading to the algorithmic setup. The algorithmic setup discussed leads to the quadratic optimization technique used to address the least squares problem that arises.

4.1 Data Description

For this work, three datasets are essentially required for the successful implementation and evaluation of the algorithm.

- A global dataset of large reservoirs including their geo-coordinates and the year they were constructed or operational in. This data is used to separate the reservoirs built after 2000 and before 2000 for evaluation purposes.
- Digital Elevation Models (DEMs) which essentially contains elevation values of rasters in and around the reservoir in consideration, and
- A land-water separation mask to segregate the land raster cells from water raster cells.

AQUASTAT[1] is a geo-referenced database which contains information about dams and their associated reservoirs on a global scale. AQUASTAT was developed using the International Commission on Large Dams[2] data, national reports and information obtained from national experts through surveys. However, this dataset is neither complete nor error-free; but gives an idea about the location of global dams built after 2000. AQUASTAT data is available in the form of spreadsheets for different continental regions which are combined to yield a comprehensive global dataset which is used in this work. A broad overview of this dataset can be seen in Figure 4.1, where the red dots represent the location of 572 reservoirs built after 2000. Two particular reservoirs extracted in Figure 4.1 are as follows:

1. Rules reservoir, Spain (up) and,
2. Porce II reservoir, Columbia (down)

indicating that the geo-coordinates of the reservoirs represented by the AQUASTAT dataset are indeed pretty accurate.

[1] AQUASTAT - A geo-referenced dams database
[2] ICOLD - International Commission on Large Dams

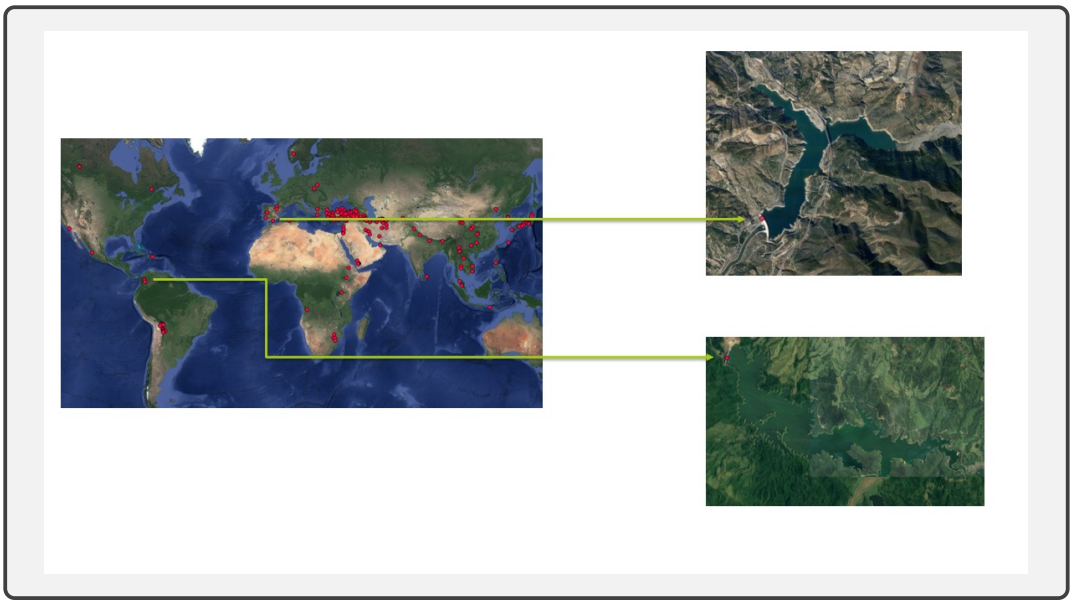


FIGURE 4.1: Overview of AQUASTAT database of reservoirs built after 2000.

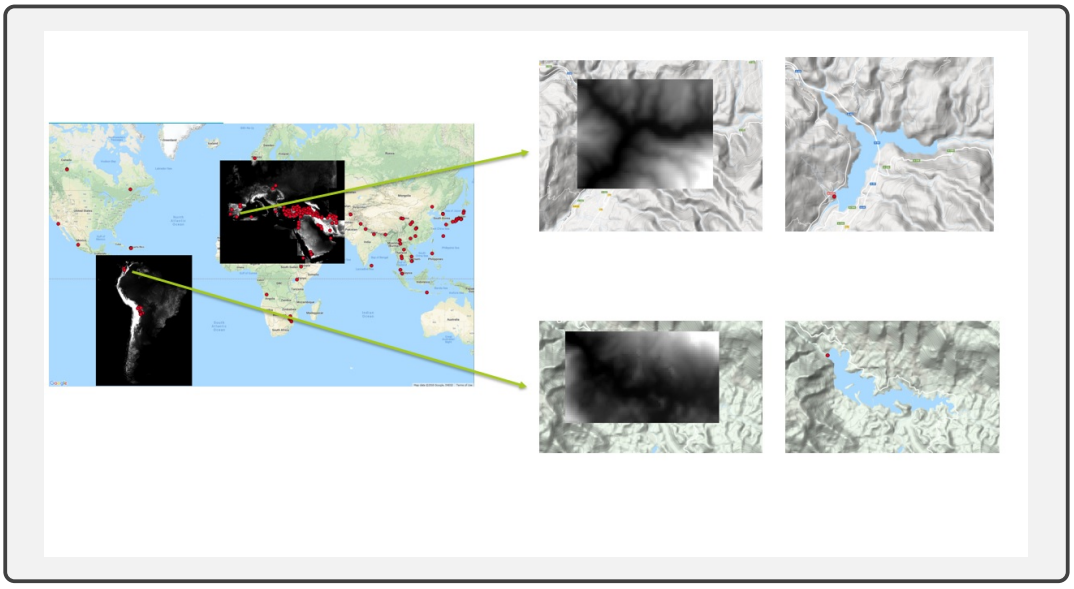


FIGURE 4.2: Illustration of cliiping procedure for HydroSHEDS data

Having acquired the geo-location of reservoirs, elevation values in the form of Digital Elevation Models (DEMs) are required for these reservoirs at the highest possible resolution. For this purpose, **HydroSHEDs** (**Hydro**logical data maps based on **SH**uttle **E**levation **D**erivatives at multiple **S**cales) is used, which is based on elevation data obtained by the Shuttle Radar Topography Mission (SRTM) at 3 arc second (approx. 90m at equator). These void-filled DEMs based on the original SRTM-3 and DTED-1 elevation models of SRTM are referenced to the WGS84/EGM96 geoid and is available on a global scale in raster format (Lehner, Verdin, and Jarvis, 2008) which is essential for this work. Based on the geo-location of the reservoirs obtained from the AQUASTAT database, the DEM for the individual reservoirs are clipped. Figure 4.2 shows the original DEMs of South America and Europe on the left. The darker pixels have relatively lower elevation values whereas the lighter pixels have relatively higher elevation values. These elevation values are computed with respect to the vertical datum i.e. EGM94. Further, on the right side of the Figure. 4.2 clipped DEMs for Rules reservoir, Spain (Up) and Porce II reservoir, Columbia (Down) can be seen. An overview of all clipped DEMs used in this work can be found in the appendix.

As the DEMs are acquired for reservoirs built after 2000, they contain elevation values of raster cells within the reservoir as well as elevation values of the surrounding raster cells. Hence, a water mask is required that separates the land raster cells from the water raster cells. For this purpose, water occurrence data developed by the European Union's Joint Research Commission is used. This water dataset broadly maps the location and temporal distribution of water surfaces on a global scale over the period of 1984 to 2015 using the Landsat imagery obtained from United States Geological Survey (USGS) and NASA. The spatial and temporal variations in the presence of surface water are captured in the form of a single dataset called **Surface Water Occurrence**. This dataset contains the probability of water occurring in the particular raster cell and can be extracted from the Global Surface Water Explorer[3]. Similar to Figure 4.2, Figure 4.3 shows the global Surface Water Occurrence data on the left and extracted data for Rules Reservoir, Spain (Up) and the Porce II Reservoir, Columbia (Down) on the right. In this data, all colored pixels are taken as water raster cells and all white pixels are considered as land raster cells, hence creating a land-water separation mask.

4.2 Data Preprocessing and Pipeline

This section briefly describes the preprocessing performed on the datasets described in the previous section and gives a block diagram of data flow for the scope of this work.

Combining the AQUASTAT data from different continents, a total of 7247 reservoirs were obtained globally. From this comprehensive list, 572 reservoirs built after 2000 were extracted. Further, reservoirs with height less than 15 meters were removed, as we consider only large reservoirs for this analysis[4]. Further, it was observed that the dams built for hydroelectric purposes have relatively narrow valleys, as cost of construction of narrow valleyed dams is considerably low. These narrow valleyed reservoirs were also removed from analysis as number of raster cells covered by these reservoirs were pretty low. Some reservoirs were operational after

[3] Global Surface Water Explorer
[4] ICOLD - Definition of a Large Dam

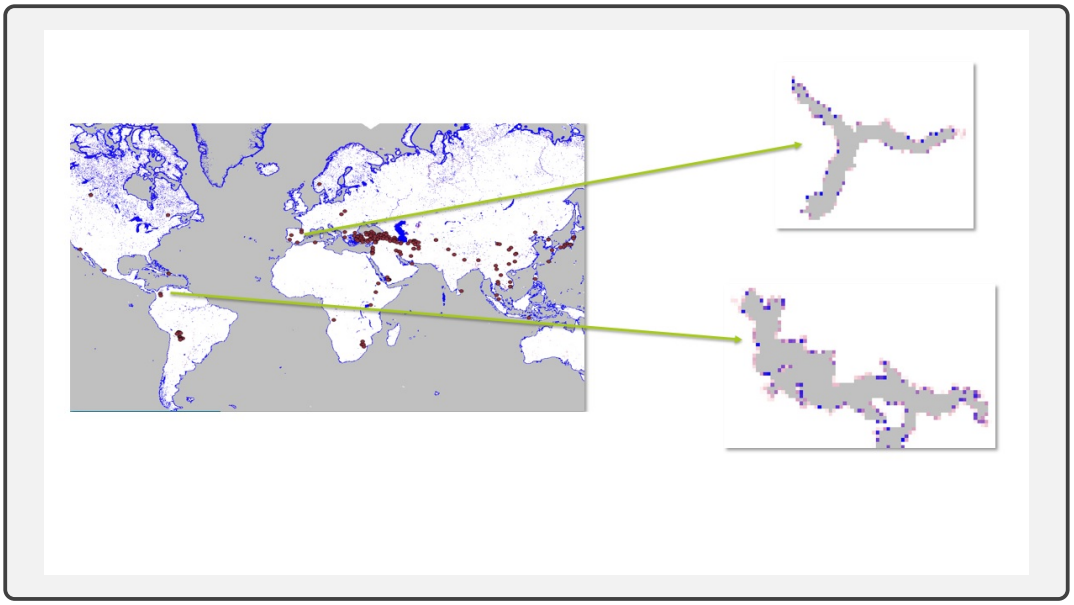


FIGURE 4.3: Illustration of Clipping procedure of the JRC Water Occurrence Data

2000 but their construction began before 2000 and hence the elevation values inside the water mask were unavailable.

These preprocessing steps yielded a database of 28 reservoirs; an overview of which is in Table 4.1. Throughout the course of this work, DEMs will be referred as HydroSHEDs data and the water-land separation masks will be referred as the JRC data. The preprocessing stage includes the clipping of HydroSHEDs and JRC data for the 28 reservoirs mentioned in Table 4.1. This preprocessing was done using QGIS[5] and an example of the resulting clipped images can be seen in Figure 4.2 and Figure 4.3. Further, the resolutions of the HydroSHEDs and JRC DEMs were different. Hence, these two raster DEMs were aligned using the **Align Rasters** tool[6] in QGIS using the nearest neighbor re-sampling approach. The resampling approach was only performed on the JRC data. Overall, at the end of the preprocessing stage, a database of 28 HydroSHEDs and JRC based raster files are obtained at 3 arc second resolution. An overview of these clipped images can be found in the appendix.

Figure 4.4 provides a block diagram of data flow for this work. For a particular reservoir, the HydroSHEDS dataset is segregated into two parts using the JRC dataset. The elevation values outside the water mask is the input to the algorithm, which in the block diagram is also termed as **landscape data** and the elevation values inside the water mask acts as the ground truth for the reservoir. Following, the surrounding landscape data is fed into the algorithm which computes the estimated elevation values using the quadratic optimization approach discussed in the next section. Finally, the **Ground truth data** and **output data** are compared using maximum storage capacities and volume-area curves to evaluate the performance of the proposed model.

[5] QGIS - An open source Geographic Information Systems
[6] Align Rasters Tool, QGIS

TABLE 4.1: The characteristics of reservoirs considered in this analysis are detailed.

Details of reservoirs				
ID	Reservoir Name	Country	Surface Area	Depth
3	Porce 2	Colombia	7,715,803	150
4	Peribonka	Canada	24,854,261	262
6	Caruachi	Venezuela	290,446,217	174
7	Cana Brava	Brazil	99,273,071	138
8	Manso	Brazil	325,388,827	177
9	Queimado	Brazil	17,026,743	97
11	Paraitinga	Brazil	5,047,028	27
14	Quebra Queixo	Brazil	5,826,459	77
18	Itapebi	Brazil	53,329,549	251
20	Itoiz Dam	Spain	7,978,982	150
21	Alqueva	Portugal	169,539,376	200
22	Rules	Spain	3,351,000	141
24	Kozjak	Macedonia	8,189,202	154
25	Capanda	Angola	161,247,179	224
26	Mohale	Lesotho	12,261,889	204
28	Maguga	Swaziland	7,930,496	104
29	Karkheh	Iran	99,861,622	142
30	Masjed Solayman	Iran	44,131,687	596
31	Karoon 3	Iran	15,236,540	168
32	Sardar Sarovar	India	27,124,529	119
34	Bansagar	India	483970455.8	180
35	Bennithora	India	16240501.06	35
36	Nina	China	2061054.772	56
37	Khuga	India	6237827.739	38
38	Thaphanseik	Myanmar (Burma)	376369244.2	117
40	Jinpen	China	2205357.219	112
45	Ham Thuan 1	Vietnam	23380857.4	165
46	Da Mi 1	Vietnam	5607964.907	82

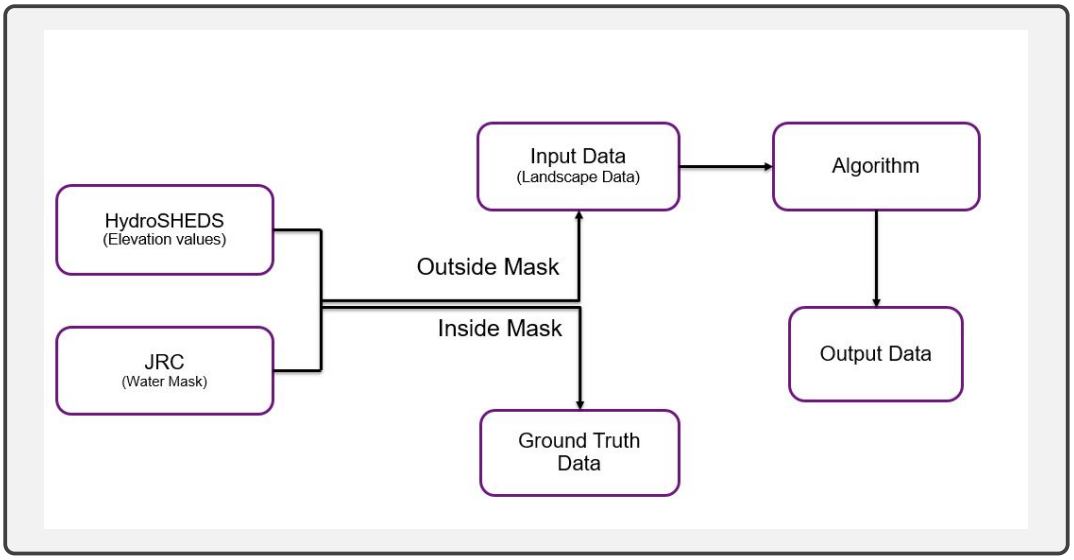


FIGURE 4.4: Data pipeline

4.3 Algorithmic Setup

In this section, the geometric approach is discussed in detail which leads to the development of the algorithmic setup.

The geometric approach is motivated from the Colorization using Optimization setting (Levin, Lischinski, and Weiss, 2004). The colorization problem is based on the premise that: neighboring pixels in space and time having similar intensities should have similar color, following which a quadratic cost function is developed in the form of an optimization problem that can be solved using state-of-the-art quadratic optimization techniques. In principle, our challenge has a similar setting. We hypothesize that the characteristics of the surrounding landscape are a good determinant to estimate the underlying topology of a reservoir. This hypothesis is in synchronization with the fact that the general shape of the valley where dams are constructed have a homogeneous valley shape which can be determined by the characteristics of the surrounding landscape; unless, geological factors such as tectonics and sedimentation affect the homogeneity of the landscape itself. However, for this work, these geological factors are not considered as they are subjective long-term changes which need further research in the field of water management.

Further, in the colorization setting, intensity values of a pixel are considered from a YUV color space, whereas in this work raster based elevation values are considered. In the colorization setting, colored annotations are the input to the algorithm, whereas in this work, the slope of the surrounding landscape are used as input to the algorithm. In the colorization setting, the average values of the neighboring intensity values are inpainted using the bi-laplacian filter, whereas in this work, slopes of neighboring raster values are inpainted using the bi-laplacian matrix. Hence, this colorization setting can be transferred conceptually to model large scale reservoirs. A summary of comparison is presented in Table 4.1.

Following the approach and premise, the algorithmic setup of the algorithm is developed. As previously stated, the input to the algorithm includes elevation values of surrounding rasters (u_b) and it yields an output of elevation values of the underlying landscape (u). We wish to impose the constraint that two neighboring rasters should have similar elevation values if the slope between them is zero.

TABLE 4.2: Summary of comparison between Colorization Approach and the Geometric Model

SETUP	COLORIZATION SETTING	GEOMETRIC MODEL
Data used	Intensity values of pixels from the YUV Color Space	Raster based DEMs (HydroSHEDs)
Input	Grayscale Image with colored annotations	Elevation values of the surrounding landscape with the land-water separation mask
Premise	Neighboring pixels in space and time having similar intensity values should have similar color	Surrounding landscape characteristics are a good determinant to predict the underlying topology of a reservoir
Output	Colored Image	Estimated elevation values of the reservoir (i.e. inside the water mask)

Hence, we minimize the difference between the elevation values of a raster and a weighted average of elevation values at neighboring rasters. Another constraint to be imposed states that the surrounding raster elevation values should remain the same and should not be affected by the algorithm. These two constraints can mathematically be formulated as:

- To minimize the average difference between the elevation values of a particular raster cell and a weighted average of elevation values at neighboring raster cells, an objective function $E(u)$ is introduced which is defined as:

$$E(u) = \frac{1}{2}\sum_{i=1}^{n}\Big(\sum_{j \epsilon N(i)} w_{ij}(u_i - u_j)\Big)^2 \quad \forall u \in C \tag{4.1}$$

- Raster elevation values of the surrounding landscape should remain the same.

$$u = u_b \quad \forall u \notin C \tag{4.2}$$

where, i is the raster cell in consideration, j are the neighboring raster cells, n is the total number of cells in the raster image, u_i is the elevation value at pixel i, u_j is the elevation value at pixel j, C are the raster cells in the water mask for which we have the elevation values and w_{ij} is the weighting function between i and it's neighboring raster cells. For this work, the weighting function is taken as unity, as we hypothesize that all neighboring raster cells equally contribute to the slope of the neighboring raster cell.

Using the graph Laplacian matrix defined previously, Eqn. 4.1 can be reformulated. For the reformulation, raster level representation of the Laplacian matrix is defined as follows:

$$(Lu)_i = \sum_{j \epsilon N(i)} w_{ij}(u_i - u_j) \tag{4.3}$$

Using Eqn. 4.3, the objective function described in Eqn. 4.1 can be restated as:

$$E(u) = \frac{1}{2}||Lu||^2 = \frac{1}{2}u^T L^T Lu \tag{4.4}$$

This equation, thus developed is a multivariate quadratic polynomial with n-dimensions. Hence, the solution to this equation and the final form of objective function is as follows:

$$L^T Lu = 0 \tag{4.5}$$

The objective function includes $L^T L$, thus indicating that the slope of the surrounding two-ring rasters is queried and not just the elevation values. The second constraint described by Eqn. 4.2 is addressed by the introduction of the selector matrix. Selector matrix is a rectangular matrix that selects the constrained pixels or, in other words, it separates the land raster cells from the water raster cells. If $u_b = u_1, u_2, ...u_m$ are the land raster cells or the constrained rasters, then the selector matrix will be as follows:

$$S_{ij} = \begin{cases} 1 & j \in u_b \\ 0 & elsewhere \end{cases}$$

Hence, Eqn. 4.2 can be rewritten as:

$$Su = u_b \tag{4.6}$$

Further, equations 4.5 and 4.6 are formulated in a matrix structure as follows:

$$\begin{bmatrix} L^T L & S^T \\ S & 0 \end{bmatrix} \begin{bmatrix} u \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ u_b \end{bmatrix} \tag{4.7}$$

Dimensions of the matrices are as follows:

- Laplacian matrix (L): $n \times n$
- Selector matrix (S): $m \times n$
- Matrix containing surrounding landscape (u_b): $m \times 1$
- Matrix containing values which are to be predicted (u): $n \times 1$
- Lagrange Multipliers (λ): $m \times 1$

where, the HydroSHEDs and JRC dataset, have total n raster cells, the JRC dataset has m raster cells labeled as water, reulting in $(n - m)$ raster cells to be labeled as land raster cells. Artificial variables in the form of Lagrange multipliers $\lambda_1, \lambda_2, ..., \lambda_m$ are included in order to complete the matrix setup dimensionally, however these variables have no significance and are hence discarded after solving the system.

Overall a quadratic least squares based optimization problem is developed using the objective function and linear constraints. The raster level representation of Laplacian matrix is formulated and used in the development of the objective function. This objective function is the basis of the geometric approach described in this work.

Chapter 5

Implementation

Following the theoretical development and understanding of the geometric model, this chapter provides details on the implementation procedure. An initial implementation of the geometric approach is established including the construction the Laplacian matrix and the details of the quadratic optimization approach. Further, implementation of RMSE, Maximum Storage Capacities and VA curves is discussed, which are generated from the elevation values developed by the algorithm. Finally, control parameters are introduced for further improvement of the model.

5.1 Laplacian Matrix

As seen in Chapter 4, the Laplacian matrix is an integral part of the algorithmic setup. Hence, an initial implementation of the Laplacian matrix was implemented in MATLAB. The psuedocode illustrating the construction of the Laplacian matrix is as follows:

Algorithm 1 Algorithm to compute the Laplacian Matrix

```
procedure LAPLACE(C)                    ▷ C is an input 2D array with n elements
    Initialize L^(nxn) = 0
    for C = 1 to n do
        N ← Set of neighboring elements
        L(N) ← −1
        Diagonal elements of L ← count(N)
    end for
    return L                            ▷ Resulting 2D Laplacian Matrix
end procedure
```

Further, the optimization framework described in Section 4.3 is re-stated in Eqn. 5.1. This optimization framework includes the bi-laplacian matrix ($L^T L$) which takes into account the immediate two-ring neighborhood of the surrounding landscape. The bi-laplacian matrix defined is of high order and is sparse in nature, as most of the elements in this matrix are zeros. Hence, sparse matrices are used as a data structure to define this matrix.

$$\begin{bmatrix} L^T L & S^T \\ S & 0 \end{bmatrix} \begin{bmatrix} u \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ u_b \end{bmatrix} \tag{5.1}$$

Using the Laplacian matrix and the optimization framework (Eqn. 5.1), the initial implementation of the algorithm was developed. The input to the algorithm includes elevation values and land-water separation mask, which yields the missing

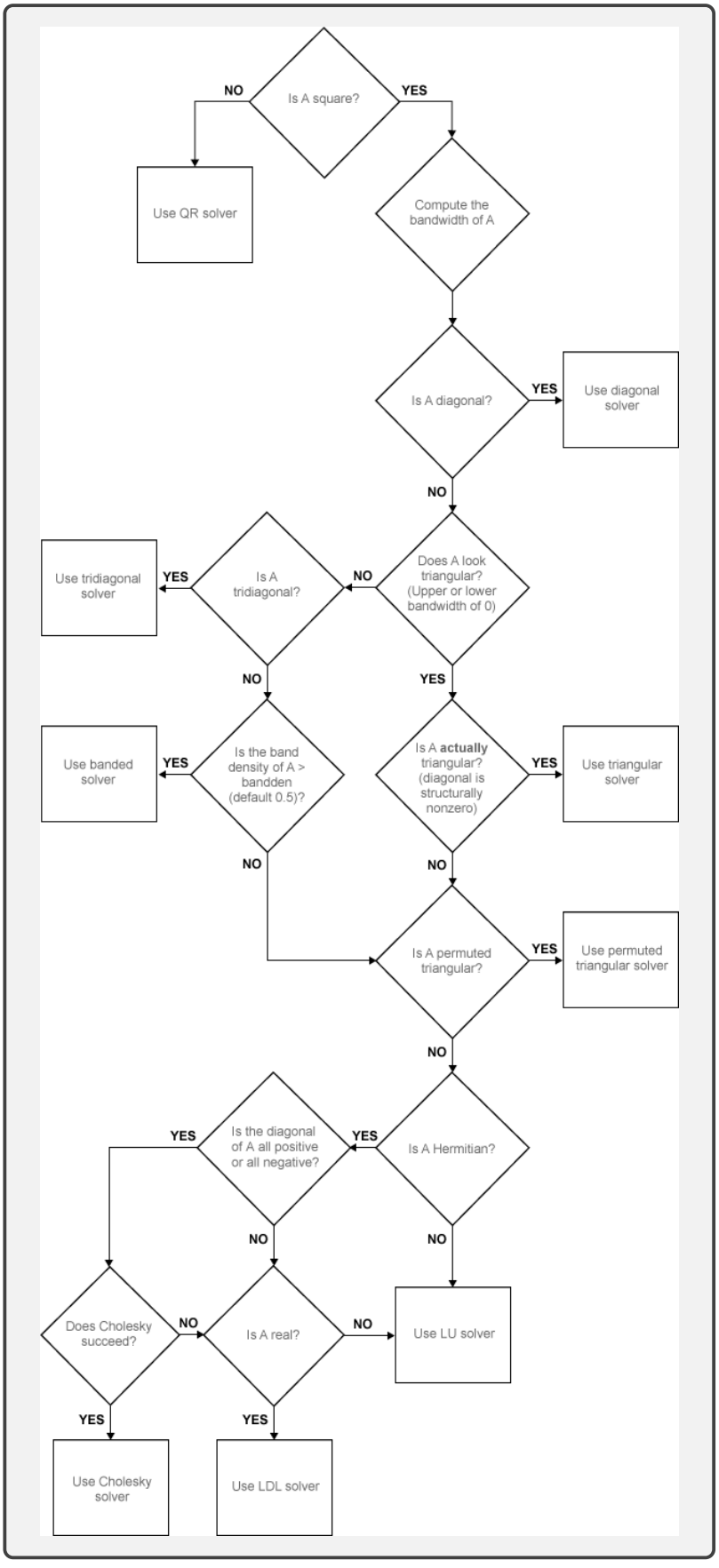


FIGURE 5.1: Optimization flowchart

elevation values of the reservoir landscape as output. Further, the algorithm defined in Eqn 5.1 uses quadratic optimization techniques to solve the system.

5.2 Quadratic Optimization

The proposed algorithm was implemented in MATLAB with the help of the backslash operator[1] which invokes a range of sparse solvers depending on the properties of the matrix generated. Eqn. 5.1 can be further simplified as:

$$Ax = B \tag{5.2}$$

The backslash operator performs small overhead checks on properties of the matrix A illustrated in detail in Figure 5.1. The matrix A, in this work, is formed by a combination of bi-laplacian ($L^T L$) and selector matrices, which has the following properties in context of the flowchart (Figure 5.1):

- A is not a diagonal matrix.
- The band density of A is less than 0.5, hence not good enough for the banded solver.
- A is not a triangular matrix.
- A is not positive definite and is hence not suitable for Cholesky decomposition.

Following the checks on matrix A, the most efficient optimization strategy is the LDL decomposition, and hence MA57 based multifrontal method is used to solve this system. The details of the MA57 approach is presented in Section 3.3 and a further review for sparse linear system solvers is presented by Davis, Rajamanickam, and Sid-Lakhdar, 2016.

Using the laplacian matrix and quadratic optimization approach, elevation values of the underlying landscape are generated. Further, using these elevation values Maximum Storage Estimates and Volume-Area relationships are generated for the computed landscape and the ground truth data.

5.3 Maximum Storage Capacities

For the estimation of maximum storage capacities, r.lake[2] function was used from GRASS GIS. This function fills the lake or cavity of a given DEM to a particular pre-defined level and hence gives the storage capacities of the filled area. The DEM obtained from the geometric model or the ground truth data of reservoirs is filled to the maximum elevation value in the water mask, thus yielding the maximum storage capacities.

[1] Backslash Operator - MATLAB
[2] r.lake - GRASS GIS

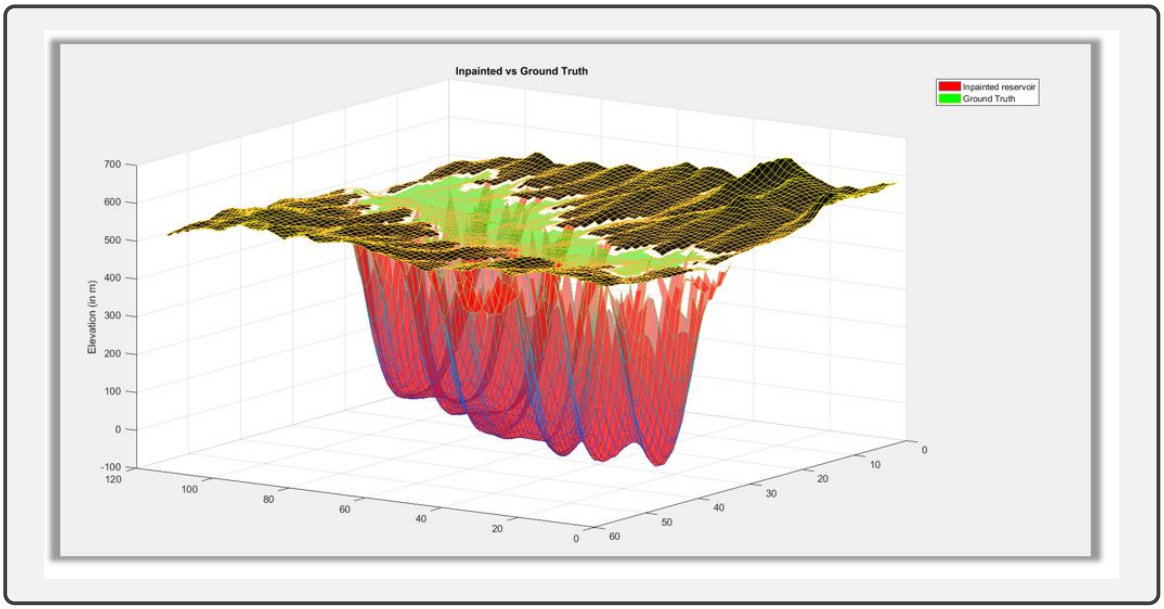


FIGURE 5.2: Effect of Kappa on the geometric model (K = 1). Red curve indicates the modeled reservoir, green curve is the ground truth data and black region is the surrounding landscape.

5.4 Volume Area Curves

For the construction of Volume-Area curves, volume and area estimates of a reservoir are required for varying reservoir water levels. The steps followed to compute VA curves in Python using GRASS GIS bindings are as follows:

- Computation of raster size - In this work, we use HydroSHEDs as our base DEM which is referenced to the WGS84 coordinate reference system. The resolution of the raster cells varies depending on the geo-coordinates of the reservoir. For instance, the resolution at the equator is approx. 3 arc seconds or 90 meters, which decreases as a function of the latitude as we move towards the poles. Hence, to compute near-accurate raster sizes we use the Haversine formula (Section 3.4) which determines the great-circle distance between two points on a sphere given their latitudes and longitudes. The Haversine formula hence computes the size of the rasters in the form of East-West resolution (ewres) and North-South resolution (nsres).

- Area of the reservoir - Area of the reservoir is computed using the count of rasters in the water mask. This count is further multiplied with the north-south resolution and east-west resolution to obtain area estimates. These area estimates vary depending on the shape and size of the water mask.

$$Area\ Estimates\ (in\ m^2) = count * (nsres) * (ewres) \tag{5.3}$$

  where, *count* is the total number of raster cells labeled as water in the JRC mask.

- Volume of the reservoir - Inundated volume of the reservoir is computed using the sum of elevation values in the water mask. This sum is further multiplied with the north-south resolution and east-west resolution to obtain volume estimates. These volume estimates vary depending on the shape/size of the water mask and the elevation values in the water mask itself.

$$Volume\ Estimates\ (in\ m^3) = sum * (nsres) * (ewres) \tag{5.4}$$

  where, *sum* is the total sum of elevation values in the HydroSHEDs data corresponding to which the raster cells are labeled as water in the JRC mask.

- Curves - Following, the maximum and minimum elevation values are computed inside the water mask. The reservoir is virtually filled to the maximum elevation value, resulting in area and volume computations at maximum storage. Further, the level of water is extracted in steps of $1m$ resulting in multiple area and volume computations until the reservoir is emptied. These values are then plotted to give VA curves which help us better understand the shape of the underlying landscape to a reservoir. A visualization of this curve is shown in Section 2.2.4.

5.5 Introduction of Control Parameters

On implementation of the geometric model on real world reservoirs, it was observed that in most cases our geometric model was underestimating depths as compared to the ground truth. To control the generic shape of the inpainted reservoir, in this section parameters such as Kappa and Beta are introduced.

- Kappa is a gravitational pull like operator which pulls the modeled reservoir downwards. This parameter is used to tackle reservoir cases which tend to underestimate depths of the reservoirs by pulling the shape of the reservoir downwards. Kappa is included in the algorithmic setup as shown in Eqn. 5.5.

$$\begin{bmatrix} L^T L + \kappa I & S^T \\ S & 0 \end{bmatrix} \begin{bmatrix} u \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ u_b \end{bmatrix} \tag{5.5}$$

  However, the model is extremely sensitive to the Kappa parameter as small variations in values can result in drastic changes. An illustration of the effect of Kappa can be observed in Figure 5.2. The red curve is the modeled or inpainted reservoir which is pulled drastically when Kappa is set to unity.

- Beta, on the other hand, can be imagined like a rubber sheet or splines which pulls the shape of the reservoir upwards and can be used to tackle models which tend to overestimate depths.

$$\begin{bmatrix} L^T L + \beta L & S^T \\ S & 0 \end{bmatrix} \begin{bmatrix} u \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ u_b \end{bmatrix} \tag{5.6}$$

For the course of this work, Beta parameter is not used because the modeled reservoirs is generally observed to underestimate depths which can be tackled using Kappa. However, the Beta parameter is introduced for future research for modeling reservoirs using the geometric approach.

Chapter 6

Results and Analysis

In this chapter, the results obtained by implementing the basic geometric model on real-world reservoirs are detailed and analyzed to obtain further insights. The results are obtained in terms of RMSE, Maximum Storage capacities and Volume-Area curves, which are compared with the ground truth data and state of the art models. The evaluation setting is depicted in Figure 6.1.

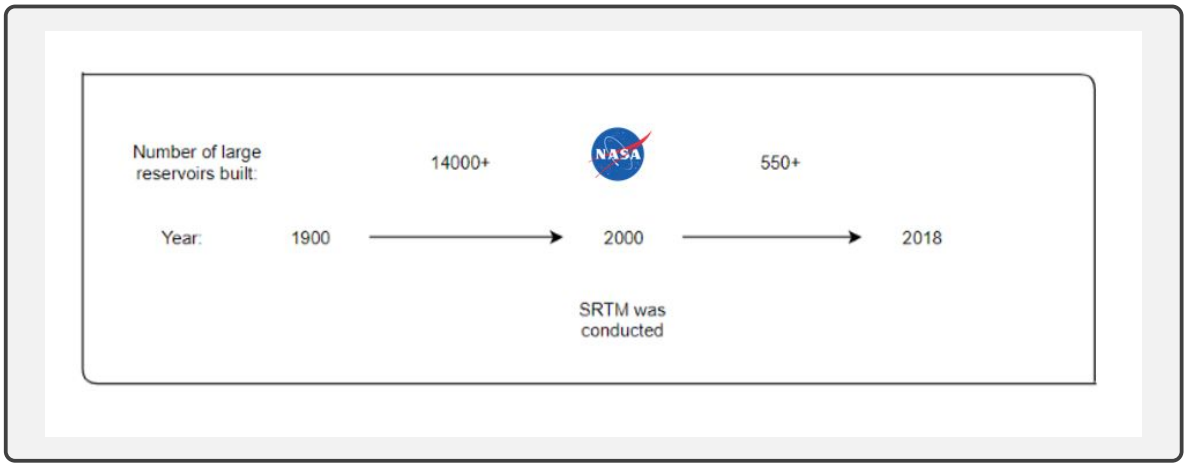


FIGURE 6.1: Evaluation Phase

The idea of the evaluation procedure is that the reservoirs built after 2000 can be characterized using the geometric model and evaluated based on the SRTM data obtained by NASA, following which the 14000+ reservoirs back in time can be characterized using this approach. The evaluation scheme of the geometric model for the reservoirs built after 2000 is divided into two phases:

1. Comparison phase - To compare the geometric model with the GRanD model (Lehner et al., 2011a) by computing their Volume-Area curves, simultaneously comparing both with the ground truth data. Please note that the curves obtained from this phase are referred to as **comparison curves** and the block diagram of data flow is given in Figure 6.2.

2. Evaluation phase - To evaluate the performance of the geometric model at varying storage depths and simultaneously comparing it with ground truth data and the GRanD model. The procedure involves using one water mask for a particular reservoir from the JRC dataset to obtain multiple water masks at different storage depths and then computing the area and volume estimates for these different water masks. The curves obtained by plotting these multiple area and volume estimates are referred to **evaluation curves** and the block diagram of data flow is given in Figure 6.3.

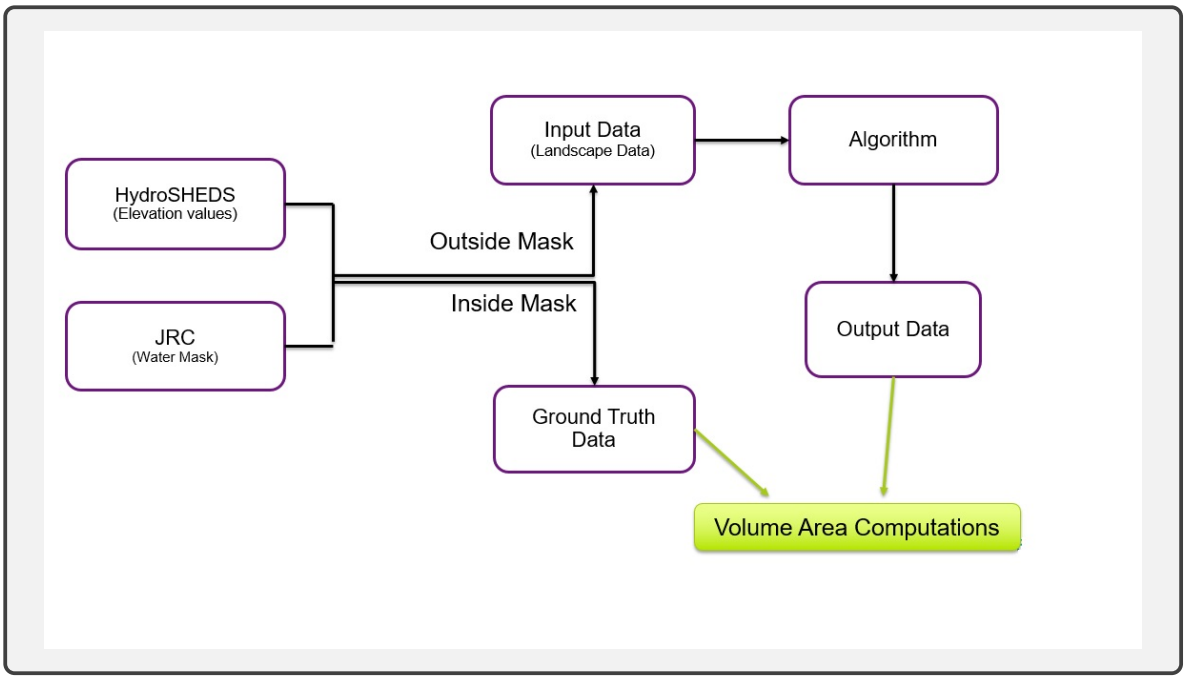


FIGURE 6.2: Block Diagram of data flow - Comparison phase

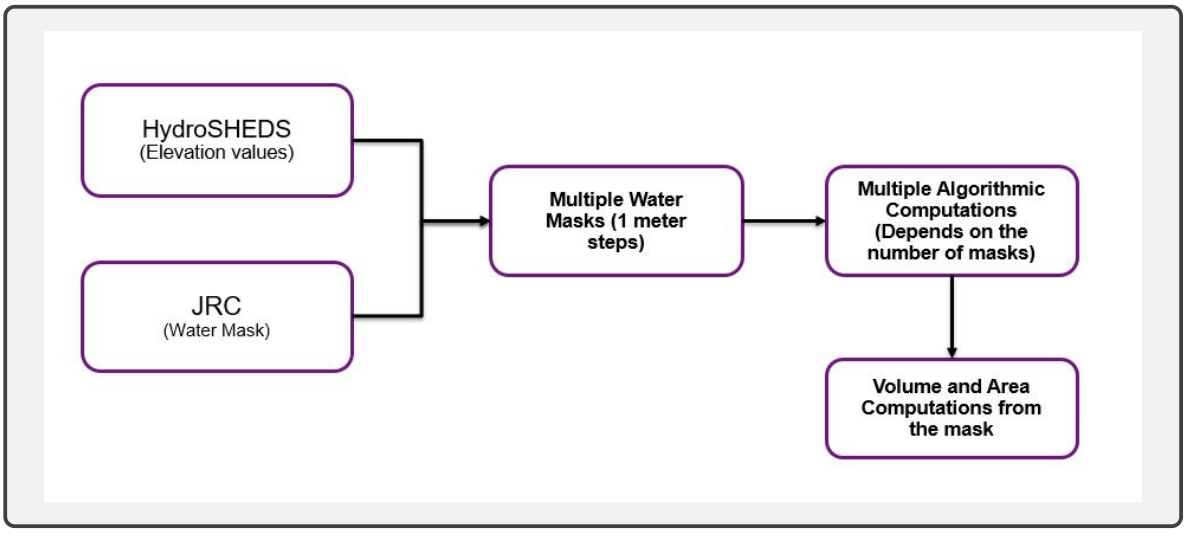


FIGURE 6.3: Block Diagram of data flow - Evaluation phase

For comparison curves, the geometric model is only computed once at the maximum storage capacity, whereas for evaluation curves, the geometric model is computed multiple times ranging from the maximum depth of the reservoir to the minimum depth of the reservoir. Comparison curves are traditional volume-area curves which characterize the reservoirs and are of significant interest in the field of water management, such as to account for anthropogenic impacts in Global Hydrological Models (GHMs). Evaluation curves are defined in this work to test the performance of the geometric model at different storage levels and are hence used to test if the geometric model is sensitive to seasonal changes.

6.1 Basic Geometric Model

This section analyzes the results obtained by the implementation of the proposed basic geometric model. The geometric model is compared to the ground truth data and the GRanD model in terms of RMSE, maximum storage estimates, comparison curves at maximum capacity and evaluation curves. Further, the performance of the basic geometric model on different reservoir cases is discussed.

6.1.1 RMSE Analysis

RMSE analysis is done to examine the performance of the basic geometric model by comparing the computed elevation values with the ground truth elevation values on a per-raster level. In Table 6.1, RMSE at maximum storage level is computed for the 28 reservoirs considered in this work. The performance of the basic geometric model gives an average RMSE of 15.5 m on the considered set of 28 reservoirs. This set of reservoirs has an average depth of 155 m. Hence, on initial analysis depths with a confidence interval of 10 percent are estimated, which is relatively good considering the fact that large reservoirs in considered in this work.

RMSE analysis gives a basic idea about the performance of the proposed geometric model, however, it is difficult to interpret whether the modeled reservoir is underestimating or overestimating depths compared to the ground truth. This is because RMSE penalizes both overestimation and underestimation of modeled depths by squaring the errors on a per-raster level.

Variation in errors can be attributed to various geological factors of the landscape. For instance, Porce 2 reservoir (Dam ID = 3) contains an island underneath the water which is difficult to model and hence gives relatively large RMSE. On the other hand, reservoirs with Dam ID = 11, 35 and 37 have relatively low RMSE because the depth of the reservoirs are low. For further analysis, the Depth:RMSE ratio is given in Table 6.1, which takes into account the depth variation of the reservoirs. Higher the ratio, better the performance of the model.

Modeling absolute and relative errors of reservoirs is beyond the scope of this work. However, to better analyze the performance of the geometric model, maximum storage estimates and volume area curves - Comparison and Evaluation - are computed.

6.1.2 Maximum Storage Estimates

In this section, maximum storage capacities are computed for the 28 reservoirs considered in this work. Maximum storage estimates are initially computed for the modeled reservoirs using the basic geometric approach and compared with the ground truth storage capacities. Further, maximum storage capacities using the GRanD

TABLE 6.1: Root Mean Squared Error and Mean Absolute Error between the Computed elevation values and Ground truth elevation values is listed. Surface area and depth for the reservoirs are also given.

RMSE and MAE for 28 Reservoirs						
ID	Reservoir Name	Surface Area ($in\ m^2$)	MAE ($in\ m$)	RMSE ($in\ m$)	Depth ($in\ m$)	Ratio (Depth:RMSE)
3	Porce 2	7,715,803	14.54	19.14	150	7.84
4	Peribonka	24,854,261	13.17	18.56	262	14.12
6	Caruachi	290,446,217	7.78	11.08	174	15.70
7	Cana Brava	99,273,072	8.89	12.87	138	10.72
8	Manso	325,388,828	7.85	10.9	177	16.24
9	Queimado	17,026,744	6.59	9.03	97	10.74
11	Paraitinga	5,047,029	3.30	4.23	27	6.39
14	Quebra Queixo	5,826,459	13.07	18.01	77	4.28
18	Itapebi	53,329,550	15.85	21.89	251	11.47
20	Itoiz Dam	7,978,982	11.77	17.98	150	8.34
21	Alqueva	169,539,377	10.26	14.85	200	13.47
22	Rules	3,351,000	16. 86	20.32	141	6.94
24	Kozjak	8,189,203	10.61	13.82	154	11.14
25	Capanda	161,247,180	10.94	15.28	224	14.66
26	Mohale	12,261,890	15.64	19.95	204	10.23
28	Maguga	7,930,497	11.82	15.48	104	6.72
29	Karkheh	99,861,622	11.7	17.09	142	8.31
30	Masjed Solayman	44,131,688	33.49	50.27	596	11.86
31	Karoon 3	15,236,540	15.48	20.13	168	8.35
32	Sardar Sarovar	27,124,529	9.68	13.09	119	9.09
34	Bansagar	483,970,456	15.97	27.67	180	6.51
35	Bennithora	16,240,501	3.59	5.64	35	6.21
36	Nina	2,061,055	7.19	10.13	56	5.53
37	Khuga	6,237,828	2.65	3.4	38	11.18
38	Thaphanseik	376,369,244	5.52	7.47	117	15.66
40	Jinpen	2,205,357	11.56	13.92	112	8.05
45	Ham Thuan 1	23,380,857	8.20	11.03	165	14.96
46	Da Mi 1	5,607,964	8.13	10.79	82	7.60

model are computed and compared to the ground truth model. The description of computation of these maximum storage capacities are as follows:

- **Computed Volume (V_1)** - This storage capacity is obtained by taking a sum of the elevation values obtained from the geometric model inside the water mask and multiplying with the raster resolutions.
- **Ground Truth Volume (V_2)** - This storage capacity is obtained by taking a sum of the elevation values obtained from the HydroSHEDs data. Only the raster cells labeled as water are considered and multiplied with the respective raster resolutions.
- **GRanD Volume (V_3)** - This storage capacity is computed using the model developed by Lehner et al., 2011 and is referred to as GRanD volume. This model takes surface area of the reservoir as input and is described as follows:

$$V = 30.684 \times A^{0.9578} \tag{6.1}$$

where, V in $10^6 m^3$ and A in km^2

Converting this equation to SI units, we have:

$$V = 17.1282 \times A^{0.9578} \tag{6.2}$$

where, V in m^3 and A in m^2

Figure 6.4 shows the scatter plot comparing the modeled volumes using the geometric approach (V_1) with the ground truth volumes (V_2). A linear best fit trend-line ($V_2 = 0.9861 * V_1 + 3 * 10^8$) yields a coefficient of determinant (R^2) of 0.9957. On the other hand, Figure 6.5 shows the scatter plot comparing the GRanD model described by Eqn. 6.2 with the ground truth model. A linear best fit trend-line ($V_2 = 16.331 * V_3 + 2 * 10^8$) gives a coefficient of determinant (R^2) of 0.9156.

In both the scatterplots, the data points representing reservoirs underneath the best fit line implies that the storage capacities computed by the model on x-axis overestimates the ground truth storage capacities represented on the y-axis. On the other hand, the reservoirs above the best fit line are underestimated by the model when compared to the ground truth volume.

When comparing the values of storage estimates found in appendix, the GRanD model underestimates the storage capacities in all reservoir cases when compared to the ground truth data. This can be explained by the fact that the regression analysis performed to obtain the GRanD equation (Eqn. 6.2) contains many small reservoirs (both in terms of storage capacities and surface area) in the dataset, and hence the model obtained is skewed towards small reservoirs. On the other hand, the geometric model overestimates the storage capacity for 11 reservoirs and underestimates the storage capacity for 17 reservoirs when compared to the ground truth data.

The storage capacities analysis when compared with RMSE analysis gives better performance and concrete results. This can be explained by the fact that, when computing storage capacities, errors caused due to overestimations and underestimations tend to compensate each other, resulting in good volume estimations. However, RMSE penalizes both underestimated elevation values and overestimated elevation values. To further analyze the performance of the geometric model, VA curves are analyzed in the next section.

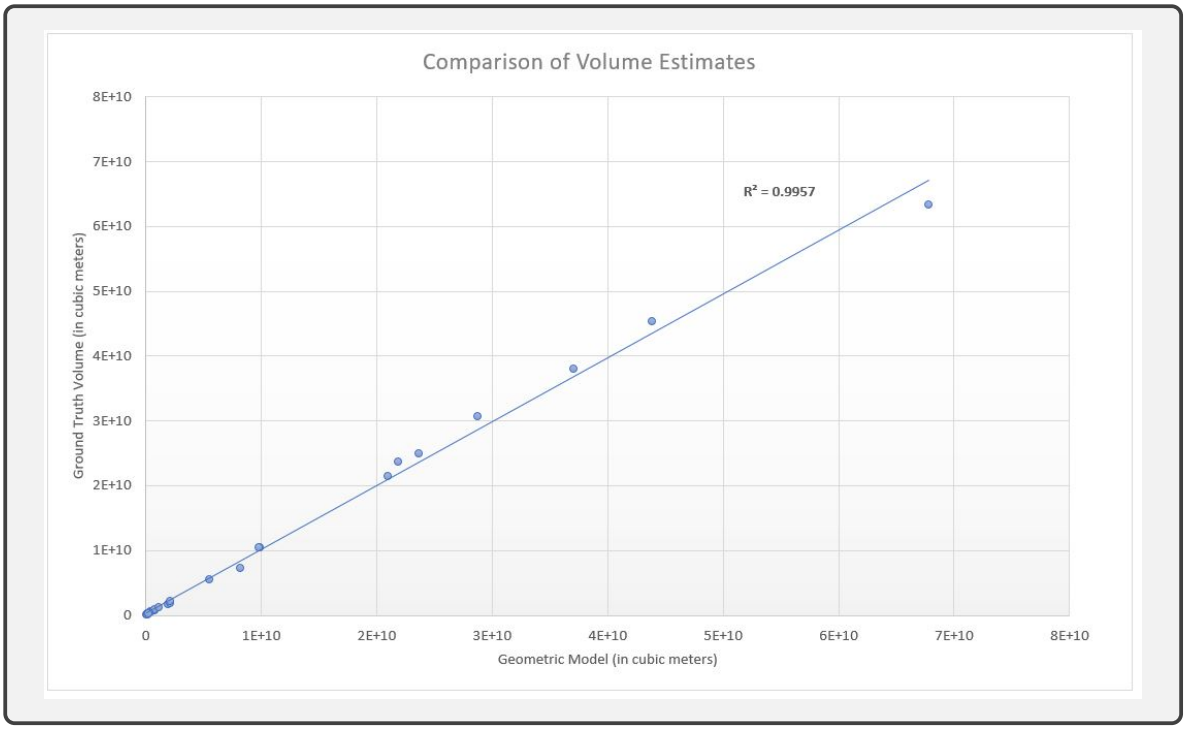


FIGURE 6.4: Ground Truth vs Geometric Model - Maximum Storage estimates

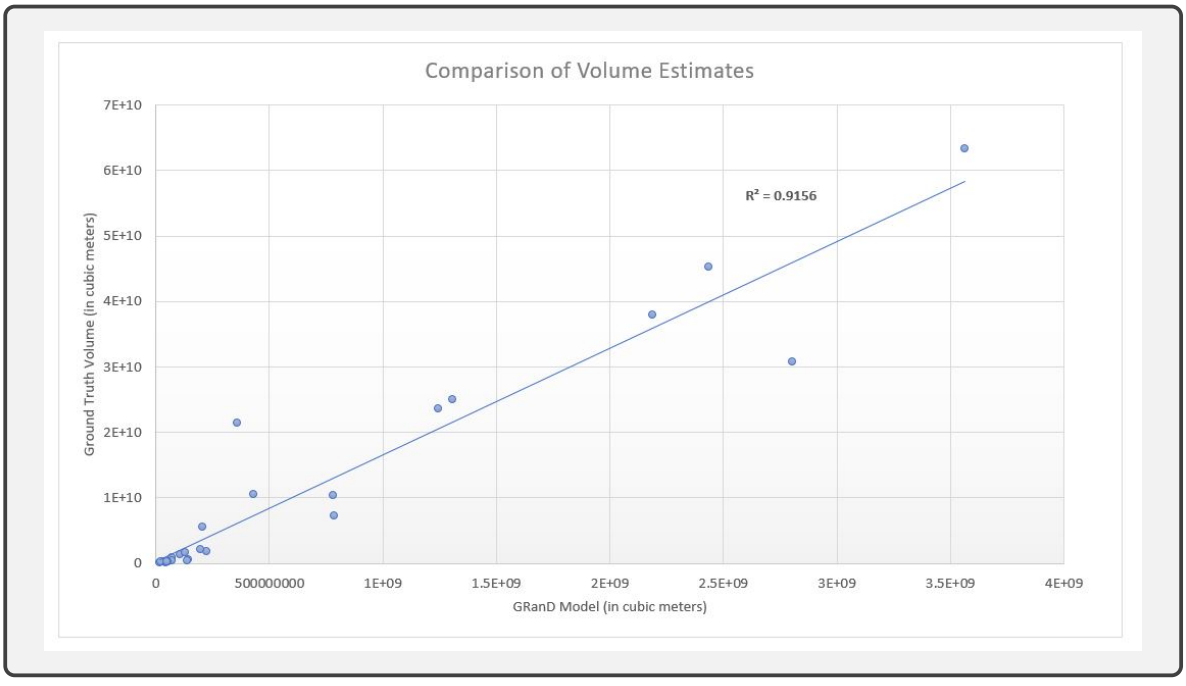
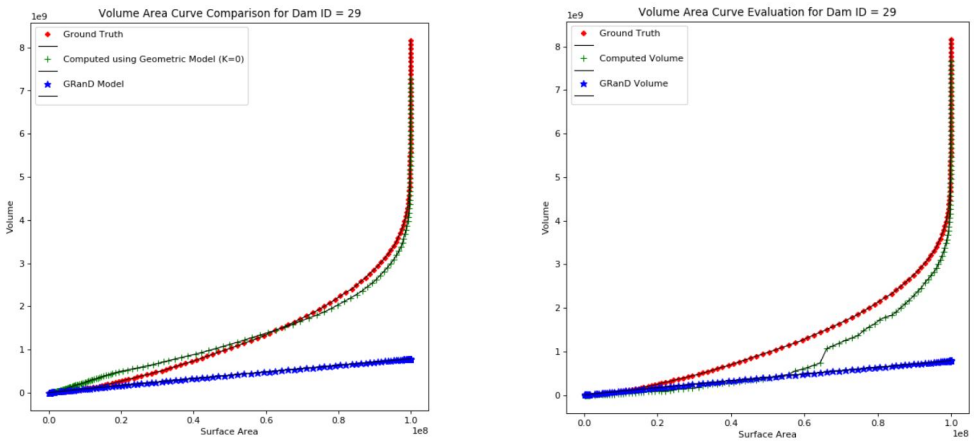


FIGURE 6.5: Ground Truth vs GRanD Model - Maximum Storage estimates

6.1.3 VA Curves Analysis

Following the analysis of maximum storage capacities of the reservoir, VA curves are computed to understand whether the proposed geometric model is accurately able to predict the shape of the reservoir. Comparison curves are built using the JRC water mask at maximum storage and then compared to the ground truth and GRanD models as described in Fig. 6.2.

An example of good comparison curve is shown in Fig. 6.7a. It can be observed that the comparison curve for the basic geometric model (shown in green) is able to predict the ground truth data (shown in red) accurately. In addition, the geometric model proposed performs considerably better than the GRanD model (shown in blue), especially when the reservoir storage capacity increases. This trend can be observed for volume area curves obtained for all reservoir cases presented in the Appendix. Further, it can be seen that the geometric model tends to overestimate storage capacities for the Karkheh reservoir when the surface area is less than 60 percent of maximum surface area and underestimates the storage capacities when the surface area is more than 60 percent of the maximum surface area. However, overall the VA curve is able to estimate the shape of the reservoir accurately.



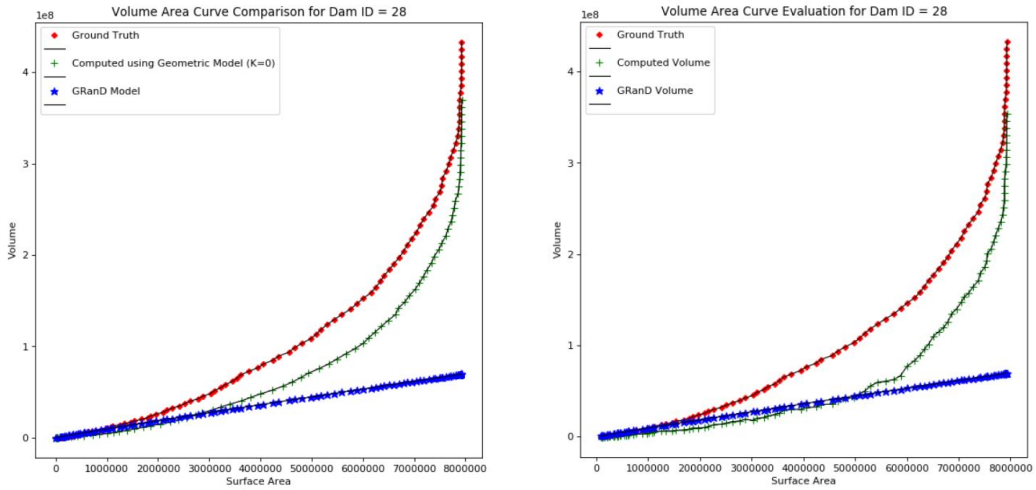
(A) Comparison Curve (B) Evaluation Curve

FIGURE 6.6: VA Curves for Karkheh reservoir, Iran

On the other hand, the comparison curve obtained for the Maguga reservoir in Swaziland shown in Fig. 6.7a is satisactory as compared to the GRanD curve when the surface area is less than one third of the maximum surface area. However, when the effective surface area increases, the geometric model tends to adapt to the shape of the reservoir which the GRanD fails to do, hence yielding accurate storage capacities.

Further, evaluation curves are computed to produce volume and area capacities for reservoirs at different depths using different water masks to understand the effect of seasonal changes on the performance of the model. For instance, these curves can be used to understand the performance of the reservoir when it is at full capacity in the monsoon season versus when it is possibly half the capacity or even lower in the dry summer season. Hence, for this analysis, we use multiple water masks across various depths which are computed as shown in Fig. 6.3.

The evaluation curve for both the Karkheh reservoir, Iran Fig. 6.6b and the Maguga reservoir, Swaziland Fig. 6.7b shows that the geometric model tends to underestimate storage capacities for varying depths. This can be observed for most of the



(A) Comparison Curve (B) Evaluation Curve

FIGURE 6.7: VA Curves for Maguga reservoir, Swaziland

reservoirs, the VA curves for which are given in the Appendix. In order, to compensate for this underestimation, kappa is introduced into the model and the next section analyzes the effect of Kappa parameter on different reservoir cases.

6.2 Global Kappa parameter

Following the observations from the previous section, we noticed that the basic geometric model tends to underestimate the storage capacities of reservoirs in most cases, irrelevant of the storage depths of the reservoir. To compensate for these underestimations, Kappa parameter has been introduced. This section analyzes the impact of the Kappa parameter on the VA curves of a particular reservoir. For the 28 reservoirs considered, Kappa has been optimized at maximum capacity, 75% capacity, 50% capacity and 25% capacity to study the effects of the Kappa parameter and analyze its impact on scale of reservoirs. The Kappa parameter was optimized using the Bayesian Optimization Approach in MATLAB[1], using RMSE as the objective function and optimizing the parameter Kappa in the range 0 to 1.

6.2.1 RMSE Analysis

The geometric model discussed in this setting is very sensitive to the value of Kappa as minor changes in Kappa would affect the RMSE, which in turn would affect the volume estimates. The value of Kappa is between 0 to 1 and even goes to the scale of 10^{-5} for some large reservoirs. Kappa parameter for this work has been optimized using the RMSE, taking HydroSHEDs as the ground truth data. Simultaneously, Kappa has been optimized for different storage capacities - 100%, 75%, 50% and 25%.

The average change of RMSE ($\Delta RMSE$) from the Table 6.2 obtained by optimizing Kappa is 1.36 meters. This is calculated as the difference between RMSE of the geometric model when Kappa is optimal to the RMSE of the basic geometric model (i.e. when Kappa is zero). This change in RMSE seems to have a marginal difference on individual raster depth estimations, however when volume estimates are considered this change would be significant.

Optimal Kappa values that have been computed for the 28 reservoirs at four different capacity values have a mean of 0.02 and a standard deviation of 0.06. Hence these Kappa values do not have a global optimum, or in other words, the Kappa values do not cluster to an optimum value. Hence, Kappa should depend on objective characteristics of the reservoir which is analyzed in the following section.

[1] Bayesian Optimization Algorithm - MATLAB

TABLE 6.2: Optimal Kappa values computed for the 28 reservoirs at different storage capacities.

Kappa Analysis for Reservoirs								
ID	Optimal Kappa	$\Delta RMSE$ (inm)	Optimal Kappa	$\Delta RMSE$ (inm)	Optimal Kappa	$\Delta RMSE$ (inm)	Optimal Kappa	$\Delta RMSE$ (inm)
	100%		75%		50%		25%	
3	0.00086	0.34	0.0005	0.12	0.013	3.988	0.14	3.172
4	0	0	0	0	0	0	0.0008	0.24
6	0.00019	0.73	0.00016	0.71	0.0002	0.73	0.002	2.59
7	0.0008	1.45	0.0008	1.46	0.00085	1.46	0.002	0.97
8	0.000055	1.77	0.000056	1.79	0.00006	1.67	0.00012	2.05
9	0.0017	1.02	0.0017	0.6	0.094	2.08	0.268	4.51
11	0	0	0	0	0	0	0	0
14	0.016	7.96	0.016	8.32	0.0215	4.42	0.052	2.247
18	0.073	3.93	0.074	3.93	0.07	3.83	0.52	9.19
20	0.0015	0.65	0.008	0.2	0	0	0	0
21	0.0044	2.28	0.004	2.27	0.0044	2.29	0.04	4.92
22	0.135	10.8	0.138	11.04	0.13	6.75	0.225	2.9
24	0.005	0.08	0.005	0.08	0.02	0.69	0.112	1.93
25	0.00014	1.86	0.00014	1.86	0.00014	1.88	0.00006	0.09
26	0.0035	2.08	0.0034	2.07	0.0168	5.39	0.084	1.1
28	0.0018	1.49	0.002	1.38	0.014	5.264	0.0328	2.383
29	0	0	0	0	0	0	0.0175	1.01
30	0	-3.11	0	0	0	0	0	0
31	0	0	0	0	0.004	1.16	0	0
32	0.008	0.31	0.006	0.2	0.004	0.06	0	0
34	0	0	0	0	0	0	0	0
35	0.000475	1.82	0.00048	1.82	0.0034	2.645	0.021	3.12
36	0	0	0	0	0	0	0	0
37	0	0	0	0	0	0	0	0
38	0	0	0.00008	1.11	0.00002	1.1	0.0003	2.28
40	0.02	0.18	0.015	0.11	0.022	0.25	0.0126	0.06
45	0.0004	0.08	0.0004	0.08	0.0015	0.381	0	0
46	0.0025	2.48	0.0017	1.26	0.002	0.89	0.005	0.7

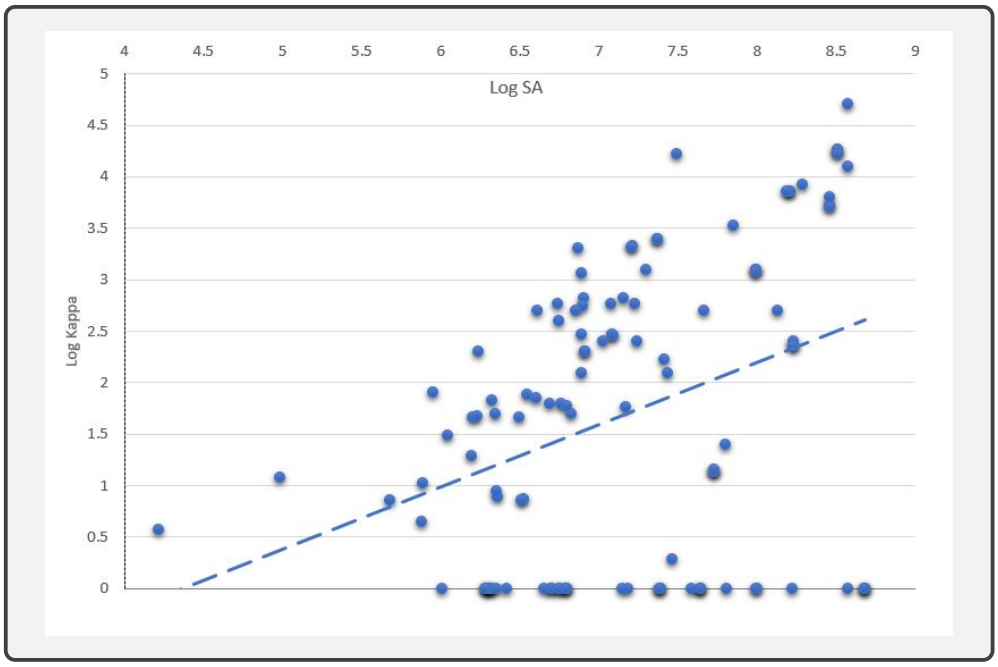


FIGURE 6.8: Log-Log plot between Kappa and Surface Area

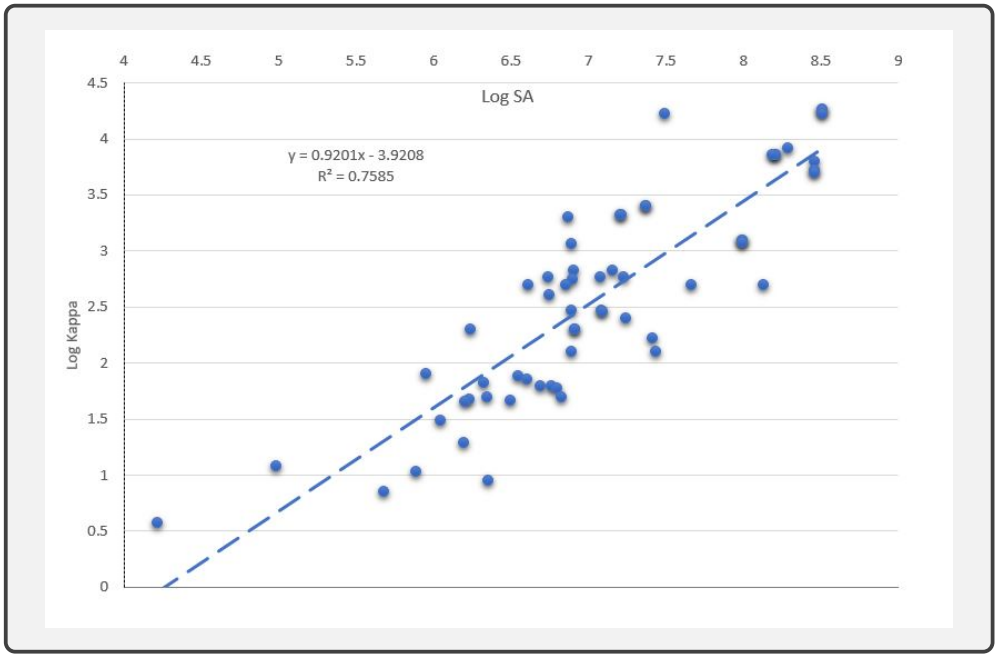


FIGURE 6.9: Log-Log plot between Kappa and Surface Area excluding exceptional cases when Kappa = 0

6.2.2 Relationship between Kappa and size of the reservoirs

In the Table 6.2, Kappa has been optimized using the ground truth data for 28 reservoirs, These Kappa values are different for these 28 reservoirs. However, in a real world setting, to model the large reservoirs bulk of which were built between 1950 to 2000 we need either a global Kappa value or a value dependent on objective characteristics of the reservoir.

Based on the hypothesis that size of the reservoir is a good objective measure, we try to find a relationship between surface area of the reservoir and Kappa. The motivation to use surface area is that it can give an idea about the size of the reservoir and is readily available through remote sensing techniques. In this work, surface area is computed by taking a count of the raster cells labeled as water in the water mask and multiplying it with the respective resolution of the rasters.

From this analysis, reservoirs with optimal Kappa values (Table 6.2) equal to zero were filtered. This resulted in a dataset consisting of 78 reservoirs at varying storage capacities. A description of these filtered reservoirs can be found in Section 6.2.3. Further, a log-log relationship can be found for these 78 reservoirs, shown in Figure 6.9. The log-log relationship gives a coefficient of determination of 0.75 and the relationship developed is as follows:

$$\log \kappa = -0.92 * \log(Surface\ Area) + 3.92 \tag{6.3}$$

where, Surface area is in m^2. Further, this equation can be rewritten as:

$$\kappa * (Surface\ Area)^{0.92} = 2.51 \tag{6.4}$$

This equation signifies that as the size of the reservoir increases the value of Kappa decreases. However, this approach has significant number of outliers and the coefficient of determination is satisfactory but has scope for improvement.

6.2.3 Analysis of Overestimating Reservoirs

The analysis in Section 6.2.2 was done for reservoirs which underestimated depths. However, there were eight reservoirs for which the model overestimated depths, which are analyzed in this section.

- The Peribonka reservoir, Canada is a reservoir that is divided into tributaries. The junction where this river is distributed is relatively wide and does not have surrounding landscape.
- The Paraitinga reservoir, Brazil is relatively a small reservoir and hence can be considered as an outlier. On further analyzing the ground truth data it can be observed that the behavior is weird and possibly the water body is already considered when DEMs were conducted.
- The Karkheh reservoir particularly overestimates in two regions: (1) Where the width of the valley is large and no islands are present. (2) Where the surrounding landscape is very steep due to a peak. This can be seen in Fig. 6.10.
- The Masjed Solayman Reservoir, Iran has a narrow valley with a wide river accounted possibly accounted for in the ground truth data, hence the geometric model overestimates the storage capacities.

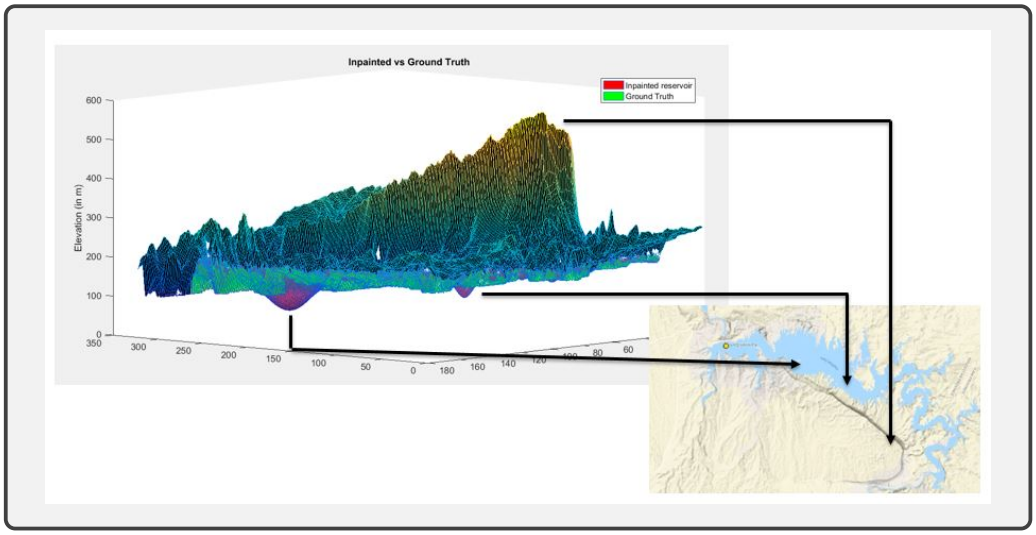


FIGURE 6.10: Karkheh Reservoir, Iran

- Thaphanseik Reservoir, Myanmar is a large reservoir with relatively less islands. The geometric model overestimates the depths in cases when the width of the valley is large. This overestimation is further intensified by the presence of a peak shown by the red arrow in 6.11 and hence resulting in geometric model overestimating the storage depths.
- Further, Nina reservoir, China and Khuga reservoir, India are relatively small reservoirs and hence the geometric model overestimates. On further analysis of the ground truth data, weird behavior can be observed at places where the geometric model overestimates.

Because of these reasons, the above reservoirs were removed from analysis in Section 6.2.2. However, due to a significant number of outliers, this approach is debatable and would be difficult to extend it to model reservoirs before 2000.

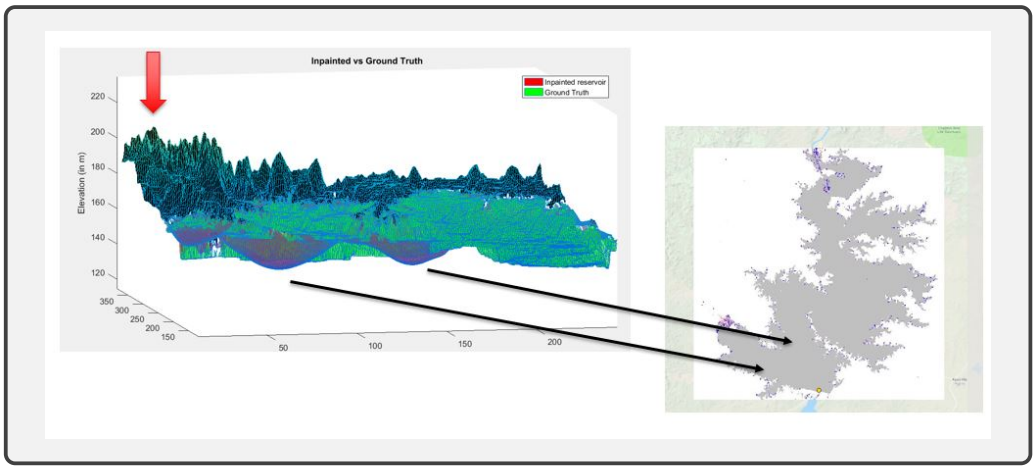


FIGURE 6.11: Thaphanseik Reservoir, Myanmar

6.3 Local Kappa Parameter

In the previous section, a global Kappa was obtained for a particular reservoir by optimizing RMSE. To further improve the modeling of reservoirs, a local Kappa setting is introduced. Further, in the previous section we concluded that Kappa decreases as the size of the reservoir increases. This leads us to the hypothesis that: the distance of the raster cells from the boundary is directly proportional to the Kappa value. To justify this hypothesis, we assume a linear relationship between Kappa and the distance to the closest land raster defined as:

$$\kappa = a + bd \tag{6.5}$$

where, d is the distance of the raster to the nearest land raster (in m) and a,b are constants. Further, we try to include the Eqn. 6.5 in the basic algorithmic setup of the geometric approach restated in Eqn. 6.6.

$$\begin{bmatrix} L^T L + \kappa I & S^T \\ S & 0 \end{bmatrix} \begin{bmatrix} u \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ u_b \end{bmatrix} \tag{6.6}$$

For this approach, the Kappa values are fed into the identity matrix using the linear relationship hypothesized between Kappa and distance to the nearest land raster. The task now is to find optimal values of a and b which minimizes the RMSE.

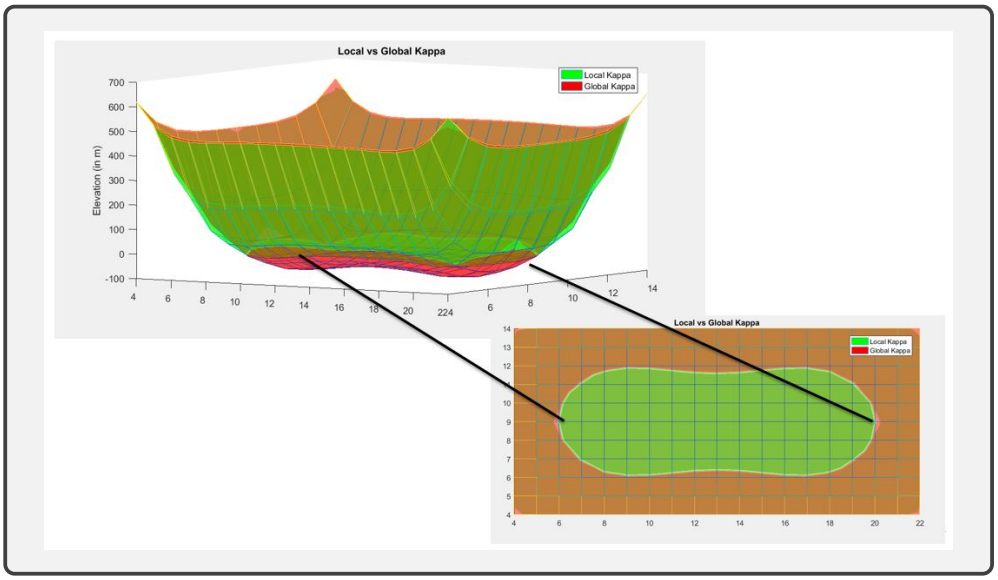


FIGURE 6.12: Top view and Side view of a reservoir illustrating the difference between Global Kappa (in red) and Local Kappa (in green)

To compute optimal values for a and b, first the model described by Eqn. 6.5 and Eqn. 6.6 was implemented. Further, RMSE was used as an objective function to find optimal values of a and b in a 2D space. This challenge was addressed using the Bayesian Optimization setting where RMSE is the objective function to be minimized and a, b are the hyperparameters.

However, on implementation of the above procedure and testing it on the 28 reservoirs, it was observed that the least RMSE approached the RMSE observed in Table 6.2. Also, the values of a were equal to the Optimal Kappa values in Table 6.2. Further the value of b was of the order 10^{-8}, indicating that localized Kappa model converged to the global Kappa model.

6.4 Discussion of different reservoir cases

In this section, the reservoirs of different shapes and sizes are discussed. Further, effects of implementation of the basic geometric model on these cases are detailed. Different reservoirs cases include - Reservoir with islands, dendritic shaped reservoirs, large reservoirs and narrow reservoirs. It is to be noted that small reservoirs and reservoirs with comparatively less surface area such as reservoirs built for hydroelectric purposes have been not considered for this analysis.

6.4.1 Reservoirs with islands

Islands between reservoirs are possible sources of errors while predicting the bathymetric landscape of the reservoir. Previous models, predicting storage capacities do not account for such islands while analysis. In the geometric model implemented, landscape changes due to large islands that can be segregated using remote sensing techniques can be used to further improve the model. The premise of the geometric model implemented was to predict the bathymetry using surrounding landscape values and with addition of islands, the amount of neighboring rasters increase, thus yielding accurate landscape predictions.

This assertion is tested on the Ham Thuan 1 reservoir in Vietnam which contains a big island in the center of the reservoir with some small islands as well. The enhanced 3D view of the big island can be seen in Fig. 6.16. The accuracy of VA curves for this reservoir can be seen in Figure 6.14. Further, it can be observed that both comparison and evaluation curves are in synchronization with the ground truth data. Also, the performance of the geometric model is very good as compared to GRanD, which does not take the islands into account. Finally, the evaluation curve computed using the geometric model slightly underestimates the evaluation curves computed by ground truth data when the surface area crosses two fifth of the maximum capacity.

Further, two reservoirs with islands include Cana Brava Reservoir, Brazil and Capanda Reservoir, Angola. These reservoirs are four times and seven times the size of the Ham Thuan Reservoir respectively in terms of the surface area. The geometric model implemented performs well for both these cases in terms of elevation and comparison curves. The VA curves for these reservoirs can be found in the Appendix.

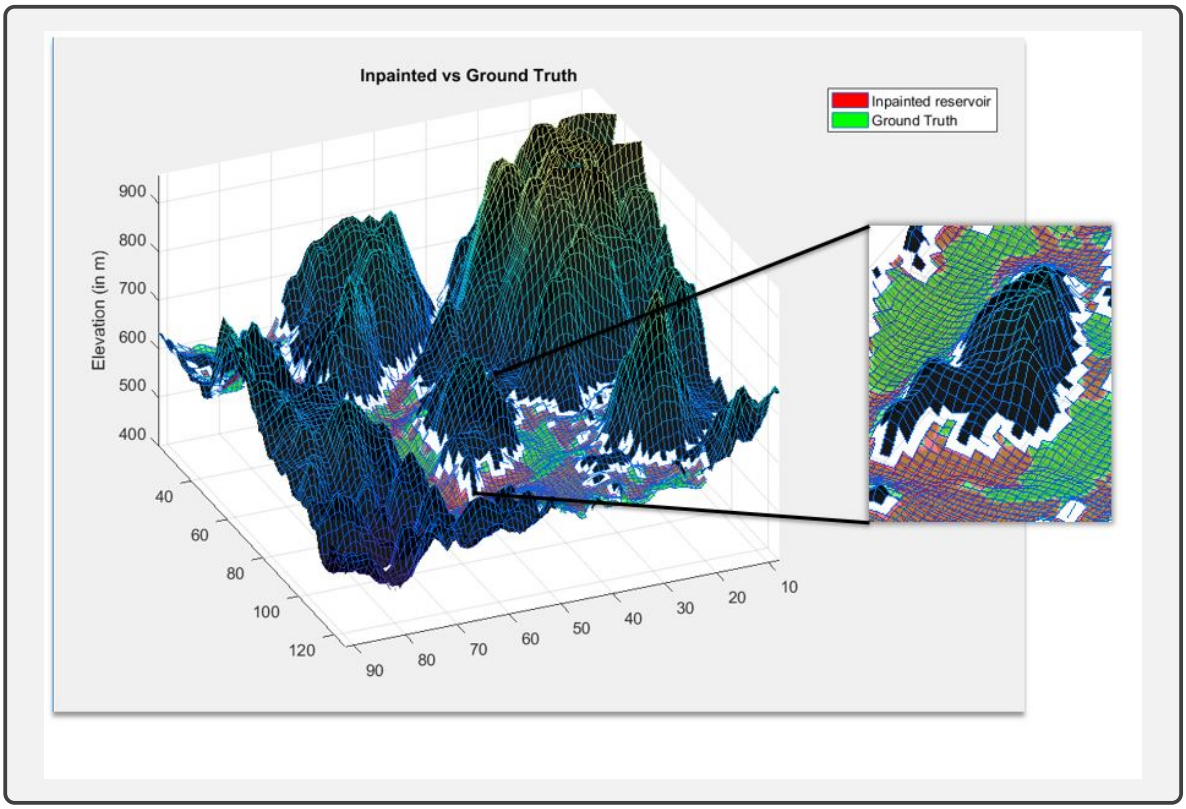
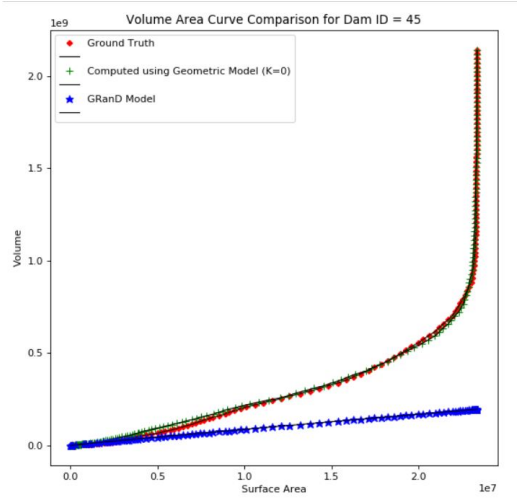
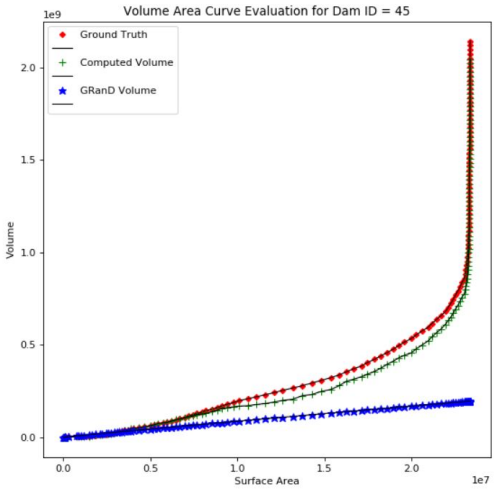


FIGURE 6.13: 3D view of Hum Thuan, Vietnam: where dark raster cells are classified as land, red raster cells is where the geometric model underestimates and green raster cells indicate that the geometric model is overestimating.



(A) Comparison Curve



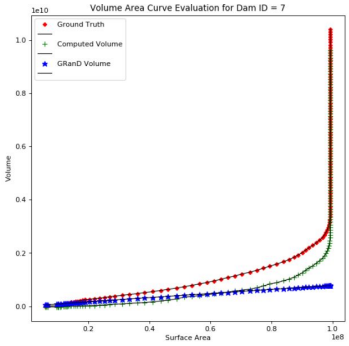
(B) Evaluation Curve

FIGURE 6.14: Hum Thuan 1, Vietnam

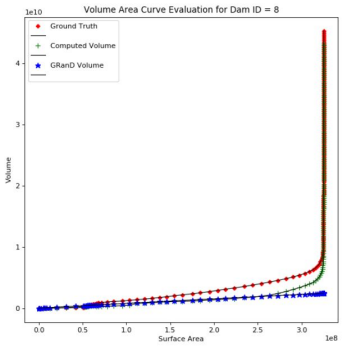
6.4.2 Dendritic shaped reservoirs

In real world reservoir cases, relatively large number of reservoirs have dendritic-shaped catchments when seen from a top view. This is because when reservoirs are built for irrigational purposes, their shape should be spread out for the water to reach a large amount of people. These dendrites, when small in number do not affect the storage capacities of the model. However, when the shape of the valley is erratic or the Horton Strahler number (indicating the outlets of the river body) are relatively large, volume of water stored in these dendritic shaped valleys is considerable and has to be accounted for. Previous work to estimate storage capacities do not specifically account for dendritic shaped valleys.

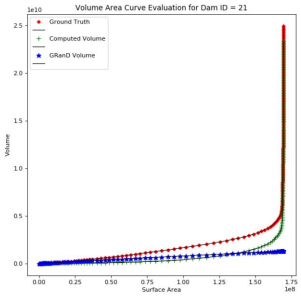
In principle, when we take the shape of dendrites into consideration, the neighbors are relatively close as compared to when we have monotonous shaped wider valleys. Hence, when geometric model is implemented, the performance is expected to be better for these dendritic shaped valleys as we have access to more neighbors which will help us better predict the underlying topology.



(A) Cana Brava Reservoir, Brazil



(B) Manso Reservoir, Brazil



(C) Alqueva Reservoir, Portugal

FIGURE 6.15: Comparison Curves for Dendritic shaped reservoirs.

Some reservoirs that have such dendritic shaped valleys include Cana Brava Reservoir and Manso Reservoir in Brazil and Alqueva Reservoir in Portugal. The VA curves for these reservoirs are shown in Fig. 6.15. The general trend observed in dendritic shaped reservoirs is stagnation of the surface area which can be observed in the form of a straight line at maximum surface area. This behavior can be observed for both ground truth data and the geometrically modeled data. The dendrites are responsible for this behavior because when we increase the volume of water in such a reservoir, surface area remains constant because the change in surface area is insignificant because of the dendrites.

Overall, the geometric model can predict the shape of these dendritic shaped reservoirs better which the GRanD model fails to predict.

6.4.3 Large Reservoirs

In this section, the geometric model implemented for large reservoirs are analyzed. Modeling large reservoirs is difficult especially when the shape of the valley is broad because the model does not have enough input characteristics from the surrounding landscape.

Some reservoirs included in this analysis include:

- Caruachi Reservoir, Venezuela
- Manso Reservoir, Brazil
- Alqueva Resevoir, Portugal
- Capanda Reservoir, Angola
- Bansagar Reservoir, India
- Thaphanseik Reservoir, Vietnam

The basic shape and the mask of these reservoirs can be seen in the appendix. The performance of these reservoirs in terms of RMSE for maximum storage capacity is 14.5 meters which is less than the overall average RMSE of 15.5 meters. Hence, the geometric model performs better for large reservoirs as compared to the relatively smaller ones. At the same time, due to skewness in the GRanD model it does not work well for large reservoir cases. The VA curves computed via the geometric model for these reservoirs thus better predict the shape of the valley. Overall, the performance of these reservoirs

However, one particular case where it does not work well is the Bansagar Reservoir from India. This reservoir is the biggest reservoir taken in consideration in this analysis. It can be noticed from the figure that the geometric model is inaccurate when the raster pixels are far away from the surrounding landscape. These overestimations can be clearly seen in the VA curves as well. However, this reservoir is an exceptional case because the surrounding landscape is fairly irregular and a highway goes from amidst the reservoir, the landscape for which is not taken into account.

6.4.4 Narrow Valley Reservoirs

Due to low construction cost, a good chunk of reservoirs built today are narrow valleyed in nature. In this section, we analyze the performance of the geometric model for reservoirs for large narrow valley reservoirs which ave significant valley width for extrapolations.

In some cases, such as the Kozjak reservoir in Macedonia and Karoon 3 reservoir in Iran, the geometric model is able to better predict the shape of the reservoir accurately. In both these cases, the performance of the GRanD model deteriorates as the size of the reservoir crosses roughly 25% of the maximum storage capacity. The good performance of the geometric model is because the valley for both these reservoirs is homogeneous and narrow. Hence, predicted elevation values are nearly accurate to that of ground truth data. The shape of the two reservoirs can be seen in the appendix.

Further, for Peribonka Reservoir, Canada the geometric model tends to overestimate storage depths as shown in Figure 6.16. However, this is because the geometric model overestimates depth when the catchment is divided into two areas. There is lack of surrounding landscape elevation values at the junction where this happens. Simultaneously, the surrounding landscape of this reservoir is very steep as the reservoir is narrow and deep. These two factors when combined leads to overestimation by the geometric model. This overestimation however can be controlled by including beta as a control parameter in the model.

Overall, the performance of the geometric model on reservoirs with narrow valley is pretty accurate both in terms of RMSE and VA curves. Further examples can be seen in appendix as the Itapebi reservoir in Brazil.

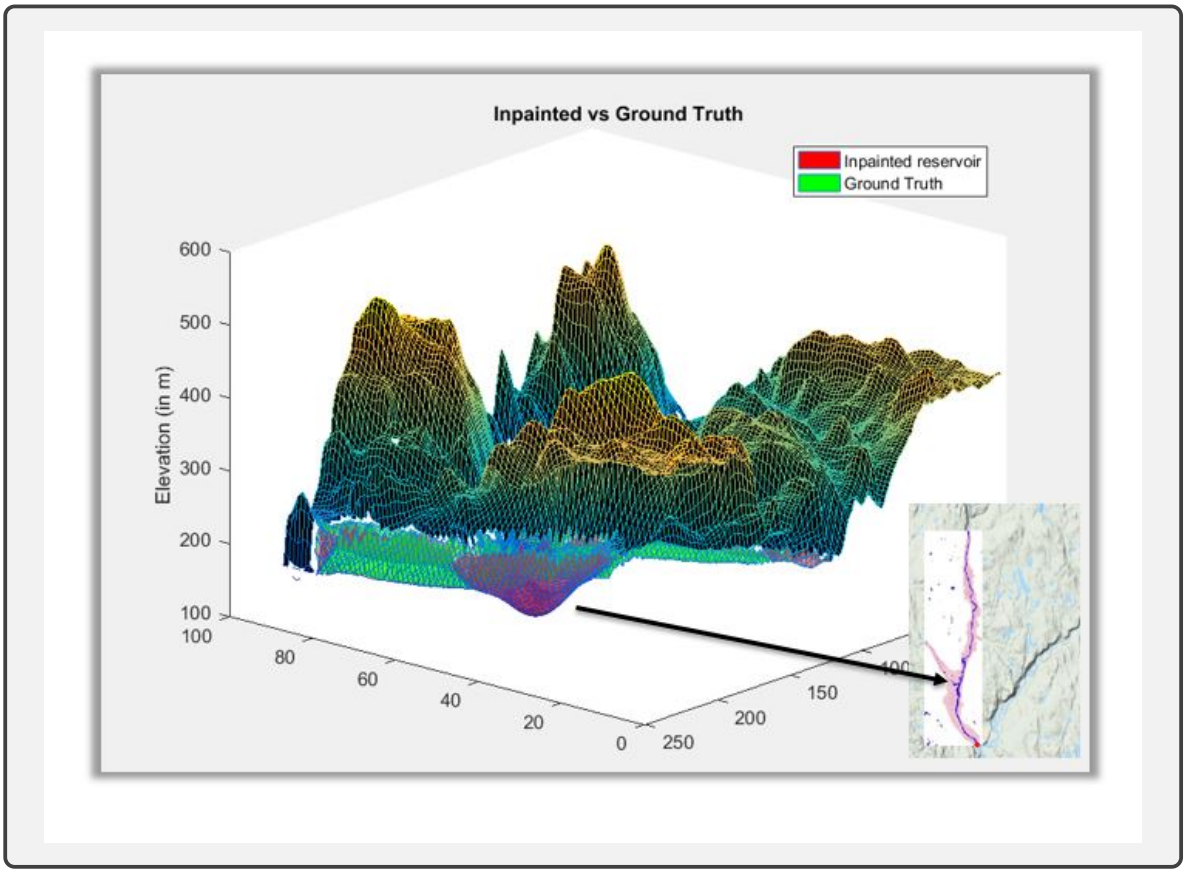
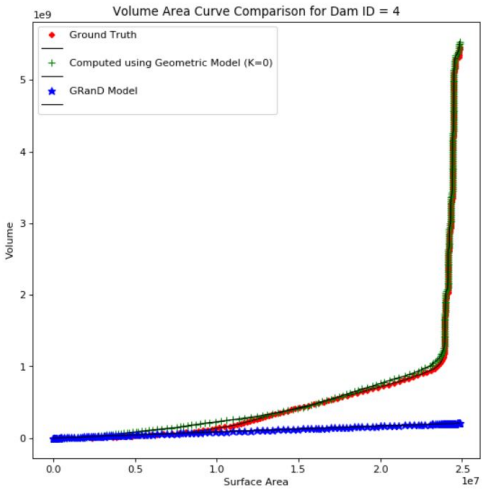
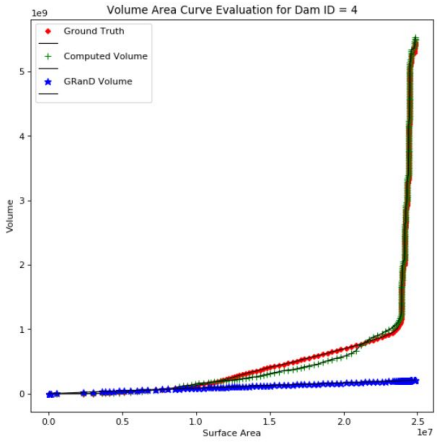


FIGURE 6.16: 3D view of Peribonka Reservoir, Canada. Arrow pointing out the area where the model overestimates.



(A) Evaluation Curve



(B) Comparison Curve

FIGURE 6.17: VA curves for Peribonaka Reservoir, Canada

Chapter 7

Conclusions and Future Work

This concluding chapter explains the scientific and technical implications of the research findings. In addition, answers to research questions is formulated in this chapter and recommendations are stated for future work.

7.1 Conclusion and Insights

In summary, a computer graphics based optimization framework is developed using the premise that the characteristics of the surrounding landscape to a reservoir is a good determinant to predict the topology within the reservoir. This framework presented is a geometric inpainting approach similar to a basic computer graphics challenge such as colorization using optimization (Levin, Lischinski, and Weiss, 2004).

The model developed is evaluated using the SRTM data {HydroSHEDs (Lehner, Verdin, and Jarvis, 2006)} for reservoirs that are built after 2000. The evaluation procedure is performed by comparing the maximum storage estimates and volume area curves with the ground truth data obtained from SRTM. Further, the geometric model is also compared with previous work in the form of GRanD model (Lehner et al., 2011a) which is a widely used model to estimate storage capacities.

In addition, the model is not affected by seasonal changes, as it works pretty well for reservoirs with varying depths tested by the evaluation curves for different reservoir cases. The model performs significantly better when compared to ground truth data for different shapes and sizes such as Reservoirs with islands, Narrow valley reservoirs, Large reservoirs and dendritic shaped reservoirs. The better performance in terms of storage capacities is further enhanced when VA curves are computed indicating that the geometric model is able to predict the shape of the underlying landscape accurately.

In general, it was observed that the model was underestimating depths as compared to ground truth data. This led to the introduction of the control parameter -Kappa- in the geometric model. Further, Kappa was optimized for the 28 reservoirs taken into consideration based on the RMSE using the Bayesian optimization approach. These Kappa values were observed to vary with the size of the reservoir on a logarithmic scale, such that the Kappa value is inversely proportional to the size of the reservoir. Hence, a log-log relationship was developed between Kappa and surface area. However, number of outliers in this equation was considerably high which provided the motivation for a modified Kappa model.

The modified Kappa model involved localizing Kappa depending on distance of the raster distance to the boundary. Hence a linear model was considered between Kappa and distance to the boundary (d): $\kappa = a + bd$. However, for finding an optimal value for a and b based on the 28 reservoirs considered in this analysis, it was observed that no improvements could be found in the model in terms of RMSE. In

fact the optimal value of b was of the scale 10^{-8} and a approached the value obtained by optimizing Kappa globally. Hence, this modified Kappa approach did not improve the performance of the geometric model.

In conclusion, the geometric model significantly improved the estimation of storage estimates and VA curves. This cross-disciplinary approach may hence be deemed as a valuable contribution to the study of reservoirs which can prove to be a platform for further research in the field of water management.

7.2 Recommendations

The geometric model presented uses surrounding landscape characteristics to predict the topology of the reservoir. This work is a platform for future research in the field of water management. While the results are promising, here are some recommendations for further research directions:

The model developed is studied for lotic ecosystems such as reservoirs in this work. However, this model can be extended in the study of lentic ecosystems such as lakes and ponds. As limited literature exists separating the lotic and lentic water bodies, this approach can provide concrete models to model water bodies in these ecosystems.

The optimization of Kappa parameter is done using reservoir characteristics like the size of the reservoirs (surface area) and distance of the raster cells to boundaries to minimize user interaction. However, further characteristics such as inscribed circle or circumscribed circle can be tested that tend to better estimate the shape and size of a reservoir.

During the development of the model itself, there is scope for further experimentation. For instance, the weights of the neighbors are considered to be equal throughout this work. Objective reservoir modeling measures can be used to describe these weights which may help in improving the performance of the models.

Further, this work involves raster based analysis of reservoirs using surrounding landscape data. Further, vector based analysis can be used to model these reservoirs on a global scale. Vector DEMs are readily available and hence graphics techniques based on diffusion curves can be implemented to understand reservoir characteristics (Bezerra et al., 2010) (Orzan et al., 2008).

The geometric model was developed to characterize reservoirs at a particular time. This analysis can be further extended by incorporation of temporal data. In principle, this would be similar to extending the computer graphics study from images to videos.

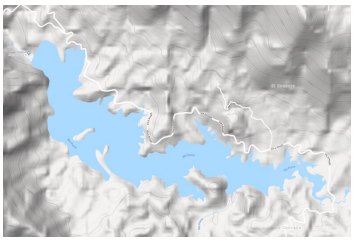
The resolution of DEMs used in the work is 3 arc seconds. In the coming years, satellite missions are expected to obtain better DEMs in terms of spatial resolution which will further improve the performance of this model. Given, DEMs with better resolutions the model can be extended to model small-scale reservoirs and narrow valleyed reservoirs as well.

Appendix A

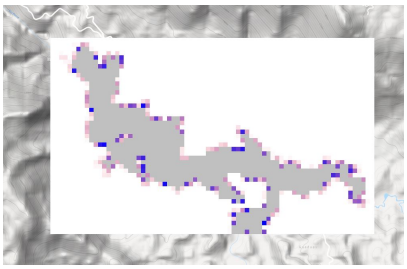
Appendix A: Reservoir Characteristics

A.1 Shape of Reservoirs and JRC Water Mask

This appendix section provides the shape of the reservoirs and the shape of the JRC water mask taken into consideration.



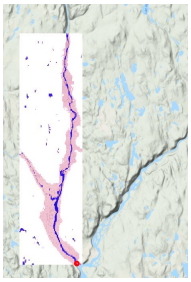
(A) Shape of the actual reservoir



(B) Shape of water mask

FIGURE A.1: Porce 2, Columbia

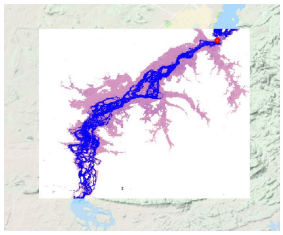
(A) Shape of the actual reservoir



(B) Shape of water mask

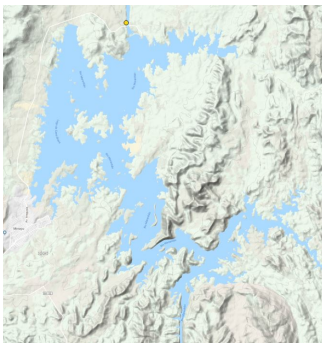
FIGURE A.2: Peribonka, Canada

(A) Shape of the actual reservoir

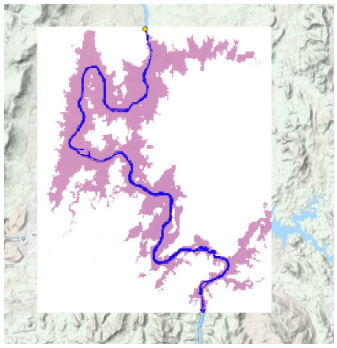


(B) Shape of water mask

FIGURE A.3: Caruachi, Venezuela

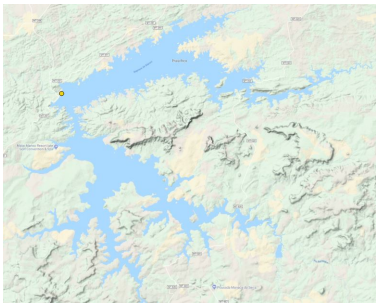


(A) Shape of the actual reservoir

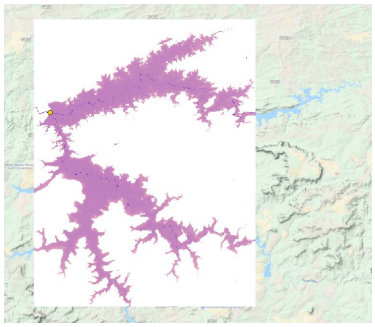


(B) Shape of water mask

FIGURE A.4: Cana Brava, Brazil



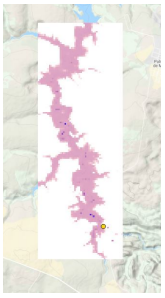
(A) Shape of the actual reservoir



(B) Shape of water mask

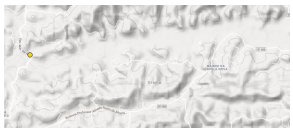
FIGURE A.5: Manso, Brazil

(A) Shape of the actual reservoir

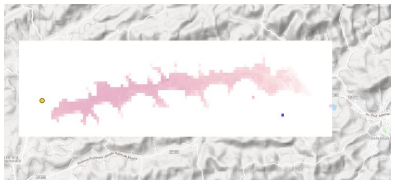


(B) Shape of water mask

FIGURE A.6: Quiemado, Brazil



(A) Shape of the actual reservoir



(B) Shape of water mask

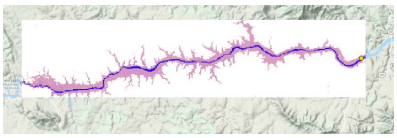
FIGURE A.7: Paraitinga, Brazil

(A) Shape of the actual reservoir

(B) Shape of water mask

FIGURE A.8: Quebra Qeixo, Brazil

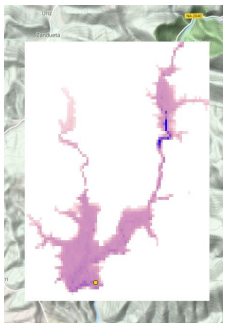
(A) Shape of the actual reservoir



(B) Shape of water mask

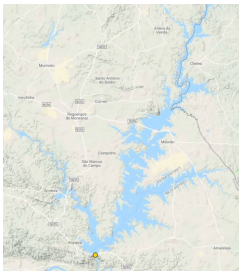
FIGURE A.9: Itapebi, Brazil

(A) Shape of the actual reservoir

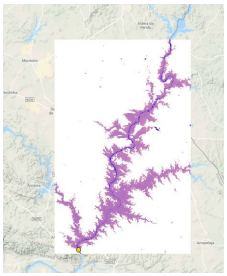


(B) Shape of water mask

FIGURE A.10: Itioz, Spain

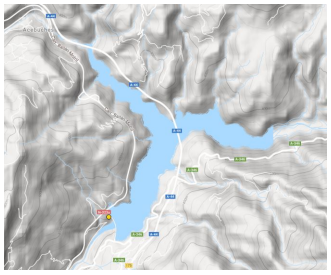


(A) Shape of the actual reservoir

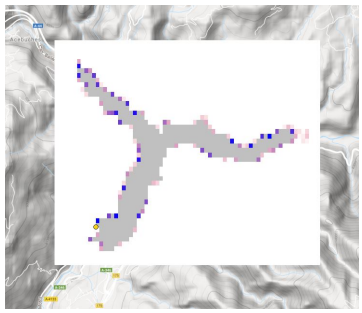


(B) Shape of water mask

FIGURE A.11: Alqueva, Portugal



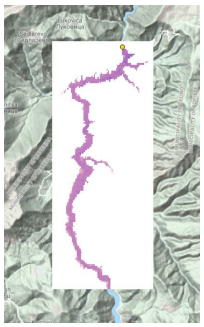
(A) Shape of the actual reservoir



(B) Shape of water mask

FIGURE A.12: Rules, Spain

(A) Shape of the actual reservoir



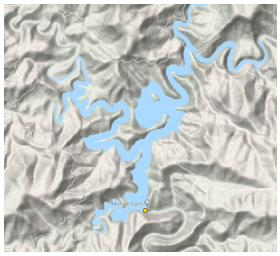
(B) Shape of water mask

FIGURE A.13: Kozjak, Macedonia

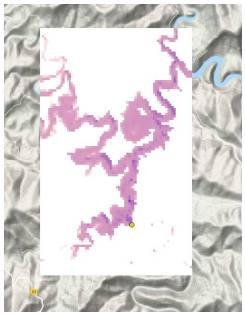
(A) Shape of the actual reservoir

(B) Shape of water mask

FIGURE A.14: Capanda, Angola



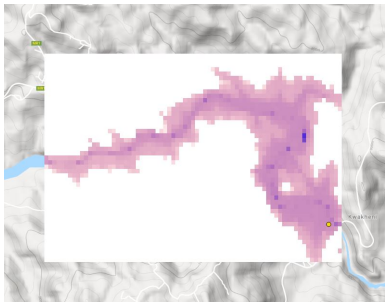
(A) Shape of the actual reservoir



(B) Shape of water mask

FIGURE A.15: Mohale, Lesotho

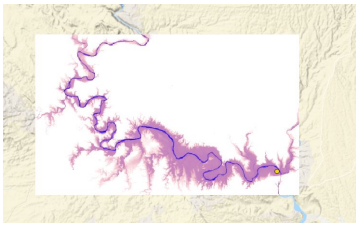
(A) Shape of the actual reservoir



(B) Shape of water mask

FIGURE A.16: Maguga, Swaziland

(A) Shape of the actual reservoir



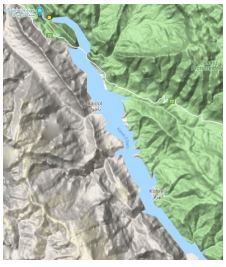
(B) Shape of water mask

FIGURE A.17: Karkheh, Iran

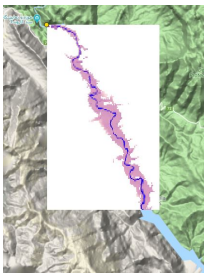
(A) Shape of the actual reservoir

(B) Shape of water mask

FIGURE A.18: Masjed Solayman, Iran

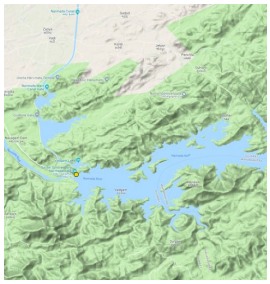


(A) Shape of the actual reservoir

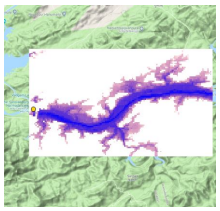


(B) Shape of water mask

FIGURE A.19: Karoon 3, Iran



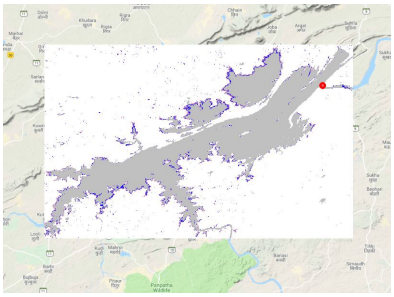
(A) Shape of the actual reservoir



(B) Shape of water mask

FIGURE A.20: Sardar Sarovar, India

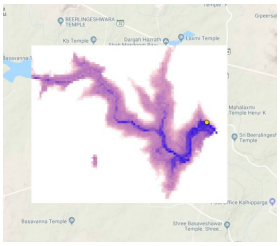
(A) Shape of the actual reservoir



(B) Shape of water mask

FIGURE A.21: Bansagar Dam, India

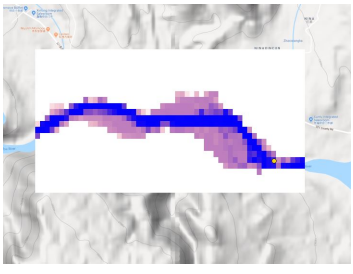
(A) Shape of the actual reservoir



(B) Shape of water mask

FIGURE A.22: Bennithora, India

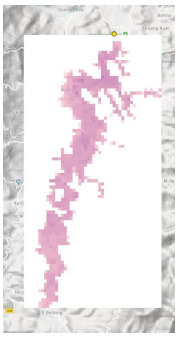
(A) Shape of the actual reservoir



(B) Shape of water mask

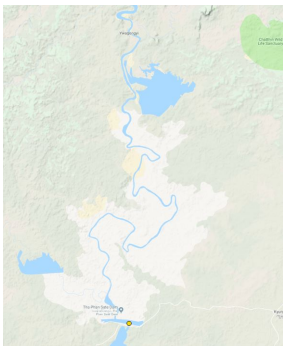
FIGURE A.23: Nina, China

(A) Shape of the actual reservoir

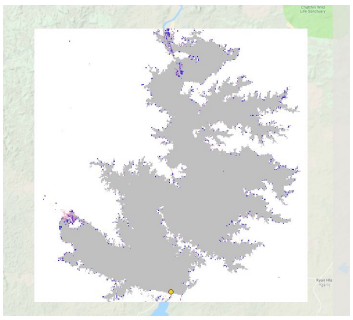


(B) Shape of water mask

FIGURE A.24: Khuga, India

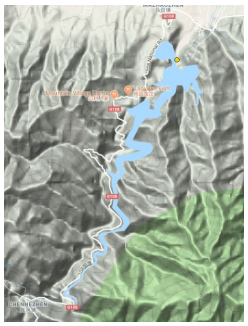


(A) Shape of the actual reservoir

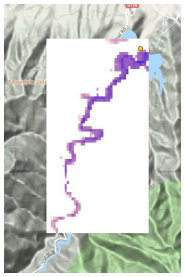


(B) Shape of water mask

FIGURE A.25: Thaphanseik, Myanmar

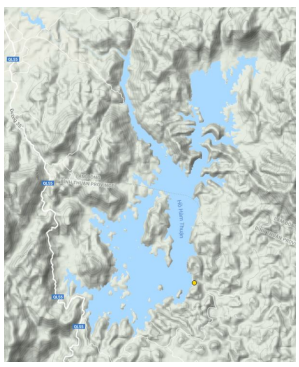


(A) Shape of the actual reservoir

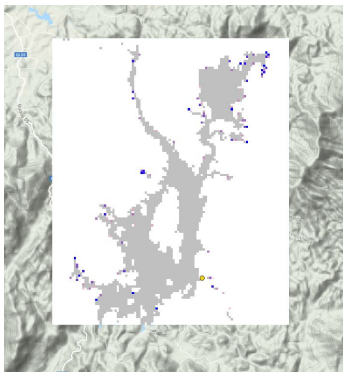


(B) Shape of water mask

FIGURE A.26: Jinpen, China

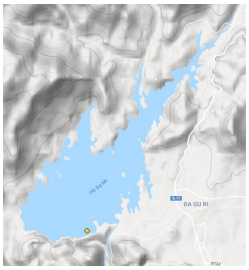


(A) Shape of the actual reservoir

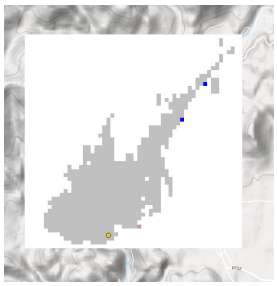


(B) Shape of water mask

FIGURE A.27: Ham Thuan 1, Vietnam



(A) Shape of the actual reservoir



(B) Shape of water mask

FIGURE A.28: Da Mi 1, Vietnam

A.2 RMSE

A.2.1 Basic Geometric Model

RMSE Analysis for Reservoirs			
ID	Reservoir Name	Country	RMSE
3	Porce 2	Columbia	19.14
4	Peribonka	Canada	18.56
6	Caruachi	Venezuela	11.08
7	Cana Brava	Brazil	12.87
8	Manso	Brazil	10.9
9	Queimado	Brazil	9.03
11	Paraitinga	Brazil	4.23
14	Quebra Quiexo	Brazil	18.01
18	Itapebi	Brazil	21.89
20	Itioz	Spain	17.98
21	Alqueva	Portugal	14.85
22	Rules	Spain	20.32
24	Kozjak	Macedonia	13.82
25	Capanda	Angola	15.28
26	Mohale	Lesotho	19.95
28	Maguga	Swaziland	15.48
29	Karkheh	Iran	17.09
30	Masjed Solayman	Iran	50.27
31	Karoon 3	Iran	20.13
32	Sardar Sarovar	India	13.09
34	Bansagar	India	27.67
35	Bennithora	India	5.64
36	Nina	China	10.13
37	Khuga	India	3.40
38	Thaphanseik	Myanmar	7.47
40	Jinpen	China	13.92
45	Ham Thuan 1	Vietnam	11.03
46	Da Mi 1	Vietnam	10.79

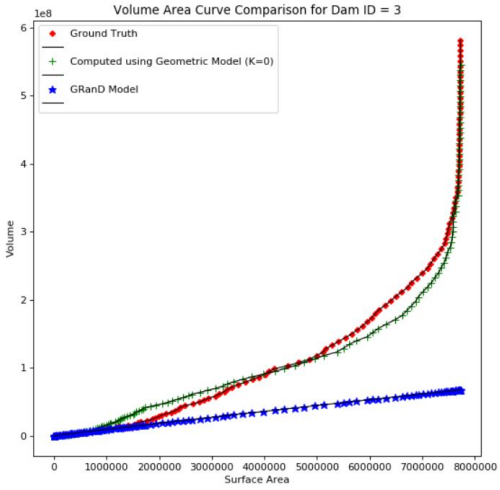
A.2.2 Geometric Model - Optimal Kappa

RMSE Analysis for Reservoirs				
ID	Reservoir Name	Country	Optimal Kappa	Corresponding RMSE
3	Porce 2	Columbia	0.00086	18.8
4	Peribonka	Canada	0	18.56
6	Caruachi	Venezuela	0.00019	10.35
7	Cana Brava	Brazil	0.0008	11.42
8	Manso	Brazil	0.000055	9.13
9	Queimado	Brazil	0.0017	8.01
11	Paraitinga	Brazil	0	4.23
14	Quebra Quiexo	Brazil	0.016	10.05
18	Itapebi	Brazil	0.073	17.96
20	Itioz	Spain	0.0015	17.33
21	Alqueva	Portugal	0.0044	12.57
22	Rules	Spain	0.135	9.52
24	Kozjak	Macedonia	0.005	13.74
25	Capanda	Angola	0.00014	13.42
26	Mohale	Lesotho	0.0035	17.87
28	Maguga	Swaziland	0.0018	13.99
29	Karkheh	Iran	0	17.09
30	Masjed Solayman	Iran	0	53.38
31	Karoon 3	Iran	0	20.13
32	Sardar Sarovar	India	0.008	12.78
34	Bansagar	India	0	27.67
35	Bennithora	India	0.000475	3.82
36	Nina	China	0	10.13
37	Khuga	India	0	3.41
38	Thaphanseik	Myanmar	0	7.47
40	Jinpen	China	0.02	13.74
45	Ham Thuan 1	Vietnam	0.0004	10.95
46	Da Mi 1	Vietnam	0.0025	8.31

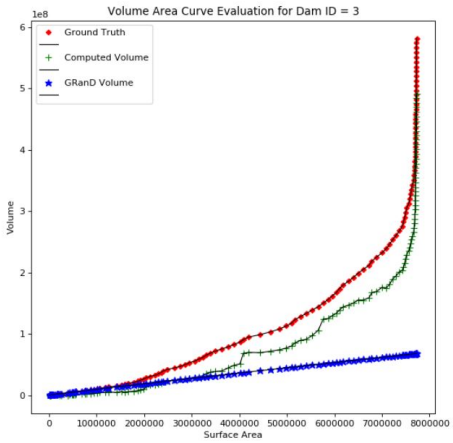
A.3 Maximum Storage Capacities

ID	Reservoir Name	Country	Surface Area	Computed	Ground Truth	GRanD
			(in m^2)	Volume (in m^3)		
3	Porce 2	Colombia	7715803.343	537991762.8	581703366	67677668.82
4	Peribonka	Canada	24854261.3	5537470378	5460633943	207504093.7
6	Caruachi	Venezuela	290446217	37107322042	37977223964	2185929122
7	Cana Brava	Brazil	99273071.74	9914004499	10403876388	781766215.1
8	Manso	Brazil	325388827.9	43854550464	45275163478	2437198890
9	Queimado	Brazil	17026743.75	417446260.3	514663027.5	144440695.8
11	Paraitinga	Brazil	5047028.882	133824923.4	98964283.95	45069153.91
14	Quebra Queixo	Brazil	5826459.237	176598904.8	244841788.4	51714979.75
18	Itapebi	Brazil	53329549.7	9787778276	10448777180	431123400.4
20	Itoiz Dam	Spain	7978982.252	771976852.7	736671492.3	69887103.29
21	Alqueva	Portugal	169539376.7	23656818762	24960064525	1305290235
22	Rules	Spain	3351000.447	259901809.3	284622167	30445551.67
24	Kozjak	Macedonia	8189202.966	821880172.4	803736964.5	71649730.87
25	Capanda	Angola	161247179.5	21921010603	23650886882	1244078212
26	Mohale	Lesotho	12261889.91	1119956436	1274159240	105470745.4
28	Maguga	Swaziland	7930496.827	370135264.5	433238104.4	69480293.51
29	Karkheh	Iran	99861622.07	8169654133	7262263421	786204852.4
30	Masjed Solay-man	Iran	44131687.66	21031840854	21419048536	359628240.8
31	Karoon 3	Iran	15236540.24	1972228496	1606345822	129861460.8
32	Sardar Sarovar	India	27124529.31	2109399143	1843774740	225624392.9
34	Bansagar	India	483970455.8	67853169334	63394673374	3564767590
35	Bennithora	India	16240501.06	278756038.8	318651243.2	138046008.5
36	Nina	China	2061054.772	106034675.4	85106988.97	19113784.28
37	Khuga	India	6237827.739	150686195.1	172538158.7	55207069.46
38	Thaphanseik	Myanmar (Burma)	376369244.2	28765850498	30727459661	2801785841
40	Jinpen	China	2205357.219	175532989.1	163715760.3	20393692.17
45	Ham Thuan 1	Vietnam	23380857.4	2138439561	2142614029	195706949.8
46	Da Mi 1	Vietnam	5607964.907	295852904.1	294199228	49855999.76

A.4 Volume Area Curves

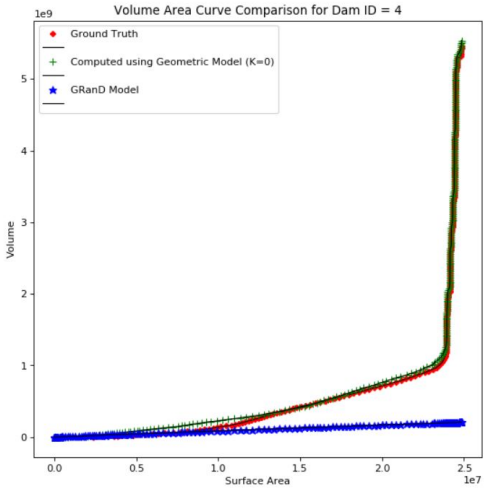


(A) Comparison Curve

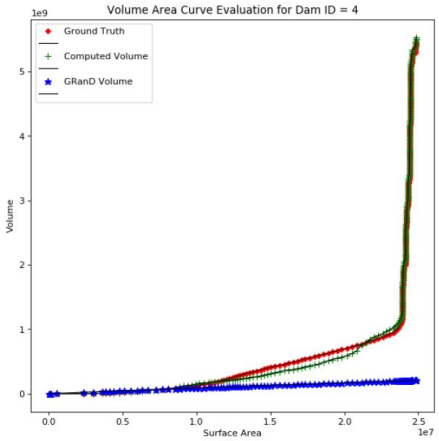


(B) Evaluation Curve

FIGURE A.29: VA Curves 3

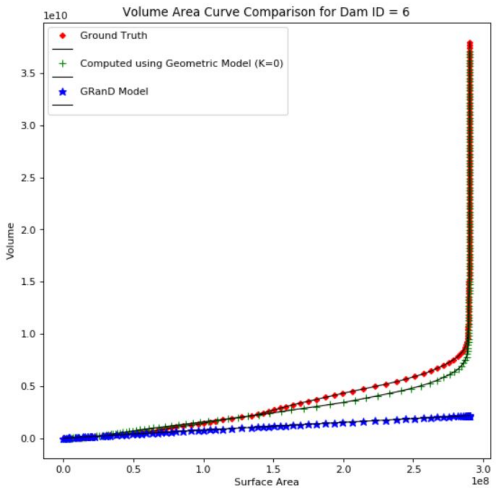


(A) Comparison Curve

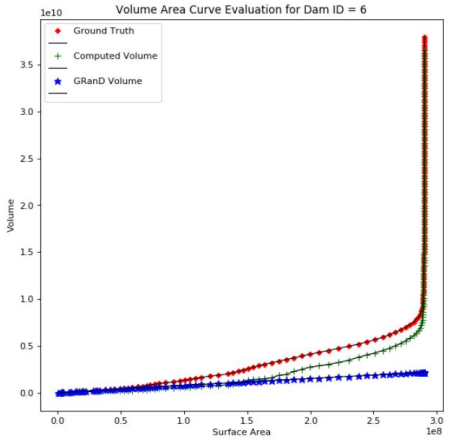


(B) Evaluation Curve

FIGURE A.30: VA Curves 4

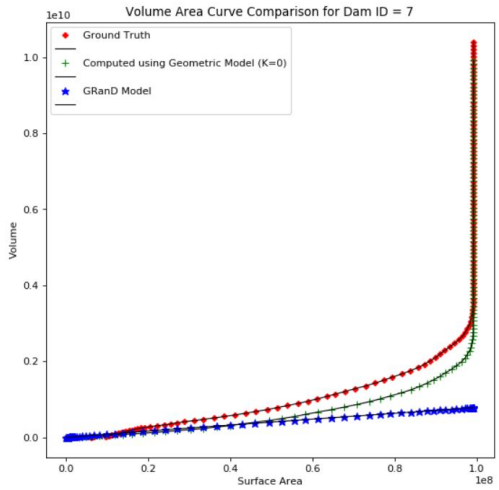


(A) Comparison Curve

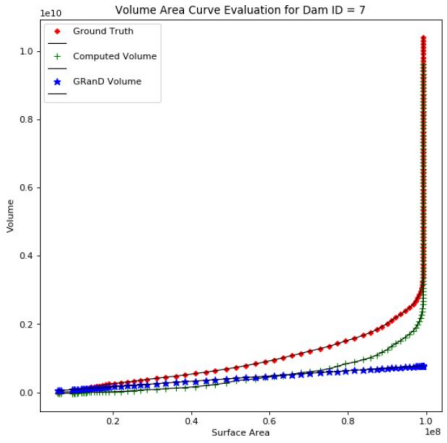


(B) Evaluation Curve

FIGURE A.31: VA Curves 6

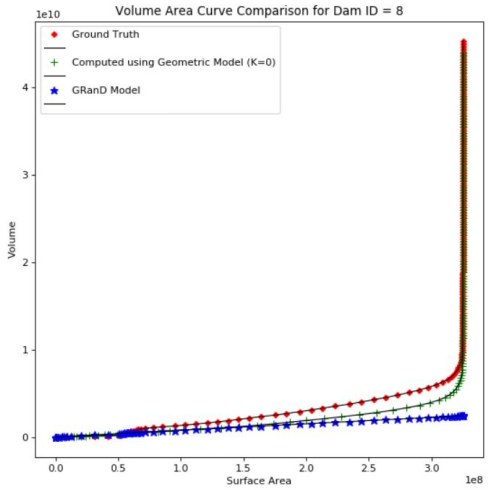


(A) Comparison Curve

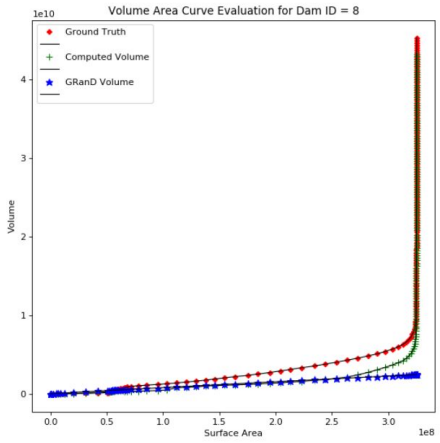


(B) Evaluation Curve

FIGURE A.32: VA Curves 7

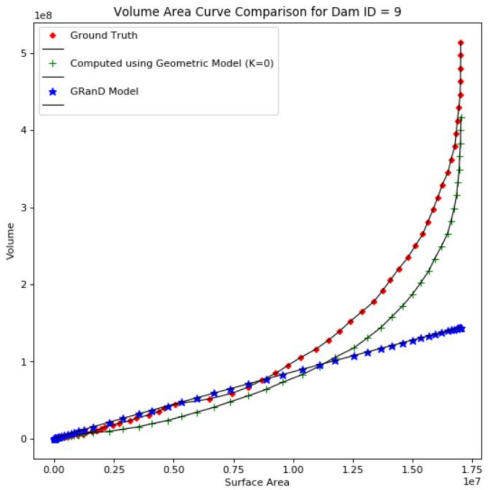


(A) Comparison Curve

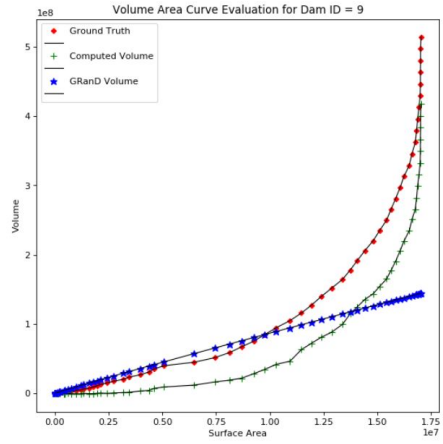


(B) Evaluation Curve

FIGURE A.33: VA Curves 8

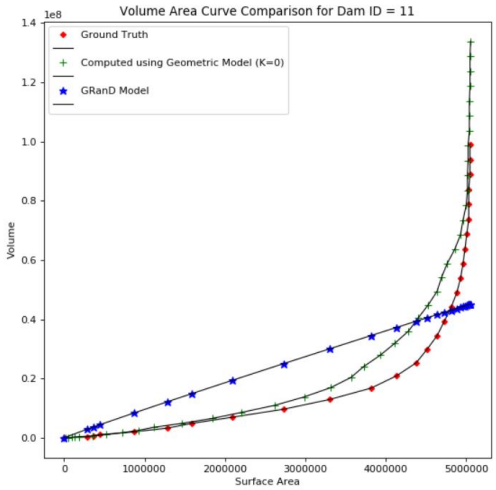


(A) Comparison Curve

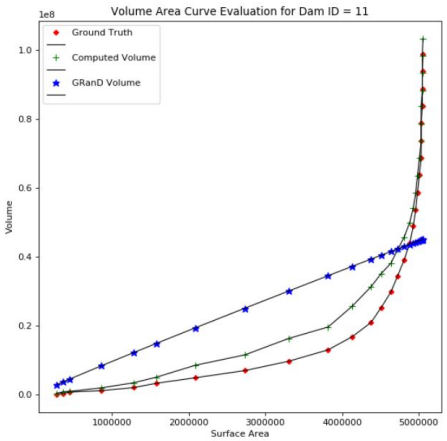


(B) Evaluation Curve

FIGURE A.34: VA Curves 9

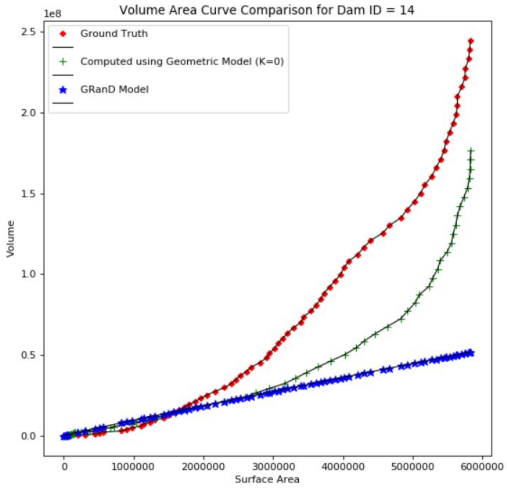


(A) Comparison Curve

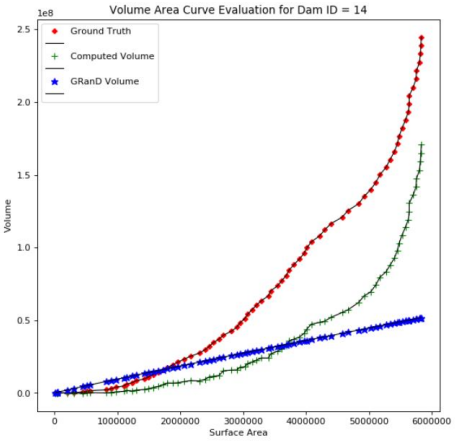


(B) Evaluation Curve

FIGURE A.35: VA Curves 11

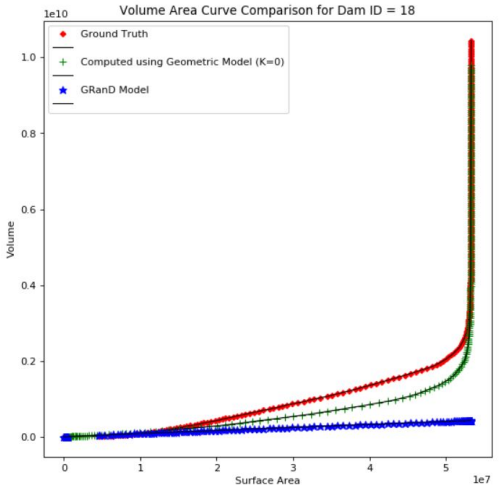


(A) Comparison Curve

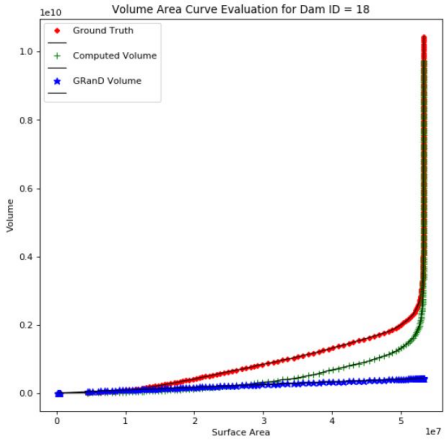


(B) Evaluation Curve

FIGURE A.36: VA Curves 14

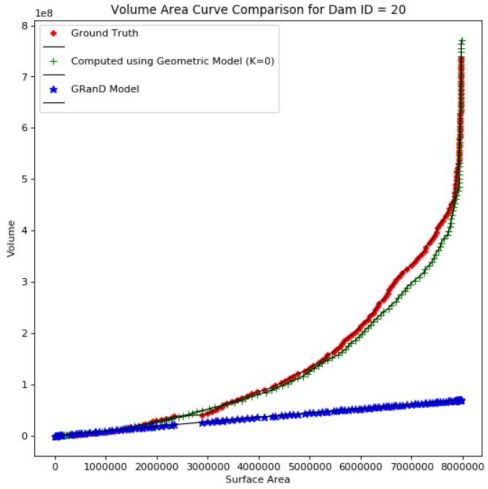


(A) Comparison Curve

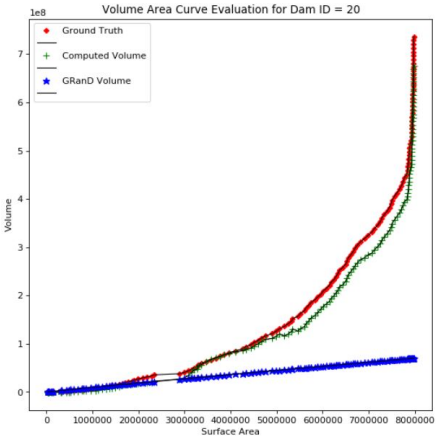


(B) Evaluation Curve

FIGURE A.37: VA Curves 18

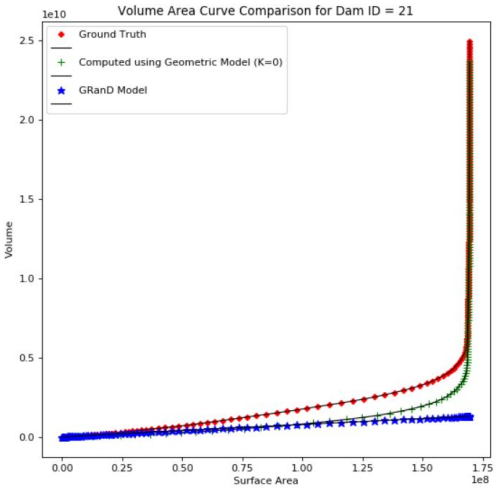


(A) Comparison Curve

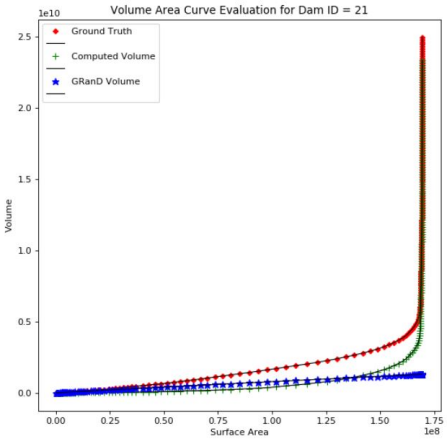


(B) Evaluation Curve

FIGURE A.38: VA Curves 20

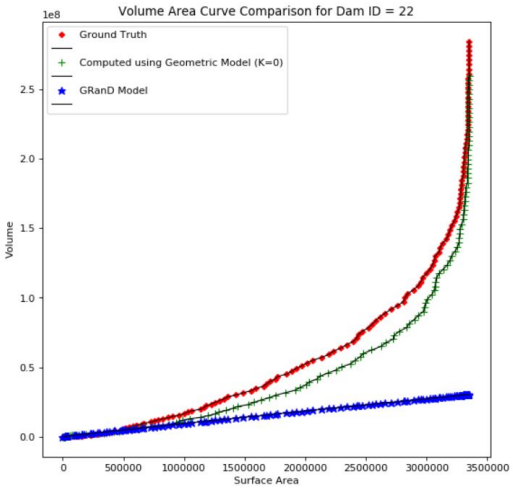


(A) Comparison Curve

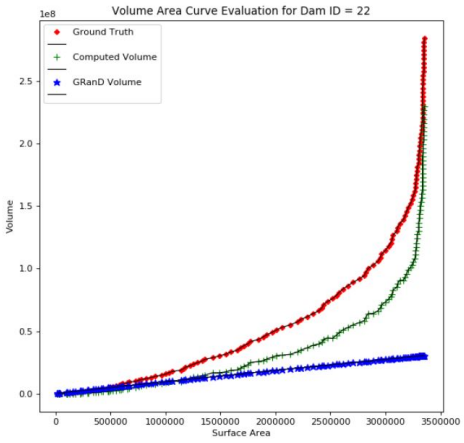


(B) Evaluation Curve

FIGURE A.39: VA Curves 21

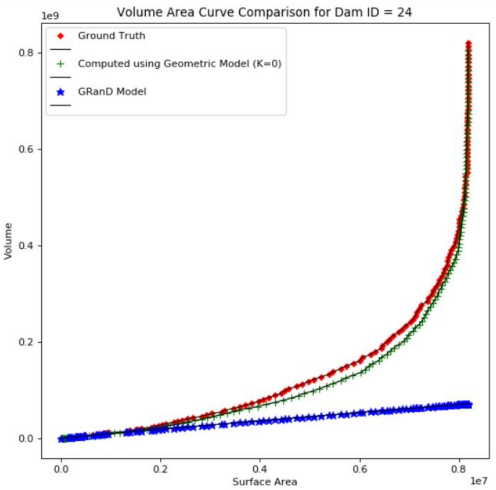


(A) Comparison Curve

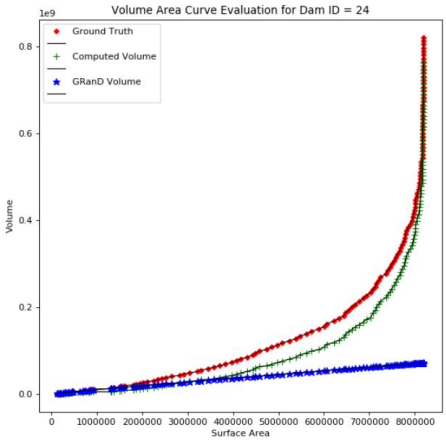


(B) Evaluation Curve

FIGURE A.40: VA Curves 22

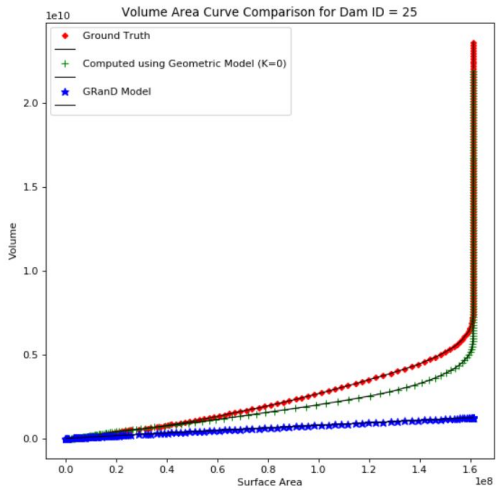


(A) Comparison Curve

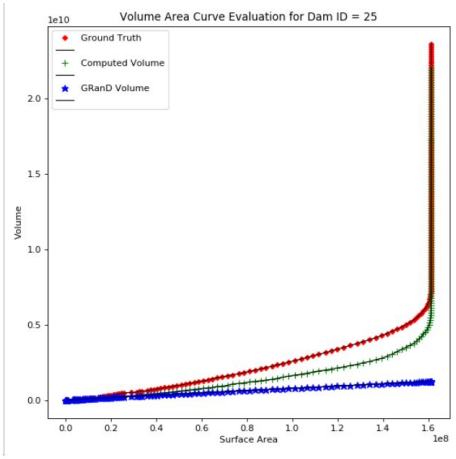


(B) Evaluation Curve

FIGURE A.41: VA Curves 24

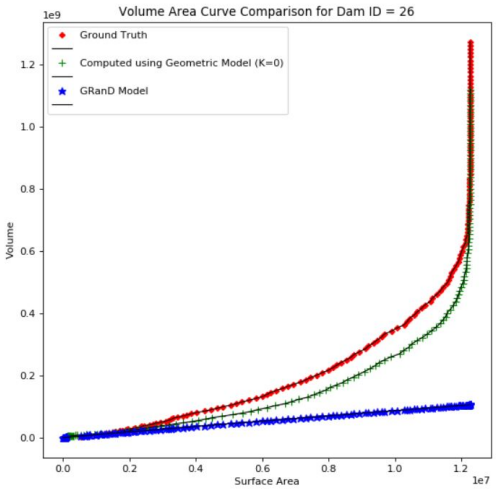


(A) Comparison Curve

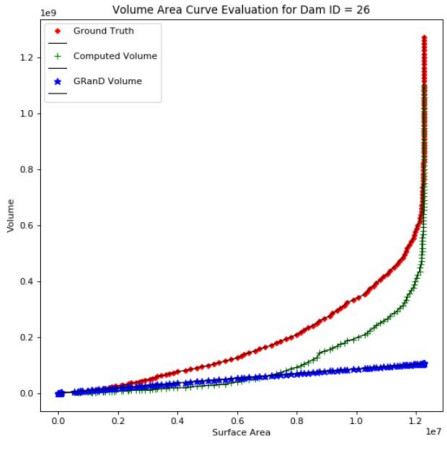


(B) Evaluation Curve

FIGURE A.42: VA Curves 25

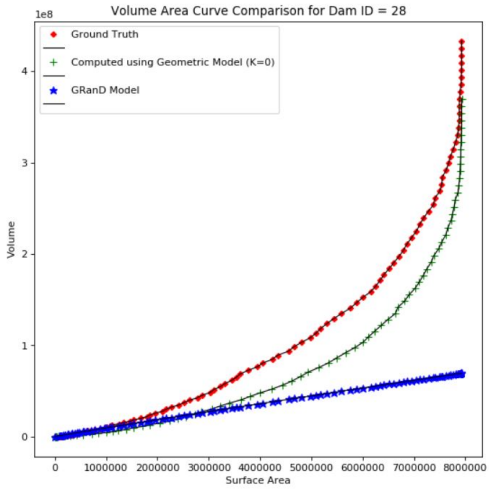


(A) Comparison Curve

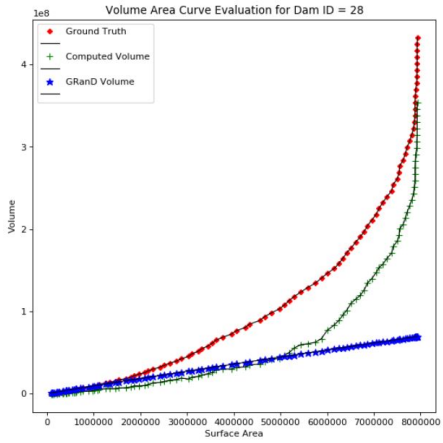


(B) Evaluation Curve

FIGURE A.43: VA Curves 26

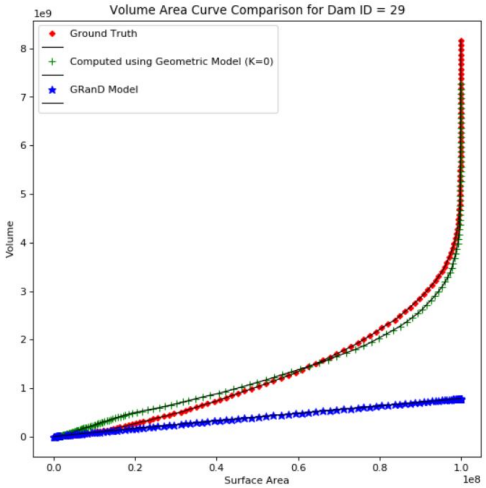


(A) Comparison Curve

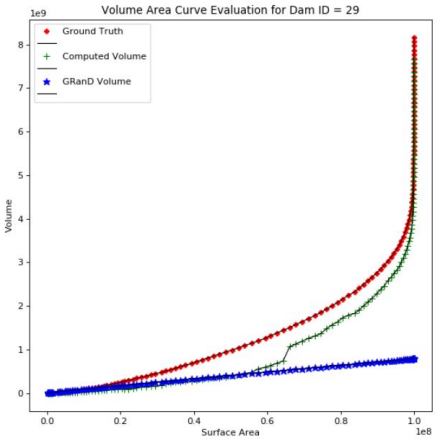


(B) Evaluation Curve

FIGURE A.44: VA Curves 28

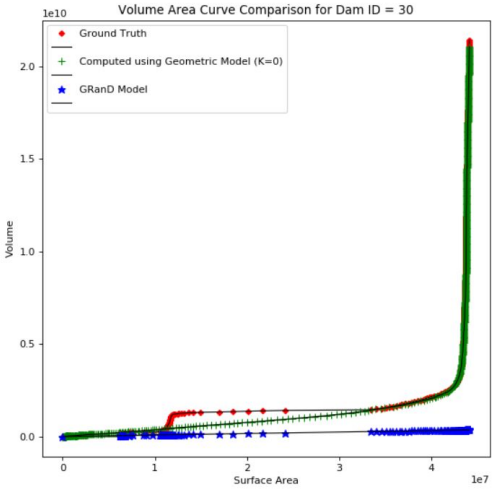


(A) Comparison Curve

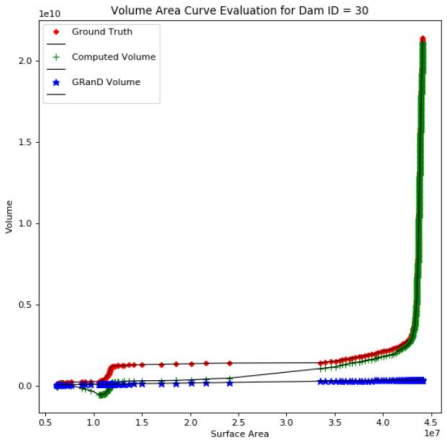


(B) Evaluation Curve

FIGURE A.45: VA Curves 29

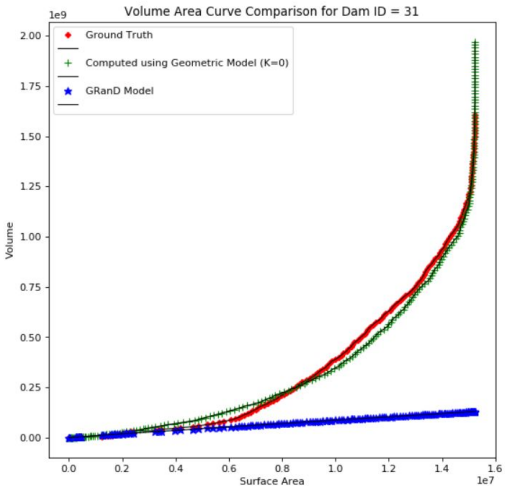


(A) Comparison Curve

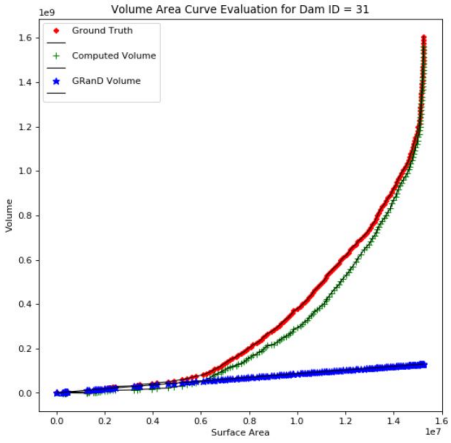


(B) Evaluation Curve

FIGURE A.46: VA Curves 30

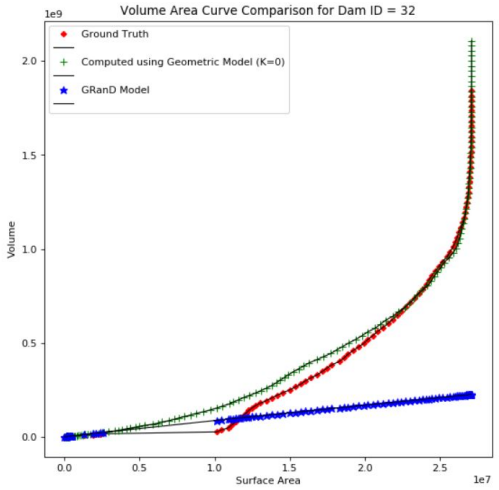


(A) Comparison Curve

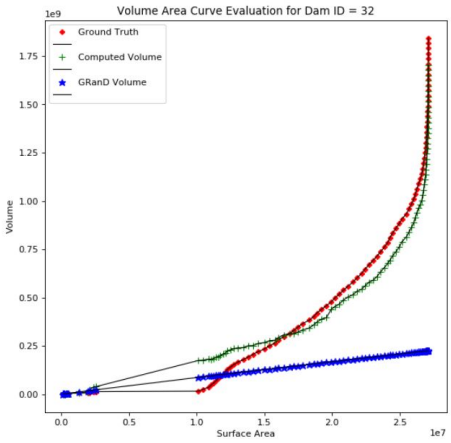


(B) Evaluation Curve

FIGURE A.47: VA Curves 31

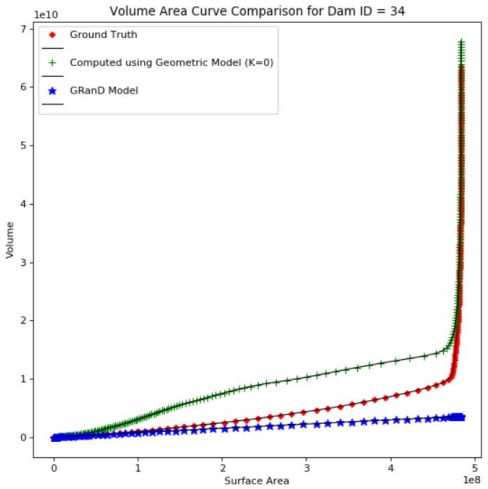


(A) Comparison Curve

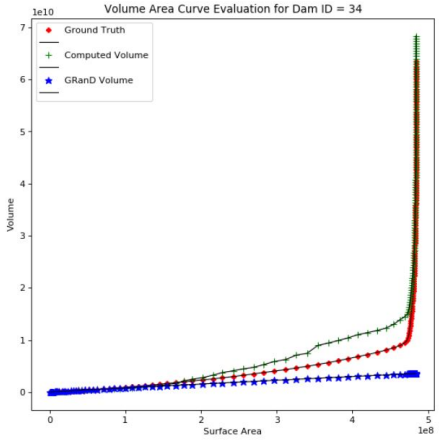


(B) Evaluation Curve

FIGURE A.48: VA Curves 32

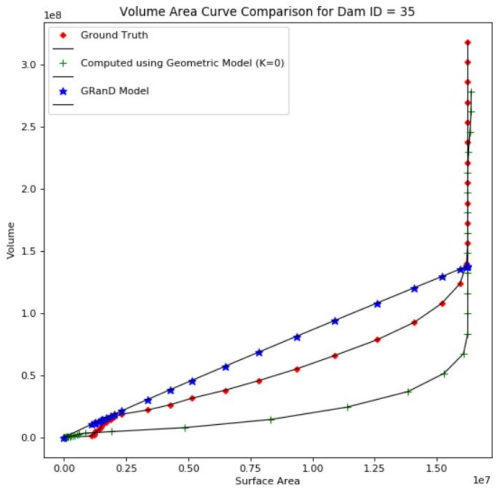


(A) Comparison Curve

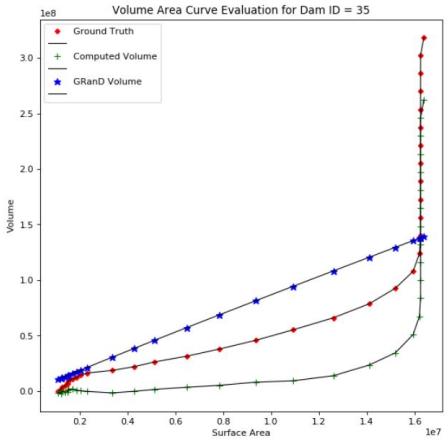


(B) Evaluation Curve

FIGURE A.49: VA Curves 34

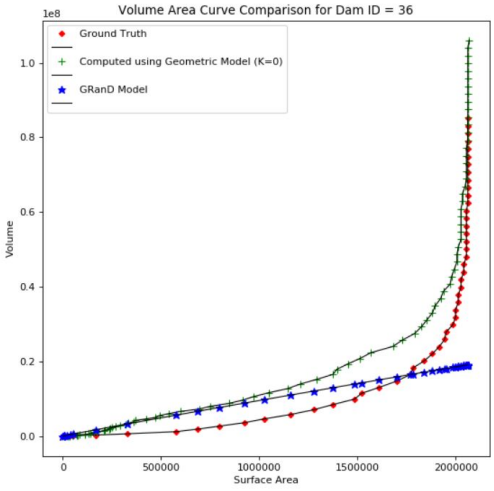


(A) Comparison Curve

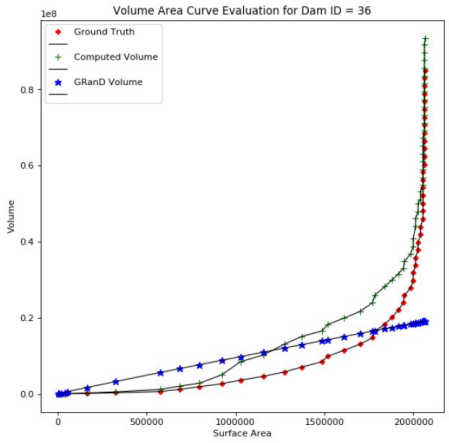


(B) Evaluation Curve

FIGURE A.50: VA Curves 35

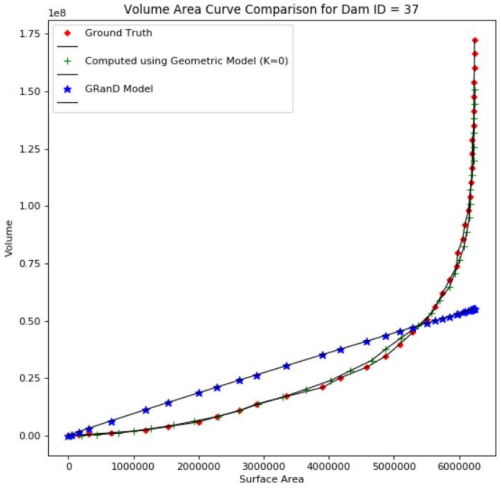


(A) Comparison Curve

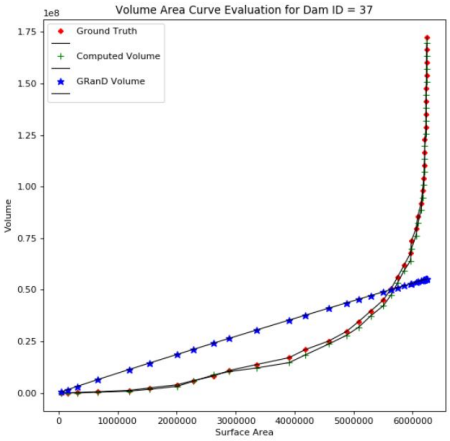


(B) Evaluation Curve

FIGURE A.51: VA Curves 36

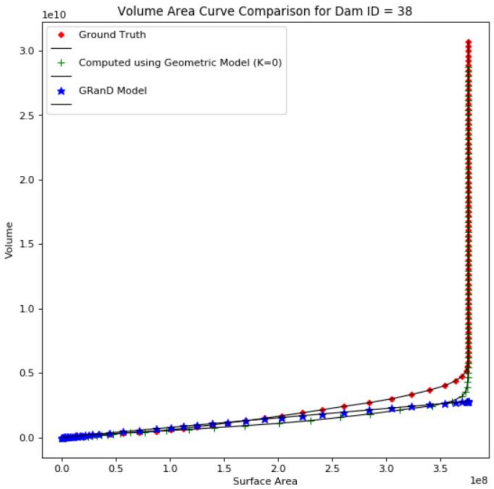


(A) Comparison Curve

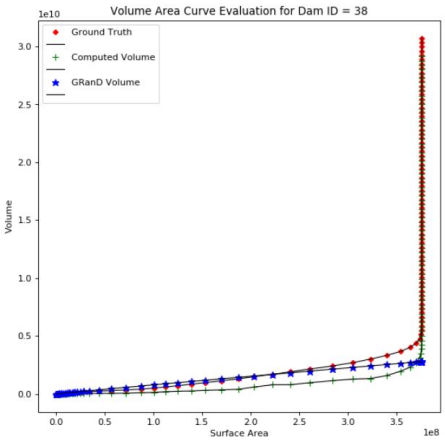


(B) Evaluation Curve

FIGURE A.52: VA Curves 37

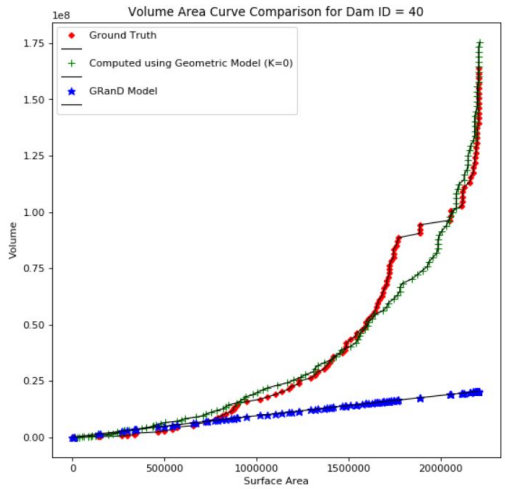


(A) Comparison Curve

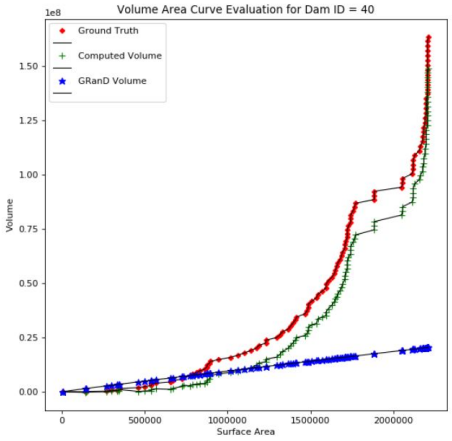


(B) Evaluation Curve

FIGURE A.53: VA Curves 38

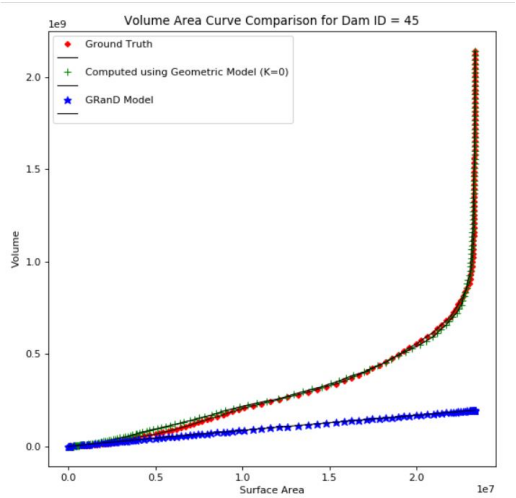


(A) Comparison Curve

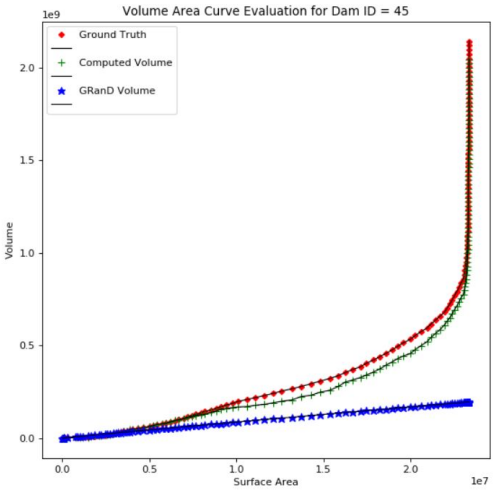


(B) Evaluation Curve

FIGURE A.54: VA Curves 40

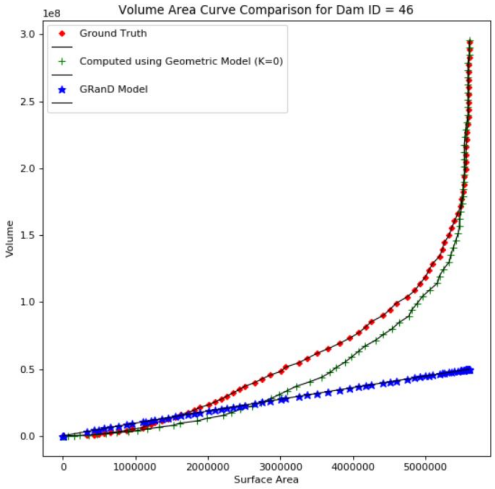


(A) Comparison Curve

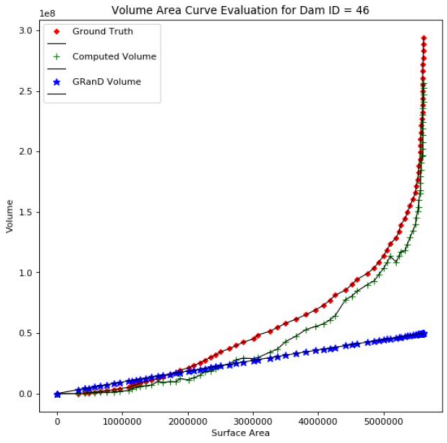


(B) Evaluation Curve

FIGURE A.55: VA Curves 45



(A) Comparison Curve



(B) Evaluation Curve

FIGURE A.56: VA Curves 46

Bibliography

Babić, Darko et al. (2002). "Resistance-distance matrix: A computational algorithm and its application". In: *International Journal of Quantum Chemistry* 90.1, pp. 166–176.

Basson, Marthinus S, Peter H Van Niekerk, and Johan A Van Rooyen (1997). *Overview of water resources availability and utilisation in South Africa*. Department of Water Affairs and Forestry.

Beck, Hylke E et al. (2016). "Global-scale regionalization of hydrologic model parameters". In: *Water Resources Research* 52.5, pp. 3599–3622.

Bemmelen, CWT et al. (2016). "Determining water reservoir characteristics with global elevation data". In: *Geophysical Research Letters* 43.21.

Bezerra, Hedlena et al. (2010). "Diffusion constraints for vector graphics". In: *Proceedings of the 8th International Symposium on Non-Photorealistic Animation and Rendering*. ACM, pp. 35–42.

Bhat, Pravin et al. (2010). "Gradientshop: A gradient-domain optimization framework for image and video filtering". In: *ACM Transactions on Graphics (TOG)* 29.2, p. 10.

Bhatia, Ramesh et al. (2008). *Indirect economic impacts of dams: Case studies from India, Egypt, and Brazil*. Academic Foundation.

Bierkens, Marc FP et al. (2015). "Hyper-resolution global hydrological modelling: what is next? "Everywhere and locally relevant"". In: *Hydrological processes* 29.2, pp. 310–320.

Chao, Benjamin F, YH Wu, and YS Li (2008). "Impact of artificial reservoir water impoundment on global sea level". In: *science* 320.5873, pp. 212–214.

Coerver, Hubertus M, Martine M Rutten, and Nick C van de Giesen (2018). "Deduction of reservoir operating rules for application in global hydrological models." In: *Hydrology & Earth System Sciences* 22.1.

Crane, Keenan, Clarisse Weischedel, and Max Wardetzky (2013). "Geodesics in heat: A new approach to computing distance based on heat flow". In: *ACM Transactions on Graphics (TOG)* 32.5, p. 152.

Cvetković, Dragos M, Michael Doob, and H Sachs (1998). *Spectra of Graphs: Theory and Applications, 3rd rev. enl. ed*.

Danielson, Jeffrey J and Dean B Gesch (2011). *Global multi-resolution terrain elevation data 2010 (GMTED2010)*. Tech. rep. US Geological Survey.

Davis, Timothy A, Sivasankaran Rajamanickam, and Wissam M Sid-Lakhdar (2016). "A survey of direct methods for sparse linear systems". In: *Acta Numerica* 25, pp. 383–566.

Dean, Walter E and Eville Gorham (1998). "Magnitude and significance of carbon burial in lakes, reservoirs, and peatlands". In: *Geology* 26.6, pp. 535–538.

Dijk, Albert IJM van et al. (2014). "A global water cycle reanalysis (2003–2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble". In: *Hydrology and Earth System Sciences* 18.8, pp. 2955–2973.

Döll, Petra, Kristina Fiedler, and Jing Zhang (2009). "Global-scale analysis of river flow alterations due to water withdrawals and reservoirs". In: *Hydrology and Earth System Sciences* 13.12, pp. 2413–2432.

Dorcey, Tony et al. (1997). *Large dams: Learning from the past, looking at the future: workshop proceedings, Gland, Switzerland, April 11-12, 1997*. The World Bank.

Downing, John A et al. (2006). "The global abundance and size distribution of lakes, ponds, and impoundments". In: *Limnology and Oceanography* 51.5, pp. 2388–2397.

Duff, Iain S (2004). "MA57—a code for the solution of sparse symmetric definite and indefinite systems". In: *ACM Transactions on Mathematical Software (TOMS)* 30.2, pp. 118–144.

Eilander, Dirk et al. (2014). "Remotely sensed monitoring of small reservoir dynamics: A Bayesian approach". In: *Remote Sensing* 6.2, pp. 1191–1210.

Farbman, Zeev et al. (2008). "Edge-preserving decompositions for multi-scale tone and detail manipulation". In: *ACM Transactions on Graphics (TOG)*. Vol. 27. 3. ACM, p. 67.

Farr, Tom G et al. (2007). "The shuttle radar topography mission". In: *Reviews of geophysics* 45.2.

Fattal, Raanan, Dani Lischinski, and Michael Werman (2002). "Gradient domain high dynamic range compression". In: *ACM transactions on graphics (TOG)*. Vol. 21. 3. ACM, pp. 249–256.

Fiedler, Julia W and Clinton P Conrad (2010). "Spatial variability of sea level rise due to water impoundment behind dams". In: *Geophysical Research Letters* 37.12.

Gamache, Martin (2004). "Free and low cost datasets for international mountain cartography". In: *4th ICA Mountain Cartography Workshop*. Vol. 26.

Haddeland, Ingjerd, Thomas Skaugen, and Dennis P Lettenmaier (2006). "Anthropogenic impacts on continental surface water fluxes". In: *Geophysical Research Letters* 33.8.

Hanasaki, Naota, Shinjiro Kanae, and Taikan Oki (2006). "A reservoir operation scheme for global river routing models". In: *Journal of Hydrology* 327.1-2, pp. 22–41.

Jackson, Sukhan and Adrian Sleigh (2000). "Resettlement for China's Three Gorges Dam: socio-economic impact and institutional tensions". In: *Communist and Post-Communist Studies* 33.2, pp. 223–241.

Ketteringham, W. *Western Cape Water Supply System Reconciliation Strategy Study, Overview of Water Conservation and Demand Management in the City of Cape Town. Prepared for the DWAF by UWP Consulting (Pty) Ltd. DWAF Report No. P WMA 19*. Tech. rep. 000/00/0507.

Lehner, Bernhard and Petra Döll (2004). "Development and validation of a global database of lakes, reservoirs and wetlands". In: *Journal of Hydrology* 296.1-4, pp. 1–22.

Lehner, Bernhard, Kris Verdin, and Andy Jarvis (2006). "HydroSHEDS technical documentation, version 1.0". In: *World Wildlife Fund US, Washington, DC*, pp. 1–27.

Lehner, Bernhard, Kristine Verdin, and Andy Jarvis (2008). "New global hydrography derived from spaceborne elevation data". In: *Eos, Transactions American Geophysical Union* 89.10, pp. 93–94.

Lehner, Bernhard et al. (2011a). "Global reservoir and dam (grand) database". In: *Technical Documentation, Version* 1.

— (2011b). "High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management". In: *Frontiers in Ecology and the Environment* 9.9, pp. 494–502.

Levin, Anat, Dani Lischinski, and Yair Weiss (2004). "Colorization using optimization". In: *ACM transactions on graphics (tog)*. Vol. 23. 3. ACM, pp. 689–694.

Liebe, Jens R et al. (2009). "Suitability and limitations of ENVISAT ASAR for monitoring small reservoirs in a semiarid area". In: *IEEE Transactions on Geoscience and Remote Sensing* 47.5, pp. 1536–1547.

Lischinski, Dani et al. (2006). "Interactive local adjustment of tonal values". In: *ACM Transactions on Graphics (TOG)*. Vol. 25. 3. ACM, pp. 646–653.

Liu, Rong and Hao Zhang (2007). "Mesh segmentation via spectral embedding and contour analysis". In: *Computer Graphics Forum*. Vol. 26. 3. Wiley Online Library, pp. 385–394.

McCully, Patrick et al. (1996). *Silenced rivers: The ecology and politics of large dams.* Zed Books.

Messager, Mathis Loïc et al. (2016). "Estimating the volume and age of water stored in global lakes using a geo-statistical approach". In: *Nature communications* 7, p. 13603.

Nazemi, Ali and Howard S Wheater (2015). "On inclusion of water resource management in Earth system models–Part 2: Representation of water supply and allocation and opportunities for improved modeling". In: *Hydrology and Earth System Sciences* 19.1, pp. 63–90.

Orth, Rene and Sonia I Seneviratne (2013). "Predictability of soil moisture and streamflow on subseasonal timescales: A case study". In: *Journal of Geophysical Research: Atmospheres* 118.19, pp. 10–963.

Orzan, Alexandrina et al. (2008). "Diffusion curves: a vector representation for smooth-shaded images". In: *ACM Transactions on Graphics (TOG)*. Vol. 27. 3. ACM, p. 92.

Pérez, Patrick, Michel Gangnet, and Andrew Blake (2003). "Poisson image editing". In: *ACM Transactions on graphics (TOG)* 22.3, pp. 313–318.

Perona, Pietro and Jitendra Malik (1990). "Scale-space and edge detection using anisotropic diffusion". In: *IEEE Transactions on pattern analysis and machine intelligence* 12.7, pp. 629–639.

Pringle, Catherine M (2001). "Hydrologic connectivity and the management of biological reserves: a global perspective". In: *Ecological Applications* 11.4, pp. 981–998.

Robinson, Natalie, James Regetz, and Robert P Guralnick (2014). "EarthEnv-DEM90: A nearly-global, void-free, multi-scale smoothed, 90m digital elevation model from fused ASTER and SRTM data". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 87, pp. 57–67.

Schumm, Stanley A and Robert W Lichty (1965). "Time, space, and causality in geomorphology". In: *American journal of science* 263.2, pp. 110–119.

Shi, Jianbo and Jitendra Malik (2000). "Normalized cuts and image segmentation". In: *IEEE Transactions on pattern analysis and machine intelligence* 22.8, pp. 888–905.

Smakhtin, Vladimir et al. (2001). "Unconventional water supply options in South Africa: a review of possible solutions". In: *Water International* 26.3, pp. 314–334.

Sood, Aditya and Vladimir Smakhtin (2015). "Global hydrological models: a review". In: *Hydrological Sciences Journal* 60.4, pp. 549–565.

St. Louis, Vincent L et al. (2000). "Reservoir Surfaces as Sources of Greenhouse Gases to the Atmosphere: A Global Estimate: Reservoirs are sources of greenhouse gases to the atmosphere, and their surface areas have increased to the point where they should be included in global inventories of anthropogenic emissions of greenhouse gases". In: *AIBS Bulletin* 50.9, pp. 766–775.

Strzepek, Kenneth M et al. (2008). "The value of the high Aswan Dam to the Egyptian economy". In: *Ecological Economics* 66.1, pp. 117–126.

Sun, Jian et al. (2004). "Poisson matting". In: *ACM Transactions on Graphics (ToG)*. Vol. 23. 3. ACM, pp. 315–321.

Syvitski, James PM et al. (2005). "Impact of humans on the flux of terrestrial sediment to the global coastal ocean". In: *science* 308.5720, pp. 376–380.

Tachikawa, Tetsushi et al. (2011). "Characteristics of ASTER GDEM version 2". In: *Geoscience and remote sensing symposium (IGARSS), 2011 IEEE international*. IEEE, pp. 3657–3660.

Tadono, T et al. (2014). "Precise global DEM generation by ALOS PRISM". In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 2.4, p. 71.

Takeuchi, K (1997). "Least marginal environmental impact rule for reservoir development". In: *Hydrological sciences journal* 42.4, pp. 583–597.

Tang, Bei, Guillermo Sapiro, and Vicent Caselles (2001). "Color image enhancement via chromaticity diffusion". In: *IEEE Transactions on Image Processing* 10.5, pp. 701–707.

Tang, Qiuhong et al. (2009). "Remote sensing: hydrology". In: *Progress in Physical Geography* 33.4, pp. 490–509.

Weisstein, Eric W. *Adjacency Matrix. From MathWorld–A Wolfram Web Resource*. Visited on 1/08/18. URL: http://mathworld.wolfram.com/AdjacencyMatrix.html.

— *Degree Matrix. From MathWorld–A Wolfram Web Resource*. Visited on 1/08/18. URL: http://mathworld.wolfram.com/DegreeMatrix.html.

Wisser, Dominik et al. (2010). "Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network-Hydrology (GTN-H)". In: *Hydrology and Earth System Sciences* 14.1, pp. 1–24.

Wood, Eric F et al. (2011). "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water". In: *Water Resources Research* 47.5.